DEVELOPMENT OF A SELF-EFFICACY SCALE FOR ASSESSING SECONDARY SCHOOL STUDENTS' SCIENCE SELF-EFFICACY BELIEFS

By

Mary Wong Siew Lian

ABSTRACT

This paper details the development of a self-efficacy scale for assessing secondary school students' science self-efficacy (SSE) beliefs. Differences in the way self-efficacy and self-concept are assessed is discussed in light of the conceptual difference between self-efficacy and other self-percepts. Steps in the formulation of the scale in accordance with theoretical guidelines on how self-efficacy should be assessed and measures of instrument validity and reliability are presented. In conclusion, the suitability of the scale developed is discussed.

INTRODUCTION

Bandura (1986, p. 391) defined self-efficacy beliefs as referring to "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances". Self-efficacy beliefs influence people's thoughts and actions; how much effort they will expend and how long they will persevere in the face of obstacles. According to the Social Cognitive theory (Bandura, 1986), self-efficacy is one of the critical components of the personal processes influencing learning. Over the past one and a half decades, there has been much debate over the measurement and mismeasurement of self-efficacy due to misconceptions concerning what self-efficacy beliefs refer to and how it should be assessed (Pajares, 1996a). In the following paragraphs, the conceptual difference between self-efficacy and other self-percepts are discussed.

THE CONCEPTUAL DIFFERENCE BETWEEN SELF-EFFICACY AND OTHER SELF-PERCEPTS

The conceptual difference between self-efficacy and other expectancy beliefs such as self-concept and self-esteem is often wrongly assumed to be the same. Pajares (1997) noted that some researchers use the two terms, self-efficacy and self-concept, synonymously while others described self-concept as a generalized form of self-efficacy. There have also been others who defined academic self-concept as self-perceptions of ability and suggested that one reason why these self-percepts affect performance is because of their effect on students' effort, persistence, and anxiety.

According to Bandura (1986), self-efficacy refers to judgment of one's capability to accomplish a certain task while self-concept is a composite view of oneself that is formed through direct experiences and evaluations adopted from significant others. Woolfolk (1998) explained that self-concept refers to the composite of ideas, feelings and attitudes people have about themselves. This is developed through personal and social comparisons in different situations. Pajares (1996b) explained that self-efficacy and self-concept are similar in that they are self-percepts or beliefs about one's perceived capability. However, they differ in that self-efficacy is defined in more specific terms than self-concept. Self-efficacy differs from self-concept in that the former is a context-specific assessment of competence to perform a specific task whereas the latter is measured at a broader level of specificity. A typical self-concept item, "Mathematics makes me feel inadequate," differs markedly from a self-efficacy question that may begin with "How confident are you that you can answer this question correctly?" Self-concept judgments, therefore, can be domain-specific but are not task-specific, while self-efficacy judgments are both task- and domain-specific. According to Marsh, Walker, and Debus (1991), another distinction between the two constructs is the source of an individual's judgment. Self-concept judgments are based on social and self-comparisons, or "frame of reference effects". For example, by comparing one’s own performance with those of others (“I am better at Science than most of my classmates”) and also by comparing one’s own performance in related areas (“I am better at Science than English”) an individual develops a judgment of self-concept. Self-efficacy judgments, on the other hand, focus on the specific ability to accomplish the criterial task.

Self-esteem is another form of self-referent thought that should be distinguished from perceived self-efficacy as the two concepts represent quite different phenomena. According to Bandura (1986, p. 410), “self-esteem pertains to the evaluation of self-worth, which depends on how the culture values the attitudes one possesses and how well one's behavior matches personal standards of worthiness.” Perceived self-efficacy, however, is concerned with the judgment of personal capabilities. Judgments of self-worth (i.e., self-esteem) and self-capabilities (i.e., self-efficacy) have no uniform relation. According to Stajkovic and Luthans (1998), another difference is that self-esteem is often portrayed as a global construct that represents a person's self-evaluations across a wide variety of different situations. In contrast, self-efficacy is the individual's conviction about a task- and context-specific capability. Self-esteem also tends to be more stable, an almost trait-like variable, whereas self-efficacy is a dynamic construct that changes over time as new information and task experiences are obtained. Woolfolk (1998, p. 73) defined self-esteem as “the value each of us places on our own characteristics, abilities and behaviors” and stressed that it is different from self-efficacy in that self-efficacy refers to “a person’s sense of being able to deal effectively with a particular task.”

ASSESSMENT OF SELF-EFFICACY BELIEFS

Self-efficacy beliefs are usually assessed by asking individuals to report the level, generality, and strength of their confidence to accomplish a task or succeed in a certain situation. According to Bandura (1986), level refers to the level of task difficulty that a person believes he or she is capable of executing (e.g., from simple to moderately difficult tasks). The second dimension, generality, refers to the extent to which self-efficacy beliefs relate to, or transfer across a wide range of tasks or domains. Strength of self-efficacy refers to whether judgment about the magnitude or level of difficulty is strong (likely to persevere in the face of obstacles) or weak (easily give up). The stronger the sense of personal efficacy, the greater the perseverance and the higher the likelihood that the activity will be performed successfully.

Bandura (1986, 1997), however, has cautioned researchers attempting to predict academic outcomes from students' self-efficacy beliefs that to increase accuracy of prediction, they would be well advised to follow theoretical guidelines regarding specificity and correspondence of self-efficacy assessment. Task-specific judgments of self-efficacy are superior to both domain-specific assessments and omnibus measures of general self-efficacy. This is because task-specific measures require students to generate judgments with a clear academic activity or task in mind. Generalized self-efficacy instruments basically assess people's general belief that they can make things happen without specifying what these things are. As Pajares (1996a) observed, when students do not know with any degree of accuracy what it is they are expected to do, the judgments on which they will base their capability to do it will be nebulous at best.

In academic settings, self-efficacy instruments usually ask students to rate their confidence in solving specific mathematics problems (Pajares & Miller, 1994), or perform particular reading or writing tasks (Zimmerman & Bandura, 1994), or engage in certain self-regulatory skills (Zimmerman, Bandura & Martinez-Pons, 1992). The individual is required to rate his or her confidence (level of self-efficacy) in performing a specific task on a scale ranging from 1 or 10 (uncertain) to 10 or 100 (certain) at 1 or 10-point intervals. The sum of the confidence ratings for the whole instrument is the strength of self-efficacy of the individual.

For example, in assessing students' mathematics problem-solving self-efficacy, students are required to respond, on a 6-point scale, to the question "How confident are you that you could give the correct answer to the following problems without using a calculator?" [followed by 20 algebra or geometry problems…] (Pajares & Miller, 1994), or students are asked to rate, on a 10-point scale, their self-efficacy for writing skills in answer to the question "How confident are you that you can perform each of the following writing skills?" [8 skills presented…] (Shell, Murphy, & Bruning, 1989). In a study involving 350 undergraduates, Pajares and Miller (1994) found that item specific math self-efficacy beliefs were more predictive of mathematics problem solving than were domain-specific self-concept beliefs.

THE PURPOSE OF THE STUDY

The purpose of this study was to develop a valid and reliable science self-efficacy scale for assessing the science self-efficacy of Form 2, Form 4 Science, and Lower 6 Science students.

METHOD

Sample

Since the present scale was developed for assessing the SSE of students at three Form levels namely Form 2, Form 4 Science, and Lower 6 Science, the pilot sample selected consisted of 111 secondary school students (50 males, 61females) in three intact classes, namely Form 2 (n = 37), Form 4 Science (n = 36) and Lower 6 Science (n = 38). The sample was drawn from a coeducational secondary school in Kuching that was not involved in the actual study.

Formulation of the SSE Scale

The SSE scale that was formulated had 10 science items that covered a wide range of science topics familiar to students in Forms, 2, 4 and Lower 6. The items in the scale represented three science domains (Biology, Physics and Chemistry) and five levels of questioning, namely knowledge, comprehension, application, analysis, and synthesis (Bloom's Taxonomy, cognitive domain). The scale consisted of the type of items that students commonly encounter in learning science, that is, doing calculations, analyzing graphs, diagrams, charts and texts, and designing experiments.

An initial pool of 30 items (i.e., three items for each of the 10 topics selected) was drawn up and subjected to the scrutiny of three experienced secondary school science teachers with 10 to 15 years experience in teaching at least two of the three above-mentioned Form levels of secondary school science. Selection of the 10 items for the scale was based on the teachers' comments and suggestions concerning the suitability of the items for students at the three Form levels. Their comments and suggestions concerning the items in the scale were taken into account and appropriate amendments were made on the scale. Following this step at establishing content validity, three experienced university lecturers further checked the scale for face validity.

In keeping with how academic self-efficacy is assessed (Pajares, 1996a; Zimmerman & Bandura, 1994; Zimmerman & Martinez-Pons, 1990), students were not required to write down the answers to the items in the scale. They were only required to give realistic estimates of their confidence in answering such science items correctly. The approximate time allotted for students to attend to each item was about 30 seconds only, so that they would not have the time to write down the answers to the items. They were asked to rate their confidence in answering each science item correctly on a 10-point scale. Sample items are shown in Appendix A. The scale, which was initially formulated in English, was translated into Bahasa Malaysia and the three-step back-translation procedure (Brislin, 1986) was employed to check on the accuracy of the translation. The resultant SSE scale was next pilot tested on a representative sample of students.

Pilot Testing the Scale

Before administration of the scale, the researcher took time to explain carefully to the students what they were required to do, as students were not used to this form of response to science items. Initially, students were given practice with efficacy assessment by judging their certainty of being able to jump progressively longer distances ranging from one meter to several meters (adapted from Schunk, 1983). The aim of using this concrete example was to help students learn how to use the scale's numerical values to convey the strength of their perceived efficacy.

Following this practice session, students were presented with the 10 items in the SSE Scale. They were asked to assume that they would be presented with such items in a coming science test. They were asked to rate their confidence in answering each science item correctly on a 10-point scale. Students were shown the 10 science items for about 30 seconds each and asked to judge their certainty of being able to answer such items correctly. The students were advised to be honest and to mark the efficacy value that corresponded to their level of certainty for being able to answer such items correctly. It was observed that students were able to complete the rating scale in the time given, that is, five minutes. A retest using the same scale was carried out on the same sample of students after two weeks. All the instruments were collected back after the first administration of the scale and the students were not informed that there would be a retest.

FINDINGS AND DISCUSSION

Distribution of SSE Scores

Each student's SSE score was obtained through summing his or her confidence ratings for each of the 10 items in the scale. Table 1 presents the means, standard deviations and minimum and maximum scores for the three classes and for the whole sample. The mean SSE scores shown indicate the suitability of the scale in differentiating between students' SSE at the three Form levels, with Form 2 students' mean SSE (M = 59.24) being less than that of Form 4 Science students (M = 73.50), which, in turn was less than that of Lower 6 Science students (M = 79.71). Table 2 shows the mean scores, standard deviations, and range of scores obtained with the actual sample.

Table 1
Means and Standard Deviations and Range of SSE scores for the Pilot Sample

Class	Means	*SD*	Minimum	Maximum
Form 2 (*n* = 37)	59.24	14.09	16	79
Form 4 Science (*n* = 36)	73.50	13.14	44	97
Lower 6 Science (*n* = 38)	79.71	12.31	50	97
Pilot Sample (*N* = 111)	70.87	15.67	16	97

Table 2
Means and Standard Deviations and Range of SSE scores for the Actual Sample

Class	Means	*SD*	Minimum	Maximum
Form 2 (*n* = 122)	56.49	16.63	22	98
Form 4 Science (*n* = 127)	69.39	13.39	35	95
Lower 6 Science (*n* = 63)	79.35	11.70	50	100
Actual Sample (*N* = 312)	66.35	16.34	22	100

Instrument Reliability

The scale was also checked for its reliability in measuring SSE. The internal consistency and the test-retest measure of stability (after two weeks) for each class and for the whole pilot sample are as shown in Table 3. A moderately high alpha reliability value was obtained for all three classes and for the whole pilot sample, showing that the scale had high internal consistency in measuring SSE. The result for test-retest stability was also moderately high and consistently so for all three classes. Quite similar results were obtained when the scale was administered on the actual sample (refer to Table 4).

Item-total-correlations were also computed for the pilot sample (N = 111) to check on how well each item fitted with the total construct measured. The results showed that all 10 items correlated significantly ($p < .001$) with the total, with correlation values ranging from .46 to .73. Meanwhile, Bartlett's Test of Sphericity, "a statistical test for the presence of correlations among variables" and which provides "the statistical probability that the correlation matrix has significant correlations at least among some of the variables" (Hair, Anderson, Tatham & Black, 1998, p. 99) was also computed. The results showed that there were significant inter-item-correlations ($p < .001$), with correlation values ranging from .40 to .64, thus showing that the items in the scale measured the same construct. With the actual sample (N = 312), item-total-correlations ranged from .52 to .70, $p < .001$, while the inter-item-correlations ranged from .50 to .70, $p < .001$.

Table 3

Science Self-Efficacy Scale Reliability

Instrument Reliability	Form 2 (*n* = 37)	Form 4 Science (*n* = 36)	Lower 6 Science (*n* = 38)	Pilot Sample (*N* = 111)
Internal consistency (Alpha Reliability Coefficient)	.71	.76	.73	.80
Stability (test-retest) Pearson *r*	.83*	.83*	.85*	.89*

Note. *N* of items for the scale = 10
* Correlation is significant at the .01 level (2-tailed)

Table 4

Science Self-Efficacy Scale Reliability

Instrument Reliability	Form 2 (*n* = 122)	Form 4 Science (*n* = 127)	Lower 6 Science (*n* = 63)	Actual Sample (*N* = 312)
Internal consistency (Alpha Reliability Coefficient)	.80	.72	.75	.81

Note. *N* of items for the scale = 10

CONCLUSION

The above report on the development of the SSE Scale shows how the scale was formulated according to theoretical guidelines put forward by self-efficacy researchers on how SSE should be assessed. Steps were also taken to ensure that the scale has validity in measuring SSE. Measures of instrument reliability show that the scale has moderately high internal consistency and test-retest stability while item-total-correlations and inter-item-correlations indicated that the items in the scale were significantly correlated in measuring the same construct. In addition, the data on the distribution of SSE scores obtained through administration of the scale indicates the suitability of the scale in differentiating between the SSE of students at the three Form levels.

REFERENCES

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in a cross-cultural psychology* (pp. 137-164). Newbury Park, CA: Sage Publishers.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis.* (5th ed.). New Jersey: Prentice-Hall International, Inc.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology, 16*, 331-345.

Pajares, F. (1996a). *Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence.* [On-line]. Paper presented at a symposium on Measuring and mismeasuring self-efficacy: Dimensions, problems, and misconceptions. Annual meeting of the American Educational Research Association, New York. Available: http://www.emory.edu/EDUCATION/mfp/aera2.html.

Pajares, F. (1996b). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.

Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.) *Advances in motivation and achievement* Volume 10 (pp. 1-49). Greenwich, CT: JAI Press.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86*(2), 193-203.

Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology, 8*, 76-86.

Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*, 91-100.

Stajkovic, A., & Luthans, F. (1998). Social cognitive theory and self-efficacy: Going beyond traditional motivational and behavioral approaches. *Organizational Dynamics*, *26*(4), 62-74.
Woolfolk, A. E. (1998). *Educational Psychology* (7th ed.) Boston: Allyn & Bacon.
Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31*(4), 845-862.
Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic achievement: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663-676.
Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*(1), 51-59.

Appendix A

(Sample Items)

SCIENCE SELF-EFFICACY SCALE

Name: ____________________ Class: _______________

Instructions:
Suppose that you are asked the following questions in a science test.

Please indicate **how confident you are** that you will be able to give the correct answer to such questions. You must give your answer to each question in 30 seconds. **You do not have to write out the answers to the questions.**

It is important that you **do not guess** but give a realistic estimate of whether your answer is correct. Please use the scale below:

If you are **not confident at all** that your answer is correct, mark (/) **1**
If you are **completely confident** that your answer is correct, mark (/) **10**
If the estimate of your confidence is **between 1 and 10**, mark the appropriate number from **2 to 9**.

Please **mark one number only** for each question. **Thank You!**

No.	Question	Confidence Scale
1.	At room temperature, a saturated solution of ammonium sulfate can be prepared by dissolving 90 g of ammonium sulfate in 100 cm^3 of distilled water. If 25 cm^3 of distilled water is mixed with the solution, how many grams of ammonium sulfate should be added so that the solution	1 2 3 4 5 6 7 8 9 10

	becomes saturated?	
2.	State the energy conversions taking place in the following sequence of events: A boy throws a basketball to a height of 2 m. The ball falls and hits a toy car on the floor. The toy car moves and hits the wall.	1 2 3 4 5 6 7 8 9 10