DOCUMENT RESUME

ED 472 908 SE 067 299

AUTHOR	Ritter, Jennifer M.; Boone, William J.; Rubba, Peter A.
TITLE	An Extension Analysis on the Self-Efficacy Beliefs about Equitable Science Teaching and Learning Instrument for Prospective Elementary Teachers.
PUB DATE	2001-06-00
NOTE	38p.; In: Proceedings of the Annual Meeting of the Association for the Education of Teachers in Science (Costa Mesa, CA, January 18-21, 2001). For full proceedings, see ED 453 083.
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Elementary Education; *Equal Education; Evaluation Methods; Preservice Teacher Education; Science Education; *Self Efficacy; *Teacher Education Programs

ABSTRACT

The purpose of this study was to develop, validate, and establish the reliability of an instrument to assess the self--efficacy beliefs of prospective elementary teachers with regard to science teaching and learning for diverse learners. The study builds upon the work of Ashton, Webb, and Bandura. The Self-Efficacy Beliefs about Equitable Science Teaching (SEBEST) instrument is modeled after the Science Teaching Efficacy Belief Instrument (STEBI) and the Science Teaching Efficacy Belief Instrument for Prospective Teachers (STEBI-B). Based on the standardized development procedures used and the associated evidence, the SEBEST appears to be a content and construct valid instrument with high internal reliability qualities for use with prospective elementary teachers to assess personal self-efficacy beliefs for teaching and learning science for diverse learners. (KHR)

Reproductions supplied by EDRS are the best that can be made
from the original document.

AN EXTENSION ANALYSIS ON THE SELF-EFFICACY BELIEFS ABOUT EQUITABLE SCIENCE TEACHING AND LEARNING INSTRUMENT FOR PROSPECTIVE ELEMENTARY TEACHERS

Jennifer M. Ritter, Millersville University
William J. Boone, Indiana University
Peter A. Rubba, The Pennsylvania State University

ED 472 908

While the 1997 TIMSS data for grade four suggests that we are moving toward being "first in the world in mathematics and science achievement by the year 2000," these data do not indicate whether all groups of elementary students performed equally well. By contrast, the most recent National Assessment of Educational Progress (NAEP, 1996) results for 9-year-olds (fourth grade) show differences in science proficiency by race, ethnicity and gender. NAEP found that for the 9 year-old group, males out performed females, and White, non -Hispanic children scored higher than Black and non-Hispanic children with Hispanic children scoring the lowest. The NAEP results also show that 13 year-old males did better then females, that White non-Hispanic children scored higher than Hispanic children and that Black children scored the lowest.

Studies have shown gender inequity with higher academic achievement for boys than girls, classroom interactions between teacher and students that favor boys, sexual stereotyping, and gender bias in curricular materials. (American Association of University Women, 1992; Kahle & Meese, 1994; Kelly, 1985; Tobin, K., & Garnett, P. 1987) Several studies have documented that teachers interact with male students more then females (American Association of University Women, 1992, Brophy & Good, 1970; Datta, Schaefer, & Davis, 1968; Dweck & Bush, 1976; Martin, 1972; Sadker & Sadker, 1985), especially White males (Irvine, 1990; Sadker & Sadker, 1981). Jackson and Cosca (1974) and Sadker and Sadker (1981) found that

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

P. Rubba

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

teachers interact with, call on with greater frequency, praise more highly, and intellectually challenge students who are middle class, male, and White.

Additionally, teachers have been found to lack knowledge about the history, ethnicity and culture of their children (Pearson, 1985). Allen and Seumptewa (1988) found that many of the non-Native American teachers who teach Native American students are in a quandary with the differences in the way that the children learn. These teachers often leave the reservation because they do not feel that they connect with the students. Stegemiller (1989) concluded from an analysis of 31 studies that teacher expectations for students are based on four factors: social class, attractiveness, ethnicity and perhaps gender. Thus, a white boy who comes from a middle or high socioeconomic class and is academically average to above average, has multiple advantages with the teacher over a minority girl or a student who comes from a low socioeconomic home or is academically challenged.

The inequality in interaction between teachers and students who are from low socioeconomic homes, ethnically and culturally diverse, and girls is compounded by the curriculum of science, which has been neglected in the elementary classroom (Tilgner, 1990; Westerback, 1982). This neglect is evident in the limited time teachers spend on teaching science, teachers lack of confidence in their ability to understand science content and to be able to teach that content effectively and their negative attitude toward the science curriculum.

The teachers' beliefs and interactions are critical elements in the success of all students. Elementary teachers have been known to have negative attitudes toward science (Shrigley, 1974), do not care for science (Tilgner, 1990), and do not have confidence in their ability to teach science (DeTure, Gregory, & Ramsey, 1990; as cited in Park, 1996). This in turn causes elementary teachers to avoid teaching science to children (Czerniak & Chiarelott, 1990;

Westerback, 1982, 1984) or spend less time teaching science as compared to other subjects (Good & Tom, 1985; Weiss, 1987; Westerback, 1984). Czerniak & Chiarelott, (1990) found that the negative attitudes of teachers can be correlated to students negative attitudes about science. An attitude according to Enochs and Riggs (1990) "is a general positive or negative feeling toward something" (p. 625). A belief as defined by Koballa and Crawley (1985) is "information that a person accepts to be true" (p.223). Both, however, influence behavior. Thus, teachers' attitudes, beliefs and interaction are critical elements in the success of scientific literacy for all students. It is, however, the goal of this instrument to examine the beliefs of prospective teachers as opposed to the attitudes.

Bandura's self-efficacy theory was based on a relationship that he proposed existed between personal self-efficacy and the actions and behaviors of these patients. Bandura postulated that "self-efficacy beliefs influence the course of action people choose to pursue, how much effort they put forth in given endeavors, how long they would persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding, how much stress and depression they experience in coping with taxing environmental demands, and the level of accomplishments they realize" (p. 3).

Bandura (1995) contrasts people with different senses of efficacy as follows:

> People who have a low sense of efficacy in given domains shy away from difficult tasks, which they view as personal threats. They have low aspirations and weak commitment to the goals they choose to pursue. When faced with difficult tasks, they dwell on their personal deficiencies, the obstacles they will encounter, and all kinds of adverse outcomes rather than concentrate on how to perform successfully. They slacken their efforts and give up quickly in the face of difficulties. They are slow to recover their sense of efficacy following failure or setbacks. Because they view insufficient performance as deficient aptitude, it does not require much failure for them to lose faith in their capabilities. They fall easy victim to stress and depression (p. 11).

On the other hand:

> People who have strong beliefs in their capabilities approach difficult tasks as challenges to be mastered rather than as threats to be avoided. Such an affirmative orientation fosters interest and engrossing involvement in activities. They set themselves challenging goals and maintain strong commitment to them. They invest a high level of effort in what they do and heighten their effort in the face of failures and setbacks. They remain task-focused and think strategically in the face of difficulties. They attribute failure to insufficient effort, which supports a success orientation. They approach potential stressors or threats with the confidence that they can exercise some control over them. Such an efficacious outlook enhances performance accomplishments, reduces stress, and lowers vulnerability to depression (Bandura, 1995, p. 39).

Bandura's philosophy of the self-efficacy construct included his theory that self-efficacy beliefs affect how people think, act, feel and motivate themselves concerning all aspects of their lives. He interpreted, however, efficacy beliefs as having varying levels of importance. The most fundamental beliefs are those around which people structure their lives (Bandura, 1997, p. 43). Such beliefs have predictive value because these types of beliefs guide which activities are undertaken and how well they are performed. Bandura found this predictive value to be of the utmost importance because it gave way to the fact that if the self-efficacy beliefs of people could be influenced, people could achieve at levels they once thought they were incapable.

The self-efficacy construct, as described by Bandura, consists of two cognitive dimensions: personal self-efficacy and outcome expectancy. Bandura (1977, 1981, 1986, 1995, & 1997) defined personal self-efficacy as "judgments about how well one can organize and execute courses of action required to deal with prospective situations that contain many ambiguous, unpredictable, and often stressful elements" (p. 201). Bandura (1977) portrays outcome expectancy as "a person's estimate that a given behavior will lead to certain outcomes. An efficacy expectation is the conviction that one can successfully execute the behavior required to produce the outcomes. Outcome and efficacy expectations are differentiated, because

individuals can believe that a particular course of action will produce certain outcomes, but if they entertain serious doubts about whether they can perform the necessary activities such information does not influence their behavior" (p. 193). Bandura (1997) also noted that people who believe that their behavior can influence the outcome of a situation act more assertively then those who believe that outcomes cannot be influenced by their behavior.

The construct of self-efficacy beliefs is grounded in social learning theory and is the product of a complex process of self-persuasion that relies on cognitive processing of diverse sources of efficacy information. These include performance accomplishments, vicarious experience, verbal persuasion and emotional and physiological arousal.

Currently, 25 of the 50 largest school districts in the United States have children of color as the majority student population (Banks, 1991). In states such as New Mexico, Texas and California children of color comprise 70 percent of the total student population (Quality Education for Minorities Project, 1990). Children of color make up 30 percent of the students in the country overall and the growth rate of the minority population segment is expected to increase to 40 percent by the year 2020 (Pallas, Natriell, & McDill, 1989). By contrast, when the demographics of the prospective elementary teacher population is examined, it is found to be predominately white, middle class and female (Banks, 1991). The elementary teacher population continues to be Caucasian, monolingual, and female with backgrounds different from those they will teach, while the face of the school population in the United States is becoming more diverse (American Association of Colleges of Teacher Education, 1987; Banks, 1991; Ducharmen and Agne, 1989; Haberman, 1987).

Science for All Americans (1989) recognizes these inequalities and proposes that scientific literacy needs to be a goal of school science education for <u>all</u> young people, "those who

in the past who have largely been bypassed in science and mathematics education: ethnic and language minorities and girls" (p. xviii). Questions concerning how scientific literacy can be achieved given inequality in interaction due to race, class and gender differences and teacher beliefs concerning the science curriculum are vital.

To ensure scientific literacy for all, it is important for elementary teachers to understand student diversity and be able to teach science for a diverse student population. Part of the solution may be in understanding the behaviors of prospective elementary teachers. Teacher beliefs appear to be good predictors of behavior (Ashton & Webb, 1986a, 1986b; Bandura, 1986; Riggs, 1988; Enochs & Riggs, 1990). Teacher self-efficacy beliefs, in particular, have been found to be valid predictors of practicing and prospective elementary teachers' behavior regarding science teaching and learning (Ashton & Webb, 1986a, 1986b; Bandura, 1986; Riggs, 1988; Riggs & Enochs, 1990).

Purpose of the Study

The purpose of this study was to develop, validate and establish the reliability of an instrument to assess the self-efficacy beliefs of prospective elementary teachers with regards to science teaching and learning for diverse learners. This is an important area of self-efficacy belief assessment for which an instrument does not exist. The study built upon the work of Ashton and Webb (1986a, 1986b) and Bandura (1977, 1986), and the instrument was modeled after the Science Teaching Efficacy Belief Instrument (STEBI) (Riggs, 1988) and the Science Teaching Efficacy Belief Instrument for Prospective Teachers (STEBI-B) (Enochs & Riggs, 1990). It was proposed to be titled Self-Efficacy Beliefs about Equitable Science Teaching (SEBEST).

According to Bandura, (1986, 1997), the construct of self-efficacy beliefs consists of the two dimensions: personal self-efficacy and outcome expectancy. Personal self-efficacy "is a judgment of one's ability to organize and execute given types of performances, whereas an outcome expectation is a judgment of the likely consequence such performances will produce" (Bandura, 1997 p.21). An aim in developing the SEBEST was for each of the dimensions of self-efficacy beliefs of prospective elementary teachers toward teaching learning science for diverse learners, i.e., personal self-efficacy, outcome expectancy (Bandura, 1986), to be represented as a subscale.

The SEBEST instrument was designed to assess preservice teachers self-efficacy and outcome expectancy beliefs with regard to teaching and learning science in an equitable manner when working with diverse learners. This is the context in which the term "equitable" was used in developing the SEBEST and it is used in this paper.

Development of the SEBEST

A seven-step plan was used to develop the SEBEST and build validity and high reliability into the instrument.

Step 1: Defining the Constructs and Content to be Measured

Diverse learners as recognized by *Science for All Americans* (1989) are "those who in the past have largely been bypassed in science and mathematics education: ethnic and language minorities and girls" (p. xviii). That definition was extended to include children from low socioeconomic backgrounds based on the research by Gomez and Tabachnick (1992). They found that the views of prospective teachers toward minority children and children from low-income families limit the children's opportunities to learn and prosper from schooling. Similarly, the work of Grant and Tate (1995) acknowledges "educational research becomes

problematic when it does not include race, class, and gender, and/or when these constructs are not rigorously interrogated" (p. 147). For example, *The IEA study of Science II: Science Achievement in Twenty-three Countries*, found that family economic factors, the educational level of the parents, the size of the family, and the amount of reading material in the home were related to achievement in science (Postlethwaite & Wiley 1992). Baker (1998) proposes that "parental attitudes and economic condition of the family could be the major determinant of whether a girl will receive an education" (p. 879).

Figure 1 presents the Content Matrix that was developed for use in this study to define the content for the SEBEST. It is composed of the self-efficacy construct (i.e., personal self-efficacy and outcome expectancy dimensions), the definition of diverse learners developed for the study (i.e., ethnicity, language minorities, gender, and socioeconomic dimensions), and the phrasing dimensions for Likert items to be included in the SEBEST (i.e., positive and negative).

Step 2: Draft Item Preparation

Information on practices that are effective for teaching science to diverse student populations explicated in science education and multicultural education research, (for example, AAUW, 1992; Kahle & Meese, 1995; Kelly, 1985; Tobin, 1996, Atwater, 1994; Brickhouse, 1994; Gomez, 1996; Hodson, 1993; Rakow, 1985; Spurlin, 1995) informed the preparation of draft items for the SEBEST. One hundred ninety-five Likert type items, modeled after those composing the STEBI (Riggs, 1988) and STEBI-B (Enochs & Riggs, 1990) were drafted with at least six representatives for each cell in the Content Matrix presented in Figure 1.

Edward's (1957, pp. 13-14) fourteen guidelines for building item clarity also were used as a guide as draft items were written to reduce item error due to ambiguity. These guidelines include

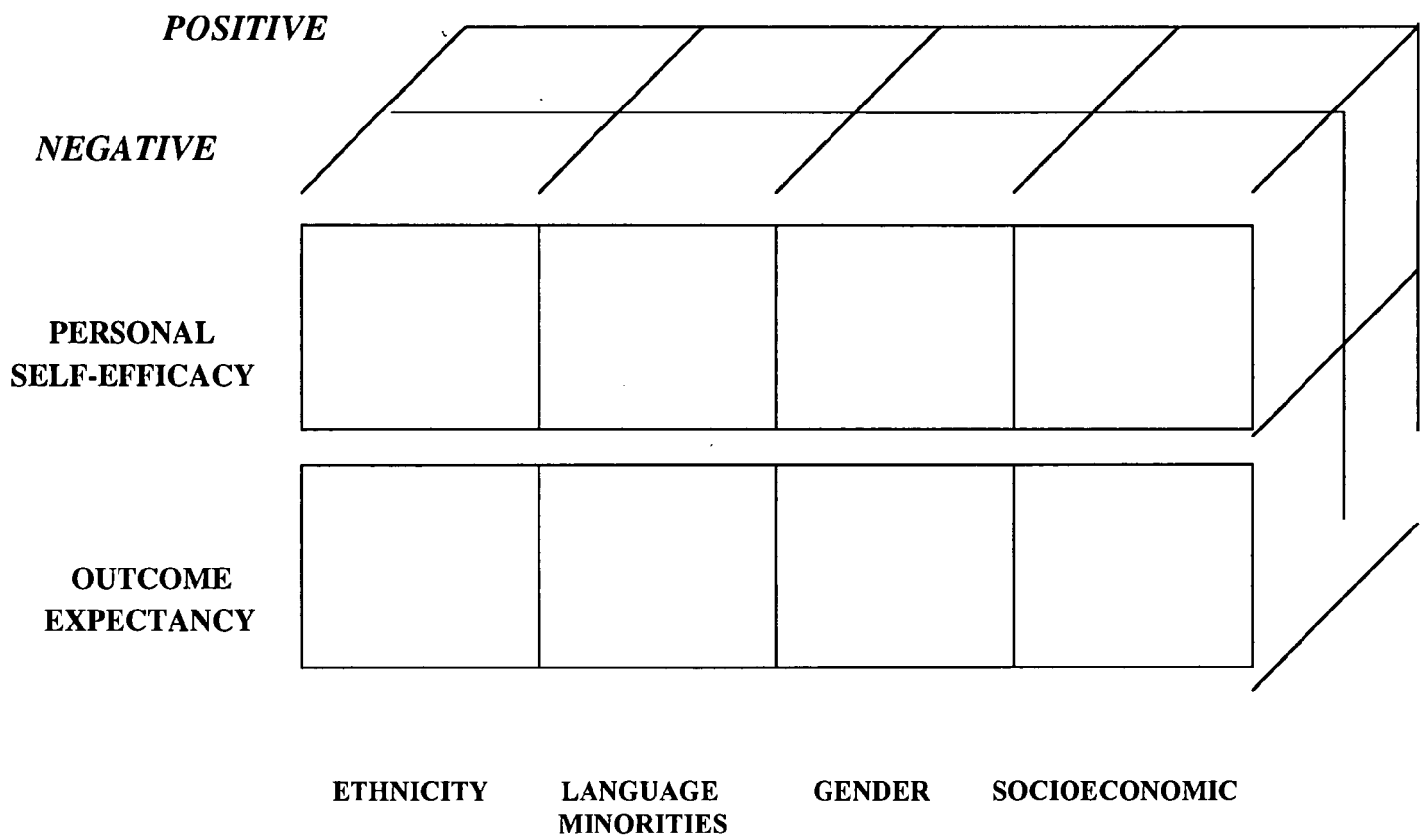

Figure 1. Content Matrix for the Self-Efficacy Beliefs about Equitable Science Teaching (SEBEST).

points, such as: (a) use items that refer to the present verses the past, (b) use simple, clear and direct language, (c) items should not use "all," "always," "none" or "never," (d) use care in using "only," "just" and "merely," and (e) avoid double negatives.

Step 3: Draft Item Review

A letter that explained the review task and the 195 draft items were submitted to 10 graduate students in Science Education at The Pennsylvania State University. Edward's criteria and the definitions of self-efficacy beliefs, personal self-efficacy, outcome expectancy, ethnicity, language minorities and gender also were included. The graduate students independently reviewed each of the draft items for clarity and comprehension by prospective elementary teachers. Comments for improvement was recorded directly on the draft items.

The feedback was used in revising the draft items. The revised items were resubmitted to the graduate students and subsequently revised until all ten graduate students judged that clarity and comprehension was achieved for at least five items in each cell of the matrix. Eighty items were identified within two rounds of review.

Step 4: Revised Item Content Validity

A panel composed of eight faculty members from inside and outside of The Pennsylvania State University representing science education, multicultural education, and self-efficacy research, was constituted for the purpose of judging the content validity of the eighty revised items. The panel members were given a letter of explanation, the revised items, the definitions of terms used within the instrument and Edwards's criteria. They worked independently to judge the content of the items and their feedback was used to revise the items. The items were to be resubmitted to the faculty members until at least four items in each cell of the Content Matrix (Figure 1) were judged content valid by five of the judges. However, this proved unnecessary

given that a sufficient number of the items, 48 with at least 6 items representing each cell in the Content Matrix, were judged content valid after one review. Those 48 content valid items constituted the "first draft" of the instrument.

Step 5: First Draft Instrument Try Out

The "first draft" instrument was administered to the 124 prospective elementary teachers in the five sections of SCIED 458--Teaching Elementary School Science and the 102 prospective elementary teachers in the nine sections of Elementary Student Teaching at The Pennsylvania State University during the second week of November 1998. These accessible groups represented the intended population for the final instrument. The resulting data were used in formulating the SEBEST as described in Step 6, below.

Step 6: SEBEST Formulation

The task in Step 6 of the development was three-fold: to identify a subset of the 48 items that: a) was construct valid, b) had high internal consistency reliability, and c) was representative of the Content Matrix presented in Figure 1. Factor analysis was used to help identify a construct valid subset of items. Coefficient Alpha, a measure of internal consistency, was used to examine the reliability of groups of items, item to total score correlation was used to determine the contribution of an item to total instrument score, and Chi Square was used to check item representation across the Content Matrix. Because the three qualities can be antithetical to one another – for example, the most construct valid and reliable set of items might not be representative of the Content Matrix -- these statistical techniques were applied multiple times and in combination to help select items for the SEBEST that gave the instrument the strongest profile across all three qualities.

The data used for these analyses were collected in step 5 by administering the 48-item "first draft" instrument to the 226 prospective elementary teachers in The Pennsylvania State University Elementary-Kindergarten Teacher Education (EK ED) program. Again, these included the students in the five sections of SCIED 458--Teaching Science in the Elementary School (n = 124) and in the nine sections of Elementary Student Teaching (n = 102) during the Fall semester of 1998. Usable data were secured from 217 of these prospective elementary teachers -- 120 of the students in SCIED 458 and 97 of the students in Elementary Student Teaching. The mean score on the 48 items among the 217 prospective elementary teachers was 151.45 with a standard deviation of 10.97 (scores on the 48 five-point Likert item instrument could range between 48 and 240).

Initial Factor Analysis Results

These data were subjected to Principal Component Factor Analysis using Varimax Rotation. The analysis generated 14 factors with an Eigenvalue of 1.00 or greater, that accounted for 64% of the variance in the instrument results. Because the desire was to select the smallest subset of the items that were construct valid, had high reliability and were representative of the Content Matrix, a Scree Plot was used to visually examine the number of factors and determine the number of significant factors. The Scree plot for the analysis is presented in Figure 2.

According to William and Goldstein (1984) the number of significant factors, or number of component factors to be retained, is indicated by a significant change in the slope of the plot set by a algorithm in the SPSS program -- the point at which the Scree plot curve breaks and forms a relatively straight line by a series of smaller, non-significant, Eigenvalues. In Figure 2, the point at which the contour of the curve changes significantly is marked with an arrow -- at an

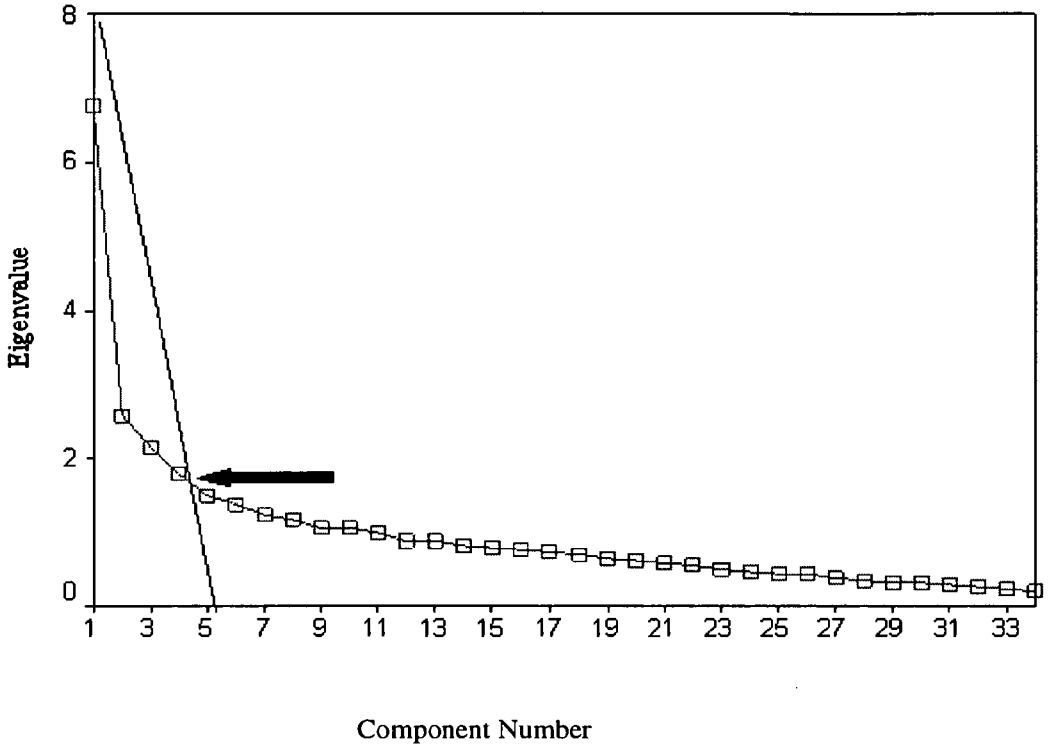

Figure 2. Scree Plot.

Eigenvalue of 1.7 and four components. Four factors were identified as significant using this method. Twenty-eight items loaded on these four factors. From a factor analysis perspective alone, the instrument might include 28 items.

The contribution each of the 48 items made to total instrument scores and reliability also was examined to determine the possible composition of the instrument from a reliability perspective. Thirty-four items were judged to be appropriate for inclusion in the instrument based on this perspective. That is, the 34 items that had the highest item to total instrument score correlations generated the highest Coefficient Alpha reliability for the total instrument and two subscales, i.e., personal self-efficacy, outcome expectancy.

Second Factor Analysis Results

The 34 items were subjected to Principal Component Factor Analysis using Varimax Rotation. These items loaded across four factors, which accounted for 39.2% of the variance in the data. Table 1 shows the item loading across the 14 factors, the variance accounted for by each factor and cumulatively across them, and the Content Matrix category for each item. As is noted, Factor 1 accounted for 11.6% of the variance in the instrument results, Factor 2 for 9.5%, Factor 3 for 9.3%, and Factor 4 for 8.9%. These percentages showed balanced variance across the four factors.

Additionally, the item factor loadings for the four factors were pure, with each factor being associated with either the Personal Self-Efficacy (PSE) or Outcome Expectancy (OE) dimension of Self-Efficacy. Eleven items loaded on Factor 1, all associated with Personal Self-Efficacy (PSE), particularly socioeconomic status, gender and ethnicity. Factor 1, therefore, was identified with PSE. Ten items, all of which were associated with Outcome Expectancy (OE) with language minorities, socioeconomic status, gender and ethnicity represented loaded

Table 1
Factor Analysis Results for 34 Items

Items No. Old	New	Factor 1	Factor 2	Factor 3	Factor 4	Matrix Cell*
1	1			0.71		PSE:LM
2	2	0.49				PSE:SES
3	3				0.47	OE:G
4	4		0.54			OE:E
5	5	0.63				PSE:E
6	6			0.65		PSE:LM
7	7	0.56				PSE:G
8	8		0.64			OE:E
9	9			0.76		PSE:LM
10	10	0.49				PSE:G
12	11				0.65	OE:E
14	12		0.50			OE:LM
15	13		0.63			OE:G
16	14		0.68			OE:E
17	15			0.65		PSE:LM
18	16	0.30				PSE:G
19	17				0.61	OE:SES
21	18				0.64	OE:G
22	19	0.69				PSE:E
24	20	0.38				PSE:E
25	21		0.42			OE:G
26	22	0.67				PSE:SES
28	23		0.45			OE:SES
29	24	0.79				PSE:E
30	25				0.49	OE:G
31	26			0.60		PSE:LM
34	27	0.67				PSE:E
40	28		0.42		0.29	OE:SES
41	29		0.53			OE:LM
42	30				0.64	OE:E
43	31		0.36		0.55	OE:SES
44	32	0.38				PSE:G
45	33		0.72			PSE:LM
48	34				0.39	OE:LM
% of Variance		11.6	9.5	9.3	8.9	
% Cumulative Variance		11.6	21.1	30.3	39.2	

* PSE = Personal Self Efficacy; OE = Outcome Expectancy
E = Ethnicity; G = Gender; LM = Language Minority; SES = Socioeconomic Status;

Factor 2. Six items identified with PSE loaded on Factor 3, all associated with language minorities. Eight items associated with OE, but from across the Content Matrix, loaded on Factor 4. The reliability of the PSE items that loaded on Factor 1 was .82 and on Factor 3 was .80. The reliability for the OE items that loaded on Factor 2 was .72 and on Factor 4 was .75.

Chi Square Results

Table 2 shows the distribution for the 34 items across the Content Matrix presented in Figure 1. A Chi-Square test was used to determine whether the 34 items were balanced across Personal Self-Efficacy/Outcome Expectancy and Ethnicity/Language Minority/Gender/ Socioeconomic Status for the PSE and OE dimensions of the Content Matrix. The resulting statistic, $X^2 = 2.71$, $df = 7$, was not significant at the .05 level of probability. This was interpreted as evidence that each of the two dimensions of the self-efficacy construct and each of the four diverse groups of learners were represented in the 34 item instrument to no significant difference.

The SEBEST Instrument

The task in Step 6 was to identify a subset of the tryout items that was construct valid, had high internal consistency reliability and was representative of the Content Matrix. Thirty-four items achieved this goal -- gave the instrument the strongest profile across all three qualities -- and so were used to compose the Self-Efficacy Beliefs about Equitable Science Teaching or SEBEST instrument. The 34 item SEBEST is presented in Appendix A. The even items compose the Personal Self-Efficacy or PSE Subscale, and the odd items compose the Outcome Expectancy or OE Subscale.

Table 2
Distribution of the 34 Items Across the Content Matrix

Dimensions/ Items	Ethnicity	Language Minority	Gender	Socioeconomic
Personal Self Efficacy	#7,#19,#27 #29,#33	#1.#5,#9,#13 #21,#25	#11,#15,#23 #31	#3,#17
Items 3, 9, 13, 21, 23, 27, 29, and 33 need to be reversed coded.				
Outcome Expectancy	#4,#12,#16 #22,#32	#18,#30, #34	#2,#8,#20, #26,#28	#6,#10, #14,#24
Items 4, 6, 8, 12, 14, 18, 22, 28 and 30 need to be reversed coded.				

Coefficient Alpha Reliability Results

Coefficient Alpha was used to assess the reliability of the 34 item SEBEST and its two subscales using data secured from the 217 prospective elementary teachers. The reliability of the entire instrument was found to be .87. The reliability was .83 for the 17 PSE items or subscale and .78 for the 17 OE items or subscale. A reliability of .87 indicates that 76% of a respondent's score is true score variance and 24% due to error. Similarly, a reliability of .83 indicates 69% true score while 31% is error, and a reliability of .78 indicates that 61% is true score and 39% is due to error.

According to standards presented by Helton, Workman and Matuszchk (1982), a reliability coefficient of .90 or higher is desired for classroom classification decisions, although this benchmark is rarely met. Remmers, Gage and Rummel (1965) support a reliability coefficient of .80 or higher for school use and .70 or higher for research instruments, especially

if group performance is only an issue. Reliability coefficients above .90 are considered necessary to make individual decisions with instrument results; above .80 are considered for research; and above .70 for initial group decisions that will be tested through additional means. (Nunnally, 1970) The reliability coefficient of .87 on the 34 item SEBEST, and .83 and .78 on its subscales were interpreted as being well within the acceptable reliability range for a research instrument.

Step 7: Further Study of Reliability

The internal consistency and test-retest reliability of the 34 item SEBEST were examined with data from two other samples of prospective elementary teachers (samples of convenience) during the Spring of 1999. One consisted of 23 prospective teachers enrolled in the Urban Early and Middle Childhood Education Program (URBED) at The Pennsylvania State University Delaware Campus, a teacher education program with an urban education focus. These prospective elementary teachers were at the mid-point of their student teaching experience in an urban elementary school. They had completed all of the required coursework for a BS degree and elementary teacher certification in Pennsylvania, including URBED 403--Using Science and Mathematics Knowledge and Assessment in Urban Settings along with an associated in-school (urban) clinical experience during the Fall semester of 1999. The purpose for including the urban preservice elementary teachers was to widen the diversity of the respondents to the instrument.

The other sample consisted of 102 prospective teachers enrolled in the Elementary-Kindergarten Teacher Education Program (EK ED) at The Pennsylvania State University University Park Campus. These prospective elementary teacher were at the mid-point of completing SCIED 458--Teaching Science in the Elementary School, along with mathematics

and social studies teaching and learning courses and an associated in-school clinical experience. The vast majority would be student teaching during the next semester (Fall 1999) and graduating with a BS in Elementary-Kindergarten Education and Pennsylvania elementary teacher certification. The EK ED students also completed the SEBEST twice: at mid-semester and at the end of the semester.

It should be noted that while one preservice teacher sample came from an "urban" teacher education program, both programs were conceptually similar, including the science pedagogy courses. Additionally, the purpose was not to compare the two samples on the instrument, but rather to study its reliability.

The Coefficient Alpha reliability for the SEBEST at mid-semester with the URBED prospective elementary teachers was .90, .81 for the PSE subscale, and .88 for OE subscale. At mid-semester, the reliability of the SEBEST with the EK ED prospective elementary teachers was .88 -- .83 for the PSE subscale, and .85 for OE subscale. The reliability of the SEBEST with the EK ED prospective elementary teachers at semester's end was .92 -- .87 for the PSE subscale, and .86 for OE subscale.

A Pearson-Product Moment correlation coefficient was calculated using data from the EK ED prospective teachers who completed both SEBEST administrations (n = 90) to estimate test-retest reliability, which was estimated to be .70 -- .70 for the PSE subscale and .67 for the OE subscale. These are considered to be estimates given the respondents were engaged in a methods course between the test and retest.

Further Test of Construct Validation

A Rasch analysis of the data was conducted in order to further evaluate the functioning of the instrument. Specifically considered were Rasch fit statistics, the distribution of Rasch calibrated survey items with regard to the latent trait, item reliability, and a principal component analysis of standardized residual correlations. These statistics have been used in a wide range of studies to evaluate the functioning of scales: see *Rating Scale Analysis* (Wright and Masters, 1982) for a full discussion of these issues. The analysis was conducted through use of the Rasch computer program Winsteps (Linacre and Wright, 2000).

Rasch fit statistics provide insight with regard to the functioning of an instrument. These statistics applied to survey items defining a scale help one evaluate whether or not an item is responded to in an idiosyncratic manner by respondents when all other responses to items are considered. In essence this statistic helps one learn if all items authored to define a latent trait (or variable) in fact do so. Analysis of both the personal self-efficacy and outcome expectancy scale revealed that no items appeared to generate high fit statistics. This suggests strength in the functioning of the scale.

Evaluating the manner in which items define a specific latent trait is another commonly used technique of assessing the functioning of tests and rating scales. In both scales (outcome expectancy and personal self-efficacy) there is a good distribution of items defining the latent trait. That means there are a range of items that are, for instance, easy to agree with and there are a range of items which are less easy to agree with (in relation to other items presented on the scale). Rasch item reliabilities were calculated for both the outcome expectancy subscale (.81) and the personal self-efficacy subscale (.98). These statistics help suggest good reliability.

One slight improvement, which might be explored in subsequent versions of this scale, is the use of a separate rating category scale for outcome expectancy and personal self-efficacy.

The reason for this is that respondents appear to have a higher probability of utilizing the strongly agree and agree categories for the outcome expectancy scale, where as these same respondents have a tendency to utilize more of the five-point scale for personal self-efficacy.

A principal component analysis of standardized residual correlations for items was computed for both subscales. Table 3 presented the outcome expectancy analysis, while Table 4 presents the statistics for personal self-efficacy. As part of that analysis, factor loading and person measures where evaluated. That analysis suggested that no SEBEST items define the factors other than those utilized for each subscale.

Conclusions

Based on the standardized development procedures used and the associated evidence, the SEBEST appears to be a content and construct valid instrument, with high internal reliability qualities, for use with prospective elementary teachers to assess personal self-efficacy beliefs for teaching and learning science for diverse learners. We suggest that this scale can be utilized in a number of ways. First, it can be used for the computation of mean linear measures based upon a set of items, for example, computation of a student's personal self-efficacy measure and outcome expectancy measure. This is the traditional way in which the SEBEST and others (e.g. Enochs and Riggs, 1990) have been used. However, we also suggest that the SEBEST can be used to help understand what it means to have a particular belief measure based upon a set of items such as the SEBEST. Figures 3 and 4 present plots that quickly convey the relationship between a preservice teacher's mean (average) raw response to SEBEST subscale items and their predicted responses to individual items in the subscale. The data for these two "most probable response plots" were collected from prospective elementary teachers during the

Spring of 1999 and subjected to Rasch Analysis. Each horizontal line in the two plots shows the

Table 3
Rasch principal component analysis of standardized residual correlations for personal self-efficacy items

Loading	SEBEST Item	
.74	29	I will be able to successful teach science to children of color.
.74	17	I will have the ability to help children from low socioeconomic backgrounds be successful in science
.62	23	I cannot held girls learn science at the same level as boys
.54	11	I can help girls learn science at the same level as boys
.53	31	I will be able to help girls learn science
.47	33	I will not be able to teach science successfully to White children
.40	15	I will be effective in teaching science in a meaningful way to girls
.33	27	I will not be able to successfully teach science to Asian children
.32	19	I will be able to successful teach science to Native American children.
.15	7	I will be able to meet the learning needs of children of color when I teach science.
.10	3	I do not have the ability to teach science to children from economically disadvantaged backgrounds.
-.61	21	I will not be able to teach science to children who speak English as a second language as effectively as I will be children who speak English as their first language
-.58	13	I do not know how to teach science concepts to children who speak English as a second language.
-.58	5	I can do a great deal as a teacher to increase the science achievement of children who do not speak English as their first language.
-.44	1	I will be able to effectively teach science to children whose first language is not English.
-.41	25	I will be able to effectively monitor the science understanding of children who are English Language Learners.
-.32	9	I do not know teaching strategies that will help children who are English Language Learners Achieve in science.

Table 4
Rasch principal component analysis of standardized residual correlations for outcome expectancy items

Loading	SEBEST Item	
.64	4	Even when teachers use the most effective science techniques in teaching science, some Native American children cannot achieve in science.
.57	12	Even when teachers use the most effective science techniques in teaching science, some children of color cannot achieve in science.
.27	6	Good teaching cannot help children from low socioeconomic backgrounds achieve in science.
.05	22	Children of color cannot learn science as well as other children even when effective science teaching instruction is provided.
.03	18	Children who speak English as a second language are not able to achieve in science even when the instruction is effective.
.01	16	Children of color can succeed in science when proven science teaching strategies are employed.
.01	24	A good science teacher can help children from impoverished backgrounds achieve in science at the same level as children from higher socioeconomic backgrounds.
-.62	26	Girls can develop in science at the same level as boys if they receive science instruction that is effective.
-.55	28	Girls do not have the ability to learn science as well as boys, even when effective teaching techniques are used.
-.44	32	White children can learn science as well as other children when effective science teaching is employed.
-.42	2	Girls can learn science if they receive effective science instruction.
-.26	30	children who are English Language Learners do not have the ability to be successful in science even when the science instruction is effective.
-.26	10	Effective science teaching can help children from low socioeconomic backgrounds overcome hurdles to become good science learners.
-.18	8	Girls are not as capable as boys in learning science even when effective instruction is provided.
-.17	20	Girls have the ability to compete academically with boys in science when they receive quality science instruction.
-.07	34	Children who are English Language Learners can be successful in learning science if the teaching is effective.
-.05	14	Effective science teaching cannot improve the science achievement of children from impoverished backgrounds.

predicted distribution of responses for each individual subscale item as a function of students' measures based upon the subscale (noted along the lowest horizontal line of both figures).

These "most probable response plots" (Figures 3 and 4) add understanding to SEBEST responses and scores in four important ways. First, the plots can be used to predict responses on a subset of items or an individual item in a SEBEST subscale when only selected responses to other items in the subscale have been made. For example, draw a vertical line from respondent #7 upward on Figure 3. This line shows that this prospective elementary teacher should have (from a probabilistic point of view) responded to items 21, 1, and 13 with the selection of the "agree" rating. It also shows that prospective elementary teacher #7 is likely to have responded to the remaining items with the selection of "strongly agree".

If only limited data were collected from a respondent, a set of responses to selected items would still allow one to predict what might be a candidate's answers to items that were not administered. This technique outlined in Figures 3 and 4 is currently being used in medicine to bring meaning to a "raw score" beyond what the "average" response of a candidate might be. In fact, this type of plot allows one to not necessarily administer all items of the survey, for the plot can be used to predict responses. Nonetheless, we suggest that the SEBEST be used in traditional ways, through the use of a total measure based initially upon a set of items, but also that figures such as 3 and 4 be utilized to better understand the meaning of a respondent's mean measure. Particularly powerful is to display the mean measure of subgroups of respondents on SEBEST subscales and how groups of respondents (e.g. white students, African American students) differ beyond the simple use of a mean measure.

Second, the plots show how likely a response (e.g., strongly agree, agree, uncertain, disagree, strongly disagree) is for a respondent to a set of items -- in the case of Figure 3 for the

```
|------+------+------+------+------+------+------+------|   ITEM
|                            SA    :     A    :    U DSD   Q23-R I Can Help Girls Learn
|                            SA    :     A    :    U DSD   Q31 I Can Help Girls Learn Sci
|                            SA    :     A    :   U  DSD   Q11 I Can Help Girls Learn Sci
|                            SA    :     A   :    U  DSD   Q15 I Will Be Effective Teaching Meaningful Way Girls
|                                                         |
|                         SA     :     A    :   ,U  :DSD   Q29   I Can Teach Child Of Color
|                                                         |
|                        SA   :     A    :    U    : DSD   Q7    Can Meet Learning Needs of Children Color
|                        SA   :     A    :    U    : DSD   Q17   I Can Help Low SES Students
|                       SA    :     A    :    U   :  DSD   Q33-R I Can Teach Sci To White Children
|                      SA    :      A   :    U    : D SD   Q27-R I Can Teach Asian Stu
|                    SA     :     A    :    U   :  D  SD   Q19   I Can Teach Native Am
|                    SA     :     A    :    U   :  D  SD   Q3-R  Have Ability Teach Low SES
|                                                         |
|                                                         |
|              SA   :     A    :    U    :   D     :  SD   Q25   I Can Monitor ESL
|                                                         |
|           SA   :      A    :   U    :   D     :    SD|   Q5    I Can Help Non Eng Speaker
|         SA    :      A    :   U    :   D     :     SD|   Q9-R  Do Learn Strat For Child Engl Learners
|                                                         |
|       SA    :     A    :    U    :   D     :    SD    |   Q13-R I Know Sci Concepts For Children ESL
|                                                         |
|     SA    :      A   :    U    :   D     :   SD       |   Q1    Teach Stu Not Eng
|                                                         |
|                                                         |
|  SA     :     A    :    U   :    D     :    SD        |   Q21-R I Can Teach ESL as Non ESL
|------+------+------+------+------+------+------+------|
                          |
              #7        #9              #2                   Respondent ID
               82       76              45                   Raw Score
```

Figure 3. Most Probable Response Plot for Personal Self-Efficacy.

Note: Items are ordered and spaced vertically from most easy to agree with (top of table) to least easy to agree with (bottom of table). An abbreviated text of each item is provided due to page limitations. The symbol "-R" denotes negatively phrased items whose scores were reversed prior to analysis. The PSE raw scores for three example respondents are presented along the base horizontal line of the figure (respondents #7, #9 and #2. To understand the meaning of the raw score, a vertical line can be drawn upward from any of the students. This line then shows what the predicted response of that particular student should be to each item. This brings immediate visual meaning to the commonly used raw score total for similar scales. The symbol ":" is used to show the boundary from one response category to another. From a probabilistic point of view, respondent #7 would answer "agree" to items 21, 1, and 13, and would have answered "strongly agree" to the remaining survey items. For his/her answers to the 17 PSE items, respondent #7 earned a "raw score" of 82. This type of plot shows what it means to have a score of 82.

BEST COPY AVAILABLE

```
|---------+---------+---------+---------+---------+---------|     ITEM
|                              SA      :      A   :  U :DSD    Q2    Girls Can Learn Science
|                                                       |
|                                                       |
|                        SA       :      A   :  U :  D  :SD    Q8-R  Girls Capable As Boys
|                                                       |
|                                                       |
|                     SA       :      A   :  U :  D    : SD    Q26   Girls Can Develop In Science
|                                                       |
|                     SA      :      A   :  U  :  D   :  SD    Q20   Girls Have The Ability To Compete Academically
|                    SA      :      A   :  U  :  D   :   SD    Q22-R Children Of Color Can Learn Science
|                   SA      :      A  :  U  :  D   :    SD    Q28-R Girls Have Ability To Learn Science
|                  SA       :      A   :  U  :  D   :    SD    Q6-R  Good Teachers Can Help Low SES Students Learn Sci
|                                                       |
|                 SA      :       A   : U  :  D    :     SD    Q10   Good Science Teachers Can Help Low SES Student
|                                                       |
|                SA      :      A   :   U :  D   :       SD    Q32   White Child Can Learn As Well As Others
|               SA       :      A   :  U  :  D   :       SD    Q24   Good Sci Teacher Can Help Low SES Students
|               SA       :      A   :  U :  D    :       SD    Q34   ESL Students Can Learn Science
|              SA       :      A   :  U  :  D   :        SD    Q14-R Effective Science Teaching Can Improve Sci
|              SA       :      A   :  U  : D    :        SD    Q16   Children Of Color Can Succeed
|             SA       :      A   :  U  :  D   :         SD    Q18-R ESL Can Succeed In Sci
|             SA      :       A  :  U  :  D   :          SD    Q30-R ESL Students Have The Ability
|            SA       :      A   :  U :  D    :          SD    Q4-R  Native Am Can Learn Sci
|            SA      :      A   :  U  :  D   :           SD    Q12-R Child Of Color Can Learn Sci
|---------+---------+---------+---------+---------+---------|
                  |
                 #2   #9 #7                                     Student ID
                 85    81 76                                    Raw Score
```

Figure 4. Most Probable Response Plot for Outcome Expectancy.

Note: Items are ordered and spaced vertically from most easy to agree with (top of table) to least easy to agree with (bottom of table). An abbreviated text of each item is provided due to page limitations. The symbol "-R" denotes negatively phrased items whose scores were reversed prior to analysis. The OE raw scores for three example respondents are presented along the base horizontal line of the figure (respondents #7, #9 and #2. To understand the meaning of the raw score, a vertical line can be drawn upward from any of the students. This line then shows what the predicted response of that particular student should be to each item. This brings immediate visual meaning to the commonly used raw score total for similar scales. The symbol ":" is used to show the boundary from one response category to another. From a probabilistic point of view, respondent #7 would answer "agree" to items 10, 32, 24, 34, 14, 16, 18, 30, 4 and 12, and would have answered "strongly agree" to the remaining survey items (2,8,26,20,22,28,6). For his/her answers to the OE 17 items, respondent #7 earned a "raw score" of 76. This type of plot shows what it means to have a score of 76.

BEST COPY AVAILABLE

PSE subscale. For example, the drawn line's points of intersection with response categories for respondent #7 shows that there is the greatest likelihood of an "agree" answer for item 21, a little less of a likelihood of an "agree" answer for item 1, and even less of a likelihood of an "agree" answer for item 13.

The most probable response plots can be used to very quickly bring meaning to a respondent's raw score total on a SEBEST subscale. As in the case of prospective elementary teacher #7, whose selection of rating categories (Strongly Agree, 5; Agree, 4; Uncertain, 3; Disagree, 2; Strongly Disagree, 1) earned an "82" raw score total for the PSE subscale, the plots allow one to see much more meaning in that score. Figure 4 presents similar information for example prospective elementary teachers for the OE subscale.

The SEBEST could be a valuable tool for science teacher educators working in practical and research settings to assess the personal self-efficacy beliefs of prospective elementary teachers with regards to science teaching and learning for diverse learners. Similarly, the SEBEST could be useful to multicultural teacher educators. For example, the SEBEST could be used to help identify if a particular course or program is achieving what it purports with regard to prospective elementary teacher preparation for science teaching and learning for diverse learner populations. Because the construct validity of an instrument is never fully established (Nunnally, 1970), the construct validity of the SEBEST will continue to need to be studied. In the process, the reliability of the SEBEST, including test-retest reliability, should be re-examined. Norming the SEBEST may provide some insights here and will provide additional information on the SEBEST that will be useful to users. Instruments such as the one presented in this paper should be viewed as evolving tools that will help improve the measurement that takes place in science teacher education. We suggest that subsequent

versions of this instrument should work toward including items for use with practicing elementary teachers should be pursued, a project the authors are undertaking.

References

Allen and Seumptewa (1988). The need for strengthening Native American science and mathematics education. Journal of College Science Teaching, 55, 364-369.

American Association of University Women. (1992). *The AAUW Report: How schools shortchange girls*, Washington, DC: The AAUW Educational Foundation.

American Association of Colleges for Teacher Education. (1987). Teaching teachers: Facts and figures, research about a teacher education project. Washington, DC: Author.

Ashton, P., & Webb, R. (March, 1982). Teachers' sense of efficacy: Toward and ecological model. Paper presented at the annual meeting of the American Educational Research Association, New York.

Ashton, P., & Webb, R. (1986a). Teacher efficacy attitudes, classroom behavior, and maintaining professional self-esteem. In P. Ashton, & R. Webb (Eds.). *Making a difference: Teachers' sense of efficacy and student achievement*. (pp. 55-89). New York: Longman Inc.

Ashton, P., & Webb, R. (1986b). Teachers' sense of efficacy, classroom behavior, and student achievement. In P. Ashton, & R. Webb (Eds.). *Making a difference: teachers' sense of efficacy and student achievement*. (pp. 125-144). New York: Longman Inc.

Atwater, M. M. (1994). Research on cultural diversity in the classroom. In D. Gabel (Ed.), *Handbook of research on science teaching and learning*. (pp. 558-576). New York: Macmillan.

Baker, D. (1998). Equity issues in science education. In B. Fraser & K. Tobin (Eds.) *International Handbook of Science Education Part Two* (pp. 869-895). Netherlands: Kluwer Academic.

Banks, J. (1991). Teaching multicultural literacy to teachers. *Teaching Education*, 4(1), 135-144.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J. H. Flavelll & L. Ross (Eds.). *Social cognitive development frontiers and possible futures*. Melbourne, Australia: Cambridge.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, New Jersey: Prentice-Hall.

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies*. Melbourne, Australia: Cambridge.

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

Brickhouse, N. (1994). Bringing in the outsides: Reshaping the sciences of the future. *Journal of Curriculum Studies, 26*(4), 401-416.

Brophy, J., & Good, T. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*, 356-374.

Czerniak, C. & Chiarelott, L. (1990). Teacher education for effective science instruction-a social cognitive perspective. *Journal of Teacher Education, 41*(1), 49-58.

Datta, L., Schaefer, E., Davis, M. (1968). Sex and scholastic aptitude as variables on teachers' rating of the adjustment and classroom behavior of Negro and other seventh-grade students. *Journal of Education Psychology, 59,* 94-101.

DeTure, Gregory, & Ramsey. (1990, April). The science preparation of elementary teachers. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Dweck, C. S., & Bush, E.S. (1976). Sex differences in learned helplessness: Differential debilitation with peer and adult evaluators. *Developmental Psychology, 12*, 147-156.

Ducharmen, E., & Agne, R. (1989). Professors of education: Uneasy residents of academe. In R. Wisniewski & E. Ducharme (Eds.), *The professors of teaching* (pp. 67-86). Albany, New York: SUNY Press.

Edwards, A. (1957). *Techniques of attitude scale construction.* New York, New York. Appleton-Century-Crofts.

Enochs, L. G. & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: preservice elementary scale. *School Science and Mathematics*, *90*(8), 694-706.

Gomez, M. L. (1996). Telling stories of our teaching, reflecting on our practices. *Action in Teacher Education, 18*(3), 1-12.

Good, T. L, & Tom, D.Y. H. (1985). Self-regulation, efficacy, expectations, and social orientation: teacher and classroom perspectives. In C. Ames & R. Ames (Eds.). Research on

Motivation in Education: Vol., 2. The Classroom Milieu (pp. 307-326). Orlando, Fl: Academic Press.

Grant, C.A. & Tate, W.F. (1995). Multicultural education through the lens of the multicultural education research literature. In J. Banks (Ed), *Handbook of research on multicultural education* (pp.145-166) New York: Macmillan.

Haberman, M. (1987). *Recruiting and selecting teachers for urban schools.* New York: ERIC Clearinghouse on Urban Education, Institute for Urban and Minority Education.

Helton, G.B., Workman, E.A., & Matuszck, P.N. (1982) *Psycoeducational Assessment: Integrating concepts and techniques.* Florida: Grune & Stratton.

Hodson, D. (1993). In search of a rationale for multicultural science education. *Science Education*, *77*(6), 685-711.

Irvine, J.J. (1990). *Black students and school failure.* New York: Greenwood Press.

Jackson, G., & Cosca, C. (1974). The inequality of educational opportunity in the Southwest: An observational study of ethnically mixed classrooms. *American Educational Research Journal*, *11*, 219-229.

Kahle, J. B., & Meece, J. (1994). Research on gender issues in the classroom. In D. Gabel (Ed.), *Handbook of research on science teaching and learning.* New York: Macmillan.

Kelly, A., (1985). The construction of masculine science. *British Journal of sociology of education*, *6*(2), 133-153.

Koballa, T. R. & Crawley, F.E. (1985). The influence of attitude of science teaching and learning. *School Science and Mathematics, 85*, 222-232.

Linacre, J.M. and Wright, B.D. (2001). Winsteps Computer Program. MESA Press, University of Chicago, Chicago, Illinois.

Martin, R. (1972). Student sex and behavior as determinants of the type of frequency of teacher-student contacts. *Journal of School Psychology*, *10*, 339-347.

NAEP, 1996. *Trends in academic progress.* Education Testing Service.

Nunnally, J. C. (1970). Introduction to psychological measurement. New York. McGraw-Hill.

Pallas, A., Natriell, G., & McDill, E. (1989). The changing nature of the disadvantaged population. *Educational Researcher*, *18*(5), 16-22.

Park, S. (1996). Development and validation of the Korean science teaching efficacy beliefs instrument (K-STEBI) for prospective elementary school teachers. *Dissertation Abstract International.*

Pearson P. D. (1985). *The comprehension revolution* (report No. 57). Urbana: University of Illinois at Urbana- Champaign, Center for the Study of Reading.

Postlethwaite, T. & Wiley, D. (1992). *The IEA study of science II: Science achievement in twenty-three countries.* New York: Pergamon.

Quality Education for Minorities Project. (1990). *Education that works: An action plan for the education of minorities.* Cambridge, MA: Author.

Rakow, S. J. (1985). Minority students in science: Perspectives from the 1981-1982 National Assessment in Science. *Urban Education, 20* (1), 103-113.

Remmers, H.H., Gage, N.L., & Rummel, J. F. (1965) *A practical introduction to measurement and evaluation.* New York: Harper & Row.

Riggs, I. M. (1988). The development of an elementary teachers' science teaching efficacy belief instrument. *Dissertation Abstract International.*

Sadker, D., & Sadker, M. (1981). The development and field trial of a nonsexist teacher education curriculum. *High School Journal, 64*, 331-336.

Sadker, D., & Sadker, M. (1985). Is the o.k. classroom o.k.? *Phi Delta Kappan, 66*(5), 358-361.

Science for all Americans. (1989). Washington, DC: American Association for the Advancement of Science.

Shrigley, R.L. (1974). The attitude of preservice elementary teachers toward science. *School Science and Mathematics*, 74(3), 437-446.

Spurlin, Q. (1995). Making science comprehensible for language minority students. *Journal of Science Teacher Education* , *6*(2), 71-78.

Stegemiller, H. A. (1989). *An annotated bibliography of the literature dealing with the contributing factors of teacher expectations on student performance.* (Report No. SP 031 604). South Bend: Indiana University at South Bend. (ERIC Document Reproduction No. ED 313 323).

Tilgner, P. J. (1990). Avoiding science in the elementary school. *Science Education, 74*(4), 421-431.

Tobin, K. (1996). Cultural perspectives on the teaching and learning of science', in M. Ogawa (ed.), Traditional culture, science and technology and development-Toward a new literacy for Science and Technology, University of Ibaraka, Mito City, Japan, 75-99.

Tobin, K., & Garnett, P. (1987). Gender related differences in science activities. *Science Education, 71,* 91-103.

Weiss, I. R. (1987). *Report on the 1985-1986 national survey of science and mathematics education.* Research Triangle Park, North Carolina: Center for Educational Research and evaluation, Research Triangle Institute.

Westerback, M. (1982). Studies on attitude toward teaching science and anxiety about teaching science in preservice teachers. *Journal of Research in Science Teaching. 19*(7), 603-616.

Westerback, M.(1984). Studies on anxiety about teaching science in preservice elementary teachers. *Journal of Research in Science Teaching, 21*(9), 937-950.

William, R. & Goldstein, M. (1984) Multivariate analysis: Methods and applications. New York: Wiley and Sons.

Wright, B. D., & Master, J. (1982) Rating Scale Analysis. MESA Press, University of Chicago, Chicago, IL.

Appendix A

Self-Efficacy Beliefs about Equitable Science Teaching (SEBEST)

Directions: Please indicate the degree to which you agree or disagree with each statement below by circling a response.

---

1. I will be able to effectively teach science to children whose first language is not English.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

2. Girls can learn science if they receive effective science instruction.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

3. I do not have the ability to teach science to children from economically disadvantaged backgrounds.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

4. Even when teachers use the most effective science techniques in teaching science, some Native American children cannot achieve in science.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

5. I can do a great deal as a teacher to increase the science achievement of children who do not speak English as their first language.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

6. Good teaching cannot help children from low socioeconomic backgrounds achieve in science.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

7. I will be able to meet the learning needs of children of color when I teach science.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

8. Girls are not as capable as boys in learning science even when effective instruction is provided.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

9. I do not know teaching strategies that will help children who are English Language Learners achieve in science.

   Strongly Agree Agree Uncertain Disagree Strongly Disagree

10. Effective science teaching can help children from low socioeconomic backgrounds overcome hurdles to become good science learners.

    Strongly Agree Agree Uncertain Disagree Strongly Disagree

BEST COPY AVAILABLE

11. I can help girls learn science at the same level as boys.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

12. Even when teachers use the most effective science techniques in teaching science, some children of color cannot achieve in science.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

13. I do not know how to teach science concepts to children who speak English as a second language.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

14. Effective science teaching cannot improve the science achievement of children from impoverished backgrounds.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

15. I will be effective in teaching science in a meaningful way to girls.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

16. Children of color can succeed in science when proven science teaching strategies are employed.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

17. I will have the ability to help children from low socioeconomic backgrounds be successful in science.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

18. Children who speak English as a second language are not able to achieve in science even when the instruction is effective.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

19. I will be able to successfully teach science to Native American children.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

20. Girls have the ability to compete academically with boys in science when they receive quality science instruction.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

21. I will not be able to teach science to children who speak English as a second language as effectively as I will to children who speak English as their first language.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

22. Children of color cannot learn science as well as other children even when effective science teaching instruction is provided.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

23. I cannot help girls learn science at the same level as boys.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

24. A good science teacher can help children from impoverished backgrounds achieve in science at the same level as children from higher socioeconomic backgrounds.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

25. I will be able to effectively monitor the science understanding of children who are English Language Learners.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

26. Girls can develop in science at the same level as boys if they receive science instruction that is effective.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

27. I will not be able to successfully teach science to Asian children.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

28. Girls do not have the ability to learn science as well as boys, even when effective teaching techniques are used.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

29. I will be able to successfully teach science to children of color.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

30. Children who are English Language Learners do not have the ability to be successful in science even when the science instruction is effective.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

31. I will be able to help girls learn science.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

32. White children can learn science as well as other children when effective science teaching is employed.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

33. I will not be able to teach science successfully to White children.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

34. Children who are English Language Learners can be successful in learning science if the teaching is effective.

Strongly Agree Agree Uncertain Disagree Strongly Disagree

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[X] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[ ] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").