

Measuring Educators' Beliefs about Diversity in Personal and Professional Contexts

Author(s): Cathy A. Pohan and Teresita E. Aguilar

Source: *American Educational Research Journal*, Vol. 38, No. 1 (Spring, 2001), pp. 159-182

Published by: American Educational Research Association

Stable URL: <https://www.jstor.org/stable/3202517>

Accessed: 13-01-2020 18:54 UTC

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/3202517?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/3202517?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*American Educational Research Association* is collaborating with JSTOR to digitize, preserve and extend access to *American Educational Research Journal*

## Measuring Educators' Beliefs About Diversity in Personal and Professional Contexts

Cathy A. Pohan

*San Diego State University*

Teresita E. Aguilar

*California State Polytechnic University*

*This study describes the development of two empirical measures designed to assess educators' personal and professional beliefs about diversity. A review of related studies and existing beliefs and/or attitudinal measures on diversity is provided. Steps utilized for instrument development and revisions are presented. Summaries of pilot, preliminary, and field testing developmental stages are provided, including results of our assessment of reliability and construct validity of the scales. These results provide promising support for internal consistency and (face and construct) validity of the instruments. Recommendations for educational and research uses of the measures are also provided.*

---

CATHY A. POHAN is an Associate Professor at San Diego State University, in the Department of Teacher Education, College of Education, 5500 Campanile Drive, San Diego, CA 92182. Her specializations are teacher preparation, diversity measures, learning and cognition, and multicultural education.

TERESITA E. AGUILAR is the Associate Dean in the College of Education and Integrative Studies and Professor in the Department of Ethnic and Women's Studies at California State Polytechnic University, Pomona, CA 91768. Her specializations are measurements in diversity, multicultural education, race and ethnicity, and cultural immersion experiences.

---

The role of teachers' beliefs, particularly those of preservice and beginning teachers, has been the focus of many educational studies for the past three decades (Pajares, 1992; Richardson, 1996; Wideen, Mayer-Smith, & Moon, 1998). Attitudes, beliefs, and expectations have been found to guide and direct teachers' responses toward various students (Good & Brophy, 1987; Grant, 1985; Nespor, 1987; Pajares, 1992). A substantial amount of evidence reveals that teachers hold beliefs about students that lead to differential expectations and treatment based on race/ethnicity (Guttmann & Bar-Tal, 1982; Hale-Benson, 1982; Rist, 1970), social class (Baron, Tom, & Cooper, 1985; Cooper, Baron, & Lowe, 1975; Rist, 1970), and gender differences (Brophy & Evertson, 1981; Brophy & Good, 1970; Sadker, Sadker, & Long, 1993). Clearly, if schools are to better serve the needs and interests of all students, particularly students from groups that have not fared well in the U.S. educational system, then low expectations, negative stereotypes, biases/prejudices, and cultural misconceptions held by teachers must be identified, challenged, and reconstructed.

Teacher educators nationwide have long been asking the question: How do we best help future and current teachers acquire the knowledge, skills, and attitudes that would result in culturally responsive teaching? Some scholars argue that attitudes represent a network of several beliefs that can be used to predict behavior (Cooper & Croyle, 1984; Fishbein & Ajzen, 1975). Bandura (1982) stated that beliefs mediate knowledge and action (behaviors/ skills). His theory, if related to culturally competent educators, suggests that educators' beliefs serve as filters for their knowledge bases and will ultimately affect their actions. Richardson (1996) advanced this theoretical notion by suggesting that prior beliefs are well established by the time a student enters college and that these beliefs are shaped by personal experience, schooling and instruction, and formal knowledge. Theoretically, we should not simply expect that an increase in (multicultural) knowledge would necessarily enhance the development of culturally competent educators, if educators lack a corresponding set of accepting/affirming beliefs about diversity. However, it is still important to consider the effects of multicultural interventions and courses on teachers' attitudes and beliefs about diverse others. Because of our interest in teacher beliefs, we sought measures to assess educators' beliefs or attitudes toward diversity.

Our interest in considering a two-dimensional (personal and professional) approach to assessing beliefs was based on the notion that there might be a situation in which one's personal beliefs about a given issue could be in direct conflict with his/her beliefs in a professional context. For example, in a personal context, an educator might believe that bilingualism is an asset in today's increasingly diverse and global society. Within a professional (i.e., schooling) context, however, this same educator might reject the notion of public monies being spent on bilingual education (i.e., maintenance programs). Thus, we believed it critical to measure beliefs about diversity in both personal and professional contexts.

This paper offers our definition of diversity, a brief review and critique of existing diversity measures reported in the literature, and the introduction of two beliefs about diversity scales. One scale measures beliefs about diversity in a general, more personal sense, and the second scale measures beliefs about diversity specifically within a professional, educational context. In the summary of pilot, preliminary, and field testing, we describe the procedures and results for determining the psychometric properties of our measures and suggest areas of further research on the measures themselves. We conclude with recommended uses of the measures for research or applied purposes.

### **Defining Diversity**

Consistent with our view of multicultural education as broad and inclusive of many aspects of sociocultural diversity, we were most interested in measuring subjects' beliefs about a range of diversity issues. We found that race and/or ethnicity were most frequently associated with the concept of diversity and that these concepts have been assumed to be the central concerns for the field of multicultural education. However, this narrow approach to diversity excludes the sociocultural educational discrepancies associated with social class, gender, religion, languages (other than English), and sexual orientation, which are also considered as central concerns or issues in contemporary approaches to multicultural education. In essence, our approach to defining diversity seeks to be inclusive of historically marginalized sociocultural groups; we do not ascribe to the narrower race or ethnic group approach.

### **Review and Critique of Existing Measures**

In our search for measures to assess teachers' beliefs about issues pertaining to diversity, we found several studies that used empirical and/or qualitative (Avery & Walker, 1993; Burnstein & Cabello, 1989; Ross & Smith, 1992) measures. However, our review only reports empirical measures.

From the field of psychology, we found studies that assessed preservice teachers' perceptions of the academic achievement abilities (Cooper et al., 1975) of diverse (using race and social class) students as well as practicing teachers' perceptions of academic achievement of students, on the basis of students' ethnic origin and gender (Guttman & Bar-Tal, 1982). The Bogardus Social Distance Scale was used in two additional studies (Byrnes & Kiger, 1989; Law & Lane, 1987) to investigate attitudes about tolerance for ethnic and/or racial diversity. These four studies focused on diversity issues that were limited to one or two diverse characteristics.

Tran, Young, and Dilella (1994) used a 7-point semantic differential cultural survey to investigate the elimination of stereotypical attitudes in ethnically diverse classrooms. Reliability and/or validity, however, were not discussed. In her study of 49 elementary school teachers, Washington (1981) measured attitudes toward multicultural education using a 16-item measure

scored on a 5-point Likert scale. These items were designed to measure “teachers’ opinions about concepts of school desegregation and multicultural education and the impact of these concepts on educational practice” (p. 188). Diversity, as we have defined it, was not central to her study. Further, data regarding reliability and validity of her measure were not reported in the study.

Tabachnick and Zeichner (1984) used a 47-item Teacher Belief Inventory to examine student teacher “perspectives” as they related to (a) knowledge and curriculum, (b) the teachers’ role, (c) the teacher-pupil relationship, and (d) student diversity. Data on the psychometric properties of this inventory, however, were not reported. Further, the items related to student diversity were more about curriculum/resources than personal beliefs about diverse others.

Wergin (1989) used an 81-item survey to study the attitudes of university freshmen students ( $n = 234$ ) toward cultural diversity. The 81-item survey was eventually reduced to four 5-item subscales. The subscales resulting from factor analyses were interest in cultural diversity ( $\alpha = .92$ ), ethnocentrism ( $\alpha = .89$ ), cross-cultural beliefs ( $\alpha = .85$ ), and racism ( $\alpha = .88$ ). In Wergin’s study, cultural diversity was conceptualized as foreign cultures, American culture, and race (limited to Black-White relations). There was no discussion of content and/or construct validity of the total scale or of the subscales.

Amodeo and Martin (1982) used a 44-item Cultural Attitude Test (CAT) to investigate teachers’ stereotypic attitudes about culturally (i.e., primarily ethnically) different students. This attitudinal scale measured stereotypes associated with Asian, Native, Black, Chicano, Anglo, Jewish, German, and Italian Americans. Higher scores indicated being less likely to “have stereotypic attitudes toward minority groups.” Data on reliability and validity were not reported for the CAT.

Moore and Reeves-Kazelskis (1992) used an 18-item Survey of Multicultural Education Concepts (SMEC) to study the beliefs of preservice teachers ( $n = 31$ ) about multicultural education. The SMEC contained items representing racism, sexism, stereotyping, linguistic views, special holidays, and educational practices (p. 6). High scores on this scale reflected positive beliefs and attitudes about multicultural education concepts. The SMEC offered the closest approximation to our approach to diversity because it included a range of diversity issues. However, Moore and Reeves-Kazelskis acknowledged that “validity and reliability of the SMEC have not been investigated” (p. 7).

One of the most frequently cited measures in the multicultural/beliefs literature is the 28-item Cultural Diversity Awareness Inventory (CDAI; Henry, 1986). Henry created a checklist to increase respondents’ awareness of their “attitudes, beliefs and behavior toward young children of culturally diverse backgrounds” (p. 2). A perusal of the items on this inventory suggests that cultural diversity was conceptualized as cultural groups (i.e., ethnic groups) and non-English speakers. Several items are presented within edu-

cational contexts, such as working with parents, developing curriculum, and testing procedures for children. The two-part published report of the CDAI includes the complete checklist in both English and Spanish. The second part includes suggestions and resources for improving educational services with diverse children. No information for scoring, interpretation, or for the reliability and validity of the checklist was provided in this booklet.

Larke (1990) used a modified version of the CDAI to study the cultural sensitivity levels of 51 elementary preservice teachers following a required multicultural education course. The reported data were on group percentages on the 5-point Likert scale, rather than on individual awareness scores. Larke did not discuss reliability or validity issues in her study. We found additional studies utilizing the CDAI (Davis & Turner, 1993; Davis & Whitner, 1994), which also excluded data on the reliability or validity of the inventory. Scoring procedures and data interpretation varied among the studies using the CDAI.

Our review of existing measures resulted in the following observations. Among the few studies of teachers' beliefs about diversity using empirical measures, reliability and validity data were seldom reported. Many of the measures focused on one or two specific characteristics of diversity (i.e., race, gender, ethnicity, and/or social class). Others focused on selected aspects of diverse learners (i.e., academic achievement abilities and stereotypic attitudes), curriculum and/or multicultural education, and cultural sensitivity (Table 1). We discovered that the data derived from these empirically based measures were interpreted with limited or no discussion on instrument reliability and validity.

Based on the results of the review of literature, we saw a need for sound instrumentation on diversity for research and applied purposes. We used the following three criteria to develop empirical measures on beliefs about diversity. We determined that our measures needed to (a) include a broader approach to diversity than was currently available, (b) address both personal and professional (i.e., educational contexts) beliefs regarding diversity issues, and (c) be rigorous and psychometrically sound. These criteria guided our development and refinement of the measures throughout the process.

### **Description of the Two Beliefs About Diversity Scales**

The Personal Beliefs About Diversity Scale includes 15 items relating to the following diversity issues: (a) race/ethnicity, (b) gender, (c) social class, (d) sexual orientation, (e) disabilities, (f) language, and (g) immigration. We were unable to establish a sound item on religion on this scale. For the personal beliefs scale, these issues are posed within the context of one's personal sphere or worldview (e.g., relationships, raising children, treatment by others, living conditions, and collective stereotypes).

The 25-item Professional Beliefs About Diversity Scale consists of items measuring diversity with respect to (a) race/ethnicity, (b) gender, (c) social class, (d) sexual orientation, (e) disabilities, (f) language, and (g) religion. The educational contexts (i.e., practices, resources, or approaches) included

Table 1

**Summary of Studies on Teachers' Attitudes/Beliefs about Diversity**

Study	Attitudes/beliefs investigated	Measures
Amodeo & Martin (1982)	Teachers' stereotypic attitudes about culturally different students	44-item Cultural Attitude Test
Byrnes & Kiger (1989)	Investigated differences in racial attitudes scores between teacher candidates and the general population	Bogardus Social Distance Scale and 12-item Social Scenarios Scale
Cooper, Baron, & Lowe (1975)	Importance of race and social class information on performance expectations	The Crandall Intellectual Academic Responsibility Scale
Davis & Whitner (1994)	Cultural sensitivity level of preservice teachers	The Cultural Diversity Awareness Inventory
Davis & Turner (1993)	How preservice teachers and culturally diverse families perceive one another	The Cultural Diversity Awareness Inventory
Guttman & Bar-Tal (1982)	Stereotypic perceptions of teachers	Scale to measure performance expectations
Henry (1986)	Investigating cultural awareness	The Cultural Diversity Awareness Inventory
Larke (1990)	Cultural sensitivity levels of student teachers	The Cultural Diversity Awareness Inventory
Law & Lane (1987)	Compared teacher racial attitudes with those of the general population over six decades	Bogardus Social Distance Scale
Moore & Reeves-Kazelskis (1992)	Preservice teachers' beliefs about multicultural education concepts	18-item Survey of Multicultural Education Concepts (SMEC)
Tabachnick & Zeichner (1984)	Impact of student teaching on teachers' perspectives	47-item Teacher Belief Inventory
Tran, Young, & Dilella (1994)	Eliminating stereotypic attitudes in ethnically diverse classrooms	26 paired items on a 7-point semantic differential
Washington (1981)	Teachers' opinions about school desegregation and multicultural education	16-item scale
Wergin (1989)	Assessing student attitudes toward cultural diversity	81-item survey

on the professional measure are (a) instruction, (b) staffing, (c) segregation/integration, (d) ability tracking, (e) curricular materials, and (f) multicultural versus monocultural education. These areas reflect an evolution of topics and contexts throughout the various test development phases. The items for both scales are presented in the Appendix.

### Initial Item Development

The first stage of instrument development included a survey of the socio-cultural diversity topics and issues frequently addressed in multicultural education courses and literature (Banks & Banks, 1993; Gollnick & Chinn, 1990; Nieto, 1992; Sleeter & Grant, 1988) to guide the item development. The dimensions of diversity originally selected for inclusion on both beliefs measures were (a) race, (b) ethnicity and culture, (c) social class, (d) gender, (e) sexual orientation, (f) exceptionality or issues about persons with disabilities, (g) language diversity, and (h) religion.

The same range of diversity issues was used on the professional beliefs scale, however, the items were presented within the context of schooling. Equal numbers of items per issue have not been utilized. Rather, attention was given to developing items that investigate a range of diversity issues, some of which are often considered more difficult to accept (e.g., gay/lesbian items). This range of issues provides a way of distinguishing individuals who are more accepting of a range of social diversity from those who are less accepting or less tolerant of diversity, as defined by the selected groups/topics included on the measures. Further, these beliefs measures were not designed to generate scores on the basis of individual diversity items or issues (i.e., subscales on race, social class, and gender). The original versions of the personal and professional beliefs scales consisted of 22 and 30 items, respectively (Pohan, 1994; Pohan & Aguilar, 1994).

### Initial Content Validity

The instruments were initially subjected to a preliminary review by three professors with a minimum of 4 years of teaching experience in the field of multicultural education ( $n = 2$ ) and social psychology ( $n = 1$ ). They were also reviewed by five graduate students in education who had previously completed at least one multicultural education course. The professors were asked to evaluate the instruments to determine if items (a) fell within the designated personal or professional belief domains outlined on the instruments, (b) were clear and unambiguous, and (c) were comprehensive in measuring beliefs about a range of diversity issues/topics. The graduate students were asked to complete each instrument and give (a) feedback regarding the clarity of individual items and administrative directions and (b) recommendations for the improvement of items. Data from this preliminary review led to some minor changes in the wording of items prior to the formal pilot testing.

### Response Format and Scoring

Both scales use a 5-point Likert-type format ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Several items on both scales are worded negatively to avoid a response set. These items are then reverse keyed to establish scale scores. Mean scores were utilized in much of the statistical analyses of the measures because our initial focus was on instrument development,



rather than on individual differences. Ultimately, however, individual scores are to be computed by adding items (sum scores) after recoding the negatively worded items.

Given the number of items on the scale and the response format, the range of possible scores for the Personal Beliefs About Diversity Scale is 15–75. The range of possible scores for the Professional Beliefs About Diversity Scale is 25–125. In the latest field testing, subsample means scores ranged from 56.23 to 64.41 and from 91.41 to 105.65 for the personal and professional beliefs scales, respectively.

The measures were designed to assess varying levels of acceptance for (or openness to) a range of diversity issues/topics. Low scores reflected general intolerance for diversity, whereas high scores reflected an openness or acceptance of most or all of the diversity issues. Midrange scores reflected a general tolerance or acceptance of some issues/topics and perhaps a degree of indifference for (or uncertainty toward) some of the issues/topics included in the measure. Midrange scores also indicated high acceptance of some issues/topics and low acceptance or tolerance for other issues/topics, resulting in a seemingly balanced (or midrange) score.

### **Current Scale Reliabilities**

The current version of both beliefs scales was administered to students enrolled in universities in California and Nebraska. Participants were enrolled in either an undergraduate or graduate multicultural education course. Approval through the institutional review boards was granted for this and all subsequent studies reported in this study. Because results from an analysis of variance (ANOVA) test did not reveal significant differences between the data from both universities, the data were pooled according to the two test conditions described here. More specifically, when the scales were administered at the beginning of the course, this was considered a pretest condition ( $n = 179$ ). When scales were administered at the end of the course, this was labeled as a posttest condition ( $n = 119$ ). It is critical to note, however, that pretest and posttest data were not collected from the same samples for the assessment of scale reliabilities.

Alpha coefficients on the final version of the scales and item-total correlations from the reliability tests are summarized in Tables 2 and 3. The alpha coefficients were .78 for both the pretest and posttest conditions of the Personal Beliefs About Diversity Scale. The alpha coefficients for the Professional Beliefs About Diversity Scale were .81 for the pretest condition and .85 for the posttest condition. These data support acceptable reliability for the current versions of both beliefs scales.

### **Summary of Pilot, Preliminary, and Field Testing**

In this section, we summarize the procedures, data analyses, and results from the pilot, preliminary, and field testing stages of test development. We provide an overview of our statistical analyses, giving particular attention to the tests for reliability and construct validity throughout the developmental pro-

Table 2

**Cronbach's Alphas and Item-Total Correlations for The Personal Beliefs About Diversity Scale (Final 1998 Version)**

Description of items	Pretest data ( <i>n</i> = 166)	Posttest data ( <i>n</i> = 109)
1. Interracial couples with children	.334	.482
2. Same-sex couples with children	.448	.323
3. Friendships across race/ethnicity	.262	.364
4. Friendships across sexual orientation	.549	.383
5. Immigration/refugee policy	.120	.097
6. Diversity strengthens the nation	.558	.539
7. Accepting gay/lesbian life	.546	.303
8. Accessible facilities too costly	.231	.388
9. Women in poverty	.252	.406
10. Leaders with disabilities	.478	.483
11. Men deserve higher wages	.432	.312
12. Poverty and motivation	.421	.541
13. Men are better leaders	.543	.528
14. Whites value education more	.390	.507
15. Immigrants to learn English	.352	.427
Alpha	.783	.780

cess. The scales were modified in each stage of development in an effort to develop reliable and valid measures of beliefs.

### Pilot Testing

The initial pilot testing was directed toward the questions of item clarity, scale reliabilities, and procedural issues related to administering the measures (e.g., completion time, clarity of directions). Two samples (*n* = 280) of undergraduate preservice education students enrolled in a required multicultural education course in a Midwestern land-grant university were included in the pilot study phase. Subjects were invited to participate, and did so on a voluntary basis.

The internal consistency of both belief scales was assessed using Cronbach's alpha. The alpha coefficients for the 22-item personal beliefs for Samples 1 and 2 were .77 and .74, respectively. The alpha coefficients for the 30-item professional beliefs were .86 and .87, respectively. These alpha coefficients indicated acceptable reliability for both scales.

Based on the item-total correlation data and frequency distributions, and with the goal of maximizing scale reliability, several minor revisions were made. Items with an item-total correlation index of .30 or greater were retained. If it was important to retain (due to content) a particular item with low discrimination indices, the item was reworded rather than deleted at this stage.

Table 3

**Cronbach's Alphas and Item-Total Correlations for The Professional Beliefs About Diversity Scale (Final 1998 Version)**

Description of items	Pretest data ( <i>n</i> = 166)	Posttest data ( <i>n</i> = 109)
1. Integrated classrooms	.375	.366
2. Middle-class classrooms	.354	.340
3. Gay/lesbian teachers	.367	.452
4. Importance of MCE	.300	.480
5. SPED money for gifted	.312	.378
6. Experience w/diverse students	.385	.363
7. Diverse faculty/staff	.469	.459
8. MCE for students of color	.303	.393
9. Monocultural education	.555	.599
10. People of color in texts	.432	.379
11. Physical limitations, regular classroom	.342	.259
12. Group students by ability	.325	.342
13. Tests to segregate students	.398	.426
14. Teachers adjust instruction	.454	.367
15. Males in math and science	.321	.409
16. Second language instruction	.135	.466
17. Teacher expectations by SES	.313	.387
18. Attention girls receive	.333	.331
19. More women in administration	.265	.423
20. Students of color in SPED	.463	.515
21. All fluent in second language	.418	.465
22. Fewer opportunities, SES	.222	.488
23. English only in schools	.349	.448
24. Religion and school policy	.257	.242
25. Understanding diverse religions	.458	.462
Alpha	.817	.855

*Note.* Although some items are weak in the pretest, they were retained because they were strong items on the posttest. The religion and school policy item was retained for content and because deleting it would not increase the alpha significantly. MCE = multicultural education; SPED = special education; SES = socioeconomic status.

### Preliminary Testing

The second stage of instrument development is the preliminary testing stage. Tests for scale reliabilities continued to be a central focus in this stage. We also began to address questions of instrument validity. When expected group differences are found on a given measure, this pattern may contribute to the construct validity of the measure (Anastasi, 1976; Cronbach & Meehl, 1955). For example, one might argue that individuals who have a strong ethnic identity would likely be accepting of other ethnic/cultural groups. Strength of ethnic identity might therefore be expected to be positively correlated with beliefs about diversity, if items pertaining to ethnicity are included on the beliefs measure. Based on this argument, statistical analyses

were conducted to determine whether selected variables were related to belief scores. As a separate test of validity, variables that would not be expected to influence scores were also analyzed using the appropriate statistical tests. Significance was determined by using an alpha of  $p < .05$ , although a higher level is acceptable in exploratory research.

Several samples of subjects were included in this stage of testing including undergraduate students ( $n = 92$ ), graduate students ( $n = 25$ ), and practicing educators from a rural ( $n = 29$ ) and an urban ( $n = 41$ ) school district. All subjects ( $n = 187$ ) were combined for reliability and validity tests because no statistically significant group differences were found.

Subjects were given a 14-item demographic sheet and the Multicultural Education Knowledge Test (Aguilar, 1993). The Multicultural Education Knowledge Test is a 35-item measure of topics often included in multicultural education coursework. Using Cronbach's alpha, the reliability coefficient of the knowledge test for the total pooled samples ( $n = 187$ ) was .95.

The internal consistency of both belief scales was assessed using Cronbach's alpha. The alpha coefficient for the revised 18-item personal beliefs was .81 and the alpha coefficient for the revised 23-item professional beliefs was .89. These alphas indicate acceptable to strong reliability for both scales and the coefficients were higher than in the initial pilot testing versions of the measures.

Again, all subjects were combined ( $n = 187$ ) for the statistical tests for validity. ANOVAs were used to test whether personal or professional beliefs scores were positively related to the number of multicultural courses completed. Four groups were generated: (a) no courses ( $n = 28$ ); (b) one course ( $n = 87$ ); (c) two to three courses ( $n = 57$ ); and (d) four or more courses ( $n = 16$ ). An emerging pattern of increasingly accepting personal beliefs about diversity was found when more multicultural coursework/content was completed. However, the difference was not statistically significant at the .05 level,  $F(3, 184) = 2.22, p \leq .08$ . The relationship between multicultural coursework and professional beliefs about diversity was statistically significant for this sample,  $F(3, 184) = 4.13, p \leq .007$ . More accepting (i.e., higher mean scale scores) professional beliefs were found for groups having engaged in more multicultural coursework, despite the different cell size among groups (Table 4).

ANOVAs were also used to test whether subjects with more cross-cultural experiences (i.e., domestic or foreign travel, Peace Corps or Vista volunteer, work/schooling in a foreign country) possess more accepting beliefs about diversity. Three groups were created for these analyses: (a) zero to one experience ( $n = 75$ ); (b) two experiences ( $n = 62$ ); and (c) three or more experiences ( $n = 52$ ). Personal belief scores were more accepting (i.e., higher) with increased cross-cultural experiences,  $F(2, 186) = 4.44, p \leq .01$ . This pattern did not persist for the influence of cross-cultural experiences on professional beliefs scores.

Age and gender were also tested using ANOVA and *t* tests, respectively. Beliefs did not vary as a function of age. However, both personal and pro-

Table 4  
**ANOVA for Multicultural Education Courses and Professional Beliefs Scores With Group Means (Preliminary Studies 1 and 2)**

ANOVA					
Source of variation	Sum of squares	Degrees of freedom	Mean square	<i>F</i>	<i>p</i> ≤
Between groups	2.68	3	.89	4.13	.007
Within groups	39.71	184	.21		
Total	42.39	187			

  

Group means				
Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>SE</i>
No courses	28	3.94	.53	.10
One course	87	4.10	.42	.04
Two to three courses	57	4.20	.43	.05
Four or more courses	16	4.42	.63	.15

ANOVA = analysis of variance. (preliminary Studies 1 and 2.)

professional beliefs varied as a function of gender. Compared with the males ( $n = 55$ ) in these samples, females ( $n = 133$ ) reported statistically higher (i.e., more accepting) scores on both personal ( $t = 4.81, df = 186, p \leq .001$ ) and professional ( $t = 5.79, df = 186, p \leq .001$ ) beliefs about diversity (Table 5).

Correlational analyses were conducted to determine the strength of the relationships among three variables: personal beliefs, professional beliefs, and multicultural education knowledge. The results suggest that the strongest relationship exists between personal beliefs and professional beliefs ( $r = .72$ ). Although moderate at best, a stronger relationship was reported between personal beliefs and multicultural education knowledge ( $r = .45$ ) than between professional beliefs and multicultural education knowledge ( $r = .38$ ).

### Field Testing

The third stage of scale development occurred in two waves of field testing with broader national (drawn from five states) samples. In addition to monitoring scale reliabilities, we assessed construct validity and potential threats to validity (Henerson, Morris, & Fitz-Gibbon, 1978). To study possible threats to validity, we investigated response set bias due to sequence of scale administration, as well as response bias due to social desirability.

Subjects ( $n = 756$ ) in the first wave of field testing were preservice and practicing teachers from four states. Preservice teachers from California, Florida, Nebraska, and Utah and practicing teachers from California, Florida,

Table 5

**T Tests for Gender and Beliefs Scores (Preliminary Studies 1 and 2)**

Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i> ≤
Personal beliefs						
Women	133	4.14	.41	4.81	186	.001
Men	55	3.80	.50			
Professional beliefs						
Women	133	4.25	.42	5.79	186	.001
Men	55	3.84	.48			

(preliminary Studies 1 and 2.)

and Nebraska were surveyed. In the second wave of field testing ( $n = 539$ ), four subsamples of preservice and practicing teachers were drawn from Colorado ( $n = 64$ ), California ( $n = 115$  and  $n = 109$ ), and Nebraska ( $n = 251$ ).

During the first wave of field testing, the alpha coefficients for the 22-item personal beliefs for both the preservice and practicing teachers were .80. The alpha coefficients for the 23-item professional beliefs were .82 for the preservice teachers and .77 for the practicing teachers. Following testing, both measures were revised. In the second wave of field testing, the substantially revised 16-item personal beliefs scale produced Cronbach alpha coefficients ranging from .64 to .81. among various subsamples. For the slightly reworded 23-item professional beliefs scale, the alphas ranged from .74 to .83. These data provide additional support for acceptable internal consistency.

Response set bias was investigated with the reverse sequencing of personal and professional beliefs scales. More specifically, a subset of preservice ( $n = 206$ ) and practicing teachers ( $n = 37$ ) was given the set of scales in reverse order. In the larger sampling subset, the personal belief scale was presented first. In the second and smaller sampling subset, the professional beliefs scale was presented first. In this analysis, *t* tests for independent samples were used to determine significant differences based on the sequence of scales. Neither personal nor professional beliefs scores varied significantly for either preservice or practicing teachers. The results suggest that the scores on these measures are not influenced by the order of administration.

Response bias, assessed by social desirability, was also tested at this stage of instrument development. A 10-item version of the Crowne-Marlowe Social Desirability Scale (Strahan & Gerbasi, 1972) was used to investigate whether participants felt the need to answer in a socially desirable manner. This scale was administered to all preservice subjects ( $n = 492$ ) and one subset of practicing teachers ( $n = 46$ ). The scale reliability for preservice teachers was .62; for practicing teachers, the alpha was .64. These alphas were somewhat lower than those reported by Strahan and Gerbasi, but are still acceptable for attitudinal measures. Social desirability is particularly im-

*Table 6*  
**Pearson Product-Moment Correlations for Beliefs and Related  
 Constructs (Field Test Data)**

	Preservice teachers ( <i>n</i> = 411)		Practicing teachers ( <i>n</i> = 209)	
	Personal beliefs (Items-22)	Professional beliefs (Items-23)	Personal beliefs (Items-22)	Professional beliefs (Items-23)
Professional beliefs	.77**		.67**	
Multicultural coursework	.21**	.20**	.02	-.00
Cross-cultural experiences	.14*	.11	.21**	.03
Multicultural knowledge	.16**	.19**	.24**	.07
Social desirability	.03	-.00	.11	.08
Age	.10*	.13*	-.05	-.10

\**P* ≤ .05. \*\**P* = .001.

portant since results can be misleading if some subjects gain points by responding in a socially desirable way whereas others are penalized for answering in a less desirable, yet candid way. In a case like this, “a large portion of the variance (spread) in scale scores will be response-set variance rather than substantive (i.e., attitudinal [belief]) variance” (Mueller, 1986, p. 74). As shown in Table 6, there was no significant relationship between social desirability and personal or professional beliefs for either preservice or practicing teachers.

Tests for construct validity were repeated in the field testing stage using correlational analyses with the variables of age, cross-cultural experiences, multicultural coursework, and perceived levels of knowledge about diverse topics. The Multicultural Education Knowledge Test was used to measure perceived multicultural knowledge and produced a reliability coefficient of .95. Statistically significant, although weak, relationships were found for preservice teachers for both personal and professional beliefs and age, multicultural knowledge, multicultural course work, and cross-cultural experiences (see Table 6). However, none of these variables were significantly related to practicing teachers’ professional beliefs. Only cross-cultural experiences and multicultural knowledge were significantly related to practicing teachers’ personal beliefs.

The variable of dogmatism was added to provide discriminant evidence of construct validity. Higher scores on the dogmatism scale reflect greater authoritarianism, intolerance, and/or closedness of one’s belief system (Rokeach, 1973). To the extent that our beliefs scales were designed to measure openness or acceptance of diversity, we would expect a negative correlation with a scale that measures individual differences in closedness of beliefs systems. Therefore, in the second wave of field testing, the Rokeach Dogmatism Scale (Rokeach, 1973) was included with a subsample (*n* = 115)

to measure dogmatism or general intolerance. Prior to administering the dogmatism scale, changes were made to remove gender bias from the wording of items (e.g., mankind to humankind, he to s/he). In our study, the dogmatism scale reliability score was .84. Scores on the dogmatism scale were negatively correlated with scores on the personal beliefs scale ( $r = -.24$ ;  $p \leq .008$ ). The correlation was significant, although the correlation coefficient was low to moderate.

## **Interpretation and Discussion of Psychometric Properties**

In this section, we address the questions of scale reliability and validity throughout the developmental process. This is accomplished by summarizing the data and offering a critical interpretation of our findings. A summary of the tests for reliability is provided to illustrate the manner in which revised scales produced different Cronbach alpha coefficients. The coefficients varied among samples and under pretest and posttest conditions throughout the developmental process. Our tests to establish validity are also summarized and interpreted as we consider the psychometric properties of the two beliefs measures throughout the developmental process.

### **Summary of Tests for Reliability**

Attitude scales often yield lower alpha coefficients than tests of intelligence or other nonattitudinal constructs (Mueller, 1986). Further, it is common that more items on a scale will yield a higher alpha coefficient. We expected fluctuations in alphas as we revised the scales and reduced the number of items. From pilot testing to field testing, reliability scores on the personal beliefs scale ranged from .71 to .81. The range of reliability coefficients on the professional beliefs scale fluctuated between .78 and .90. The reliability of each scale remained robust throughout the development process.

### **Interpretation of Tests for Construct Validity**

We offer the following summary and interpretations of our tests for validity at the preliminary and/or field testing stages of instrument development. It was encouraging that the personal and professional beliefs measures were strongly and positively correlated with each other both in the preliminary testing stage ( $r = .72$ ) and in the field testing stage ( $r = .77$ ,  $p = .001$  for preservice teachers;  $r = .67$ ,  $p = .001$  for practicing teachers). These correlations suggest overlap among the items, yet not to the extent that the scales could be seen as interchangeable. Each scale contains items on essentially the same range of diversity issues/topics, but we believe that the personal and professional contexts are uniquely different. For example, the items related to bilingualism often elicit an array of positions/beliefs across both personal and professional contexts. Some teachers believe that immigrants should learn English rather than maintain their first language (personal context). Yet, they also believe that students should be provided instruction in



their native language while they are becoming proficient in English (professional context). Many individuals believe that gays and lesbians should be allowed to teach (professional context) but do not believe that same-sex couples should have/raise children (personal context).

In the preliminary stage, age was not an influential factor for either personal or professional beliefs scores. There was a substantial age range among our subjects (i.e., 19–49 years). This result was encouraging because we would not have expected age to affect one's degree of openness toward diversity. In the field testing stage, we found a weak, although statistically significant, relationship between age and beliefs (on both measures) for preservice teachers ( $n = 411$ ). This pattern did not hold for practicing teachers ( $n = 209$ ). There is no theoretical premise upon which to build an argument for age as a predictor. Therefore, we will continue to assess this variable in future studies.

We had not anticipated a gender difference in beliefs scores when we developed the measure. However, throughout the test development phases, our results indicated that women were more accepting of diversity than were men on both beliefs measures. One interpretation of this difference in scores is the possibility of gender bias on the measures themselves (i.e., item bias). Another interpretation, which would support the validity of the measures, is that women are more accepting of diversity than are men. This argument is supported by Wergin (1989). In his attitudinal study, Wergin found that women held more positive attitudes than did men, which included issues of culture, ethnocentrism, and racism.

To further ascertain the question of instrument validity, we found that completion of multicultural/diversity education was more strongly associated with professional beliefs than with personal beliefs about diversity in the preliminary stage. However, field test results indicated that multicultural education coursework was significantly related to beliefs in both personal and professional contexts for preservice teachers. This pattern did not hold for practicing teachers. One interpretation is that preservice teachers receive more exposure to topics of diversity throughout their curriculum. However, it is also possible that the reality and/or complexity of classroom life causes practicing teachers to do what is most convenient or efficient instead of what the literature (or personal beliefs) suggests is best for children. For example, although teachers may know that academic limitations and discrepancies can result from an overreliance on ability grouping as a teaching strategy, they may still agree with the statement, "Generally, teachers should group students by ability levels." In this case, looking at the scores alone might lead the researcher to believe that teachers lack an understanding or awareness about ability grouping and issues of equity. Another possibility is that, as a result of classroom experiences, practicing teachers' personal and professional beliefs are more rigid and resistant to change.

Another related variable used to explore construct validity was direct cross-cultural experiences. In the preliminary stage, increased contact with diverse others was more strongly associated with personal than with pro-

professional belief scores. This pattern was also observed for practicing teachers in the field testing stage. For the preservice teachers in the field testing stage, cross-cultural experiences were significantly related to both personal and professional beliefs. Given that personal and professional beliefs are positively correlated with each other, the observed patterns with coursework and experiences suggest that both may be equally critical factors for teacher development with respect to diversity.

Repeating the correlational analyses of perceived multicultural knowledge, we found significant relationships with personal beliefs for preservice and practicing teachers, and with professional beliefs only for preservice teachers. The pattern was consistent with findings in the preliminary stage, although the correlation coefficients were substantially stronger in the preliminary stage than in the field testing stage.

One way to demonstrate construct validity involves "a set of correlations of the trait with other measures of . . . different but related traits" (McMillan & Schumacher, 1997, p. 237). Because Rokeach (1973) argued that his scale was designed to measure the degree of openness or closedness of a belief system (or general intolerance), this trait (i.e., dogmatism) is conceptually related to openness to diversity. Thus, the negative relationships found between dogmatism scores and scores on both the personal (statistically significant) and professional beliefs (although not significant) scales add support for construct validity.

The personal and professional beliefs measures meet the criteria that initially drove our research. More specifically, the scales do measure beliefs about diversity in the broadest sense of the term. The two scales, although related, measure beliefs in two different contexts that are unique from each other. Finally, the developmental process of the measures has been rigorous, leading to psychometrically promising measures.

### **Limitations Affecting Interpretations**

Although these exploratory tests provide support for validity, it is important to note that they were completed with our measures at various stages of development. Further, the beliefs measures continued to be refined throughout the developmental process on the basis of item-total correlations, range of scores, frequency distributions, and contributions to scale reliability. Consequently, the number and wording of items varied across the developmental stages of test construction.

Another factor to consider in these analyses is the limitations created by other measures used in our studies. For example, reliabilities for the social desirability scale were lower than we would have liked. Interpretation of these results should be viewed with caution. Although the dogmatism scale had a strong reliability ( $\alpha = .84$ ), we believe that the items on the available measure are quite dated. Unfortunately, we were unable to find more current reliable and valid empirical measures of the selected traits. Other measures such as multicultural coursework and cross-cultural experiences

used in the preliminary and field testing stages were primitive. On the other hand, the perceived Multicultural Education Knowledge Test consistently held strong reliability ( $\alpha = .95$ ). Our effort to validate our own scales has been limited by existing measures of related variables and constructs.

### **Recommendations for Further Testing and Validation**

As we continue to study the psychometric properties of the two beliefs measures, we have targeted three particular research areas that will inform instrument validity and reliability. First, we need to administer the measures to a more ethnically and racially diverse subject pool. Because most of our studies have been completed with preservice and/or practicing teachers, who have been predominantly White, we need to go outside the field of education or target historically Black or Hispanic schools for a more diverse subject pool. We anticipate a racial difference in scores, with subjects of color demonstrating more accepting beliefs about diversity scores. Wergin (1989) found this pattern when comparing the attitudes of Black and White respondents toward cultural diversity and racism.

The second area of study is establishing concurrent validity. Although we did not find other measures that conceptualized diversity in the broad manner in which ours did, it is possible to look at related measures and predict patterns of positively correlated measures due to some overlap in item content. Our primary limitation in this line of research is selecting measures for which reliability and validity have been or can be established. Relatedly, we recognize the need to repeat all previous tests for construct validity with the final 1998 scales.

A third area of study is investigating test-retest reliability. Most of our participants were enrolled in coursework related to diversity and/or multicultural education. The test-retest study should be done with participants in areas or coursework unrelated to the measures (e.g., mathematics or organic chemistry), so that changes in scores would not readily be attributed to a particular educational intervention.

### **Recommended Uses for the Current Beliefs Measures**

An initial step in eliminating the educational discrepancies identified in the multicultural literature (Banks, 1994; Banks & Banks, 1993; Nieto, 1996) is the baseline assessment of ideas that preservice and practicing educators believe about diverse others. Information obtained from empirical beliefs measures such as those developed in the current study could be used in a number of ways, a few of which are briefly described below. To date, all data have been collected with assurances of anonymity. However, identification of respondents poses a threat to validity.

#### **Education Programs**

Information obtained from the beliefs measures may be used to guide the development of a comprehensive diversity/equity plan, including a revised

curriculum for the development of teachers, counselors, and educational administrators. For example, if professional beliefs (and subsequent professional behavior) are directly influenced by personal beliefs, it is critical that the curriculum address deeper issues related to diversity (i.e., the “isms”—racism, classism—and oppression, prejudice, and discriminatory practices). Scores on the measures can be used to determine whether there is a need for a broader, more inclusive approach to diversity and multicultural staff/teacher development.

If personal beliefs are positively influenced by courses dealing with diversity and with direct cross-cultural experiences, program planners should expose students to various meaningful or optimal cross-cultural experiences within and outside their coursework. Participation in such educational opportunities is especially critical for students seeking a multicultural endorsement, specialization, or emphasis. The beliefs measures should be administered following these educational experiences to determine the short or long-term impact of the experiences on students' beliefs. This type of analysis requires a clear description of the educational experiences in order to enhance interpretation of data on beliefs changes.

Data obtained from the beliefs measures (i.e., item analysis) should also be used to identify staff development needs in the areas of multicultural education and diversity. The beliefs instruments may be used in conjunction with a multicultural or diversity knowledge assessment to determine a more comprehensive staff development plan and to address specific areas of ignorance, resistance, or closedness to diversity.

### **Research Uses**

The measures may be used to assess the impact of multicultural education interventions (e.g., workshops, seminars, course work, practica) through pretest and posttest measures to determine the approaches that are most effective or efficient. The impact of both long- and short-term educational interventions (i.e., weekend workshops), particularly those purporting to challenge personal beliefs directly, may be studied using these beliefs measures.

The beliefs measures may also be used in basic research, such as investigating the relationship between a person's beliefs and variables or factors that might effect educational policies or interventions. For example, one might investigate whether experiencing racially integrated classrooms in elementary or middle schools is related to higher or more accepting personal beliefs later in life. Support for a strong relationship might challenge the correct practice of perpetuating geographically, socioeconomically, or racially segregated elementary or middle level schools or classrooms.

Another research use may be to include these measures in testing more elaborate, comprehensive, or sophisticated empirical and theoretical models (i.e., using regression analyses), which may have implications for designing responsive teacher preparation curricula. It would be interesting to investigate the role or impact that beliefs about diversity, multicultural education

knowledge, and cross-cultural experiences have on enhancing culturally competent educators. Certainly each of these areas can be addressed within the design of teacher preparation programs. In terms of research, however, the statistical testing of these variables is limited by the availability of reliable and valid empirical measures for each variable tested.

The personal and professional beliefs measures, by design, should be useful as initial gauges of beliefs about diversity. In conjunction with qualitative assessments, results of the empirical beliefs measures may lead to a more thorough understanding of beliefs about diversity and their significance to effective and equitable teaching. This understanding could provide needed guidance in the development and design of educational and professional development programs intended to prepare more culturally responsive educators.

## APPENDIX

### **Personal Beliefs About Diversity Scale**

1. There is nothing wrong with people from different racial backgrounds having/raising children.
2. America's immigrant and refugee policy has led to the deterioration of America.
3. Making all public facilities accessible to the disabled is simply too costly.
4. Accepting many different ways of life in America will strengthen us as a nation.
5. It is not a good idea for same-sex couples to raise children.
6. The reason people live in poverty is that they lack motivation to get themselves out of poverty.
7. People should develop meaningful friendships with others from different racial/ethnic groups.
8. People with physical limitations are less effective as leaders than people without physical limitations.
9. In general, White people place a higher value on education than do people of color.
10. Many women in our society continue to live in poverty because males still dominate most of the major social systems in America.
11. Since men are frequently the heads of households, they deserve higher wages than females.
12. It is a good idea for people to develop meaningful friendships with others having a different sexual orientation.
13. Society should not become more accepting of gay/lesbian lifestyles.
14. It is more important for immigrants to learn English than to maintain their first language.
15. In general, men make better leaders than women.

**Professional Beliefs About Diversity Scale**

1. Teachers should not be expected to adjust their preferred mode of instruction to accommodate the needs of all students.
2. The traditional classroom has been set up to support the middle-class lifestyle.
3. Gays and lesbians should not be allowed to teach in public schools.
4. Students and teachers would benefit from having a basic understanding of different (diverse) religions.
5. Money spent to educate the severely disabled would be better spent on programs for gifted students.
6. All students should be encouraged to become fluent in a second language.
7. Only schools serving students of color need a racially, ethnically, and culturally diverse staff and faculty.
8. The attention girls receive in school is comparable to the attention boys receive.
9. Tests, particularly standardized tests, have frequently been used as a basis for segregating students.
10. People of color are adequately represented in most textbooks today.
11. Students with physical limitations should be placed in the regular classroom whenever possible.
12. Males are given more opportunities in math and science than females.
13. Generally, teachers should group students by ability levels.
14. Students living in racially isolated neighborhoods can benefit socially from participating in racially integrated classrooms.
15. Historically, education has been monocultural, reflecting only one reality and has been biased toward the dominant (European) group.
16. Whenever possible, second language learners should receive instruction in their first language until they are proficient enough to learn via English instruction.
17. Teachers often expect less from students from the lower socioeconomic class.
18. Multicultural education is most beneficial for students of color.
19. More women are needed in administrative positions in schools.
20. Large numbers of students of color are improperly placed in special education classes by school personnel.
21. In order to be effective with all students, teachers should have experience working with students from diverse racial and ethnic backgrounds.
22. Students from lower socioeconomic backgrounds typically have fewer educational opportunities than their middle-class peers.

*(Continued)*

23. Students should not be allowed to speak a language other than English while in school.
24. It is important to consider religious diversity in setting public school policy.
25. Multicultural education is less important than reading, writing, arithmetic, and computer literacy.

### Notes

Address correspondence to Teresita E. Aguilar, Ph.D., California State Polytechnic University, Pomona, CA 91768. E-mail: taguilar@csupomona.edu. Both authors contributed equally to this study. The authors extend their appreciation to Deborah Bandalos, Delia Saenz, and Roger Bruning for their reviews and critiques in the early stages of instrument development and to Sharon Evans for reviewing the final revision of the manuscript. We thank the AERJ reviewers for their comments and critiques of our initial submission. We must also acknowledge the very valuable insight and discussion provided by Barbara Plake of the UNL Buros Institute of Mental Measurement, which helped us to clarify and refine the manuscript. We also thank the University of Nebraska-Lincoln Research Council, Center for Curriculum and Instruction, and Teachers College, and the College of Education and Faculty Affairs at San Diego State University for their financial support at various stages of instrument development. This research was completed and supported, in part, by UNL.

### References

- Aguilar, T. E. (1993). *The development of the multicultural knowledge test*. Unpublished manuscript, University of Nebraska-Lincoln.
- Amodeo, L., & Martin, J. (1982). *A study of the effects of multicultural training on the factual knowledge and stereotypic attitudes of elementary and secondary teachers*. Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED 213 686)
- Anastasi, A. (1976). *Psychological testing*. New York: MacMillan.
- Avery, P., & Walker, C. (1993). Prospective teachers' perceptions of ethnic and gender differences in academic achievement. *Journal of Teacher Education, 44* (1), 27–37.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37* (2), 122–147.
- Banks, J. A. (1994). *Multicultural education: Theory and practice*. Boston: Allyn & Bacon.
- Banks, J. A., & Banks, C. M. (1993). *Multicultural education: Issues and perspectives* (2nd ed.). Boston: Allyn & Bacon.
- Baron, R., Tom, D., & Cooper, H. (1985). Social class, race, and teacher expectations. In J. B. Dusek, V. C. Hall, & W. J. Meyer (Eds.), *Teacher expectancies* (pp. 251–270). Hillsdale, NJ: Erlbaum.
- Brophy, J., & Everson, C. (1981). *Student characteristics and teaching*. New York: Longman.
- Brophy, J., & Good, T. (1970). Teachers' communication of differential expectations for childrens' classroom performance: Some behavioral data. *Journal of Educational Psychology, 61*, 365–374.
- Burnstein, N., & Cabello, B. (1989). Preparing teachers to work with culturally diverse students: A teacher education model. *Journal of Teacher Education*, September-October, Vol. 40, No. 1, pp. 9–16.

- Byrnes, D., & Kiger, G. (1989). Racial attitudes and discrimination: University teacher education students compared to the general student population. *College Student Journal*, 22, 176–184.
- Cooper, H. M., Baron, R. M., & Lowe, C. A. (1975). The importance of race and social class information in the formation of expectancies about academic performance. *Journal of Educational Psychology*, 67 (2), 312–319.
- Cooper, J., & Croyle, R. (1984). Attitudes and attitude change. *Annual Review of Psychology*, 35, 395–426.
- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Davis, L. E., & Turner, J. S. (1993, November). *An investigation of the cultural sensitivity level of elementary preservice teachers*. Paper presented at the the annual meeting of the Mid-South Educational Research Association. (ERIC Document Reproduction Service No. ED 372 054)
- Davis, L. E., & Whitner, L. V. (1994, November). *Preservice teachers and culturally diverse families: How do they perceive one another?* Paper presented at the annual meeting of the Mid-South Educational Research Association. (ERIC Document Reproduction Service No. ED 383 680)
- Fishbein, M., & Ajzen, A. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Gollnick, D. M., & Chinn, P. C. (1990). *Multicultural education in a pluralistic society* (3rd ed.) New York: Macmillan.
- Good, T., & Brophy, J. (1987). *Looking into classrooms* (4th ed). New York: Harper & Row.
- Grant, C. (1985). Race-gender status, classroom interaction, and childrens' socialization in elementary school. In L. Wilkinson & C. Marrett (Eds.), *Influences in classroom interaction* (pp. 403–422). New York: MacMillan
- Guttman, J., & Bar-Tal, D. (1982). Stereotypic perceptions of teachers. *American Educational Research Journal*, 19 (4), 519–528.
- Hale-Benson, J. (1982). *Black children: Their roots, culture, and learning styles*. Baltimore: Johns Hopkins University Press.
- Henerson, M. E., Morris, L. L., & Fitz-Gibbon, C. T. (1978). *How to measure attitudes*. Beverly Hills, CA: Sage.
- Henry, G. (1986, October). *Cultural Diversity Awareness Inventory = Inventario Sobre el Reconocimiento de Diversas Culturas*. Hampton, VA: Hampton University. Mainstreaming Outreach Project. (ERIC Document Reproduction Service No. ED 282 657)
- Larke, P. J. (1990). Cultural Diversity Awareness Inventory: Assessing the sensitivity of preservice teachers. *Action in Teacher Education*, 12 (3), 23–30.
- Law, S., & Lane, D. (1987). Multicultural acceptance by teacher education students: A survey of attitudes toward 32 ethnic and national groups and a comparison with 60 years of data. *Journal of Instructional Psychology*, 14 (1), 3–9.
- McMillan, J., & Schumacher, S. (1997). *Research in education: A conceptual introduction* (4th ed.). New York: Longman.
- Moore, T., & Reeves-Kazelskis, C. (1992, November). *The effects of formal instruction on preservice teachers' beliefs about multicultural education*. Paper presented at the annual meeting of the Mid-South Educational Research Association. (ERIC Document Reproduction Service No. ED 354 231)
- Mueller, D. J. (1986). *Measuring social attitudes: A handbook for researchers and practitioners*. New York: Teachers College Press.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Curriculum Studies*, 19 (4), 317–328.



- Nieto, S. (1992). *Affirming diversity: The sociopolitical context of multicultural education*. New York: Longman.
- Nieto, S. (1996). *Affirming diversity: The sociopolitical context of multicultural education* (2nd ed.). New York: Longman.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62 (3), 307–332.
- Pohan, C. A. (1994). *The development and validation of the Educators' Beliefs about Diversity Scale*. Unpublished doctoral dissertation, The University of Nebraska-Lincoln. UMI Dissertation Services, Vol. 55-10A, p. 3117 (AB No. 9507822).
- Pohan, C. A., & Aguilar, T. E. (1994). *Initial development and preliminary testing of two beliefs about diversity measures*. Research Monograph. Lincoln, NE: The University of Nebraska-Lincoln, Center for Curriculum and Instruction, Teachers College.
- Richardson, V. (1996). The role of attitude and beliefs in learning to teach. In J. Sikula, T. Buttery, & E. Guyton (Eds), *Handbook of research on teacher education* (2nd ed., pp. 102–119). New York: Macmillan.
- Rist, R. C. (1970). *Student social class and teacher expectations: The self-fulfilling prophecy in ghetto children. Challenging the myths: The schools, the Blacks, and the poor*. Cambridge, MA: Harvard Educational Review.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Ross, D., & Smith, W. (1992). Understanding preservice teachers' perspectives on diversity. *Journal of Teacher Education*, 43 (2), 94–103.
- Sadker, M., Sadker, D., & Long, L. (1993). Gender and educational equity. In J. A. Banks & C. A. Banks (Eds.), *Multicultural education: Issues and perspectives* (2nd ed., pp. 111–128). Boston: Allyn & Bacon.
- Sleeter, C., & Grant, C. (1988). *Making choices for multicultural education: Five approaches to race, class, and gender*. New York: SUNY Press.
- Strahan, R., & Gerbasi, K. (1972). Short, homogenous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28, 191–193.
- Tabachnik, B. R., & Zeichner, K. M. (1984). The impact of the student teaching experience on the development of teacher's perspectives. *Journal of Teacher Education*, Vol. 35, No. 1, 28–36.
- Tran, M. T., Young, R. L., & Dilella, J. D. (1994). Multicultural education courses and the student teacher: Eliminating stereotypic attitudes in our ethnically diverse classroom. *Journal of Teacher Education*, 45 (3), 183–189.
- Washington, V. (1981). Impact of antiracism/multicultural training on elementary teachers' attitudes and classroom behavior. *The Elementary School Journal*, 81 (3), 186–192.
- Wergin, J. F. (1989, November). *Assessing student attitudes towards cultural diversity*. Paper presented at the annual meeting of the Association for the Study of Higher Education. (ERIC Document Reproduction Service No. ED 313 979).
- Wideen, M., Mayer-Smith, J., & Moon, B. (1998). A critical analysis of the research on learning to teach: Making the case for an ecological perspective on inquiry. *Review of Educational Research*, 68 (2), 130–178.

Manuscript received December 7, 1998

Revision received November 1, 1999

Accepted January 24, 2000