

## DOCUMENT RESUME

ED 454 179

SP 040 053

AUTHOR Ambrosio, Anthony L.  
TITLE Final Report on the Multicultural/Diversity Assessment Project.  
INSTITUTION Emporia State Univ., KS.  
SPONS AGENCY National Inst. on Postsecondary Education, Libraries, and Lifelong Learning (ED/OERI), Washington, DC.  
PUB DATE 2001-02-06  
NOTE 27p.  
CONTRACT R309F70007  
AVAILABLE FROM Emporia State University, 1200 Commercial Street, Box 4036, Emporia, KS 66801.  
PUB TYPE Reports - Evaluative (142)  
EDRS PRICE MF01/PC02 Plus Postage.  
DESCRIPTORS \*Consciousness Raising; Cultural Awareness; \*Cultural Pluralism; \*Diversity (Student); Elementary Secondary Education; Gender Issues; Higher Education; \*Multicultural Education; Preservice Teacher Education; Sex Bias; Student Teacher Attitudes

## ABSTRACT

The Emporia State University Multicultural/Diversity Project developed a set of assessment instruments and a model evaluation plan to assess multicultural/diversity (MCD) outcomes in teacher education and general education programs. Assessment instruments and techniques were constructed to evaluate the impact of coursework on student attitudes, knowledge, and performance skills. Participants were incoming freshmen, beginning teacher education majors, non-education seniors, and teacher education seniors. Surveys examined four basic issues: the effect of the teacher education program on student attitudes about MCD issues; whether graduating students have the necessary knowledge to function in a diverse classroom/society; whether student teachers are gaining the necessary skills to function in a diverse classroom/society; and the effect of general education programs on students' attitudes and knowledge regarding MCD issues. Students demonstrated moderate attainment of MCD outcomes. Several areas of possible improvement were noted. Implications of this study include identifying and pursuing mechanisms for program change based on faculty discussion of essential knowledge and skills. Recommendations include using a battery of assessments rather than a single assessment when evaluating a program or a student, involving school practitioners in subsequent assessment development and program changes, and making students individually accountable for minimum acceptable performance levels. (SM)

PERMISSION TO REPRODUCE AND  
DISSEMINATE THIS MATERIAL HAS  
BEEN GRANTED BY

A. Ambrosio

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)

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.

## Emporia State University



### Final Report on the Multicultural/Diversity Assessment Project

2-6-01

Anthony L. Ambrosio, Ph.D.  
Director of Assessment  
Multicultural/Diversity Assessment Project  
Emporia State University  
ambrosia@emporia.edu

#### ESU Project 30 Committee

Jerry Long, Ph.D., Associate Dean of The Teachers College

Rodney Sobieski, Ph.D., Associate Dean of Liberal Arts & Sciences

Eileen L. Hogan, Ph.D., Department Chair of Early Childhood & Elementary Teacher Education

Marie C. Miller, Ph.D., Department Chair of Music

Cynthia Anast Seguin, Ph.D., Assistant Professor of Teacher Education &

Multicultural/Diversity

Specialist

Rachelle M. Smith, Ph.D., Assistant Professor of English & Director of English Composition

From a grant by The National Institute on  
Postsecondary Education, U.S. Department  
of Education (Grant # R 309 F 70007)

## Executive Summary

The purpose of the ESU Multicultural/Diversity Project (U.S. Department of Education grant #R309F70007) is to develop and refine a set of assessment instruments and a model evaluation plan to assess multicultural/diversity (MCD) outcomes in the teacher education and general education programs. The project represents a cooperative effort between the Teachers College and the College of Liberal Arts and Sciences. Assessment instruments and techniques were constructed to evaluate student attitudes, knowledge and performance skills. The project attempted to answer four basic research/evaluation questions: (a) What is the effect of the Teacher Education program on student attitudes on MCD issues? (b) Do graduating student-teachers have the necessary knowledge to function in a diverse classroom/society? (c) Are student teachers gaining the necessary skills to function in a diverse classroom/society?, and (d) What effect does the general education program have on students' attitudes and knowledge regarding general MCD issues? Evaluation results indicated that students demonstrated moderate attainment of MCD outcomes. Individual assessment results indicated several areas of possible improvement. Implications of this study include identifying and pursuing mechanisms for program change based on faculty discussion of essential knowledge and skills. Recommendations include using a battery of assessments rather than a single assessment when evaluating a program or individual student, involving school practitioners in subsequent assessment development and program changes, and making students individually accountable for minimum acceptable levels of performance.

## History and Background

Since 1991, The Teachers College and the College of Liberal Arts and Sciences (LAS) have been involved as partners in a group called Project 30 Alliance, a consortium of twenty-four member institutions that brings together faculty in Arts and Sciences with faculty in Education for the purpose of improving teacher education. Past collaborative efforts between the two ESU colleges have included: the development and delivery of an Ethnic/Gender Studies minor and an ESL/Bi-cultural endorsement, and the creation of the Minority Recruitment/Retention Task Force and activities sponsored by the Office of Minority Affairs. In addition, both colleges were committed to addressing Multicultural Education needs in their respective programs. LAS offers courses with a multicultural-intensive focus and the teacher education program infuses multicultural content throughout many classes. Since the instruction was in place, both colleges felt the need to initiate a comprehensive, program-wide multicultural/diversity outcomes assessment plan; one that incorporates the results of these assessments into the programs. Thus, the ESU Project 30 group had an established and positive collaborative focus. Therefore, it was the ideal group to spearhead an evaluation plan. But, before a plan could be constructed, a consensus of student learning outcomes had to be identified. These outcomes are discussed below.

### Student Learning Outcomes.

The learning outcomes adopted for this project originated from the efforts of the Multicultural/Diversity Task Force membership which reflected a campus-wide representation. The Task Force produced eight general learning objectives (see Table 1). Each of the learning objectives was further broken down into more specific affective, knowledge and performance objectives (see our web-site: <http://www.emporia.edu/teach/dean/grant/index.htm> for a complete

list). In addition, teacher education and general education courses were examined to identify course objectives. Information from both sources were combined and the resulting outcomes were clustered into performance areas according to similarity in content (see Table 2). The objectives then needed to be defined in observable, measurable terms (i.e., operationalized). To accomplish this, focus groups consisting of university, public school and community participants convened to answer three basic questions with regard to each of the seven performance objective areas: 1) What should graduates know?, 2) What should graduates be able to do? and 3) How would you know if they could do it (i.e., what would you observe as evidence)? In addition, a university-wide sample was surveyed by means of a project created instrument (the Knowledge-Based Diversity Questionnaire-KBDQ) to further identify critical MCD knowledge bases. The KBDQ assessed perceived importance of various content areas (including areas identified in the literature by noted professionals such as Banks, 1999 and Pritchey-Smith, 2000) as well as the degree to which faculty reported covering the content in their own courses. The results from all of these sources were utilized to drive the creation of the content knowledge and performance assessments.

---

**Table 1.** General learning outcomes for teacher preparation programs

1. The professional respects the dignity and worth of each individual and believes that all human beings possess the capacity for thought, feeling and learning.
2. The professional understands the influence of culture on personal growth and recognizes and respects cultural differences as manifested in personal points of view.
3. The professional uses knowledge of what is culturally significant, socially just and equitable in the general society and in specific cultures.
4. The professional recognizes the historical and current status of relations between minority-majority populations and how it affects the educational setting.
5. The professional understands and can utilize the component of effective education which includes global and multicultural perspectives.
6. The professional makes responsible use of assessment, measurement, and evaluation techniques and devices recognizing the effects of socioeconomic gender, ethnic, racial, and cultural factors in each of these endeavors.
7. The professional values the family as a partner in the educational process and facilitates an inclusive interface with the broader community.
8. The professional recognizes the importance of continued professional development and advocacy in maintaining cultural competence in educational practice.

---



---

**Table 2.** Seven assessment areas.

<u>Area</u>	<u>Description</u>
<i>Family Environment:</i>	Knowledge/skills associated with understanding the effect of the family on a child's education. This could also include skills to involve parents/families in the educational process.
<i>Community Environment:</i>	Knowledge/skills associated with understanding/utilizing the relationship between community and school. This may include ways in which a teacher can search for information about the community, use the community in the classroom, etc.
<i>Teacher/Student Expectancies:</i>	Knowledge about the effect of student and teacher expectations on the learning process. This includes strategies/skills to overcome the negative effects of these expectations (or enhance their positive effects).
<i>Assessment Interpretation:</i>	Knowledge/skills needed for the appropriate use of tests and test results (classroom and standardized) in a diverse classroom.
<i>Creating a Learning Environment:</i>	Knowledge/skills related to creating a learning environment in which all students in a diverse classroom can learn.
<i>Instructional Strategies/Methods:</i>	Knowledge about or skills related to specific teaching strategies/methods that promote learning in a diverse classroom.
<i>Bias In Education:</i>	Knowledge/skills related to identifying and remediating biased teaching methods and materials.

---

### **Method**

This project attempts to answer four programmatic evaluation questions:

1. What is the effect of the Teacher Education program on student attitudes on Multicultural/Diversity issues?
2. Do graduating student-teachers have the necessary knowledge to function in a diverse classroom/society?
3. Are student-teachers gaining the necessary skills to function in a diverse classroom/society?
4. What effect does the general education program have on students' attitudes and

knowledge regarding Multicultural/Diversity issues?

It is important that the sampling strategy and assessment strategy are designed in combination to make it possible to answer our basic evaluation questions. For example, if we wanted to see the effects of the teacher education program on MCD knowledge, we would examine the differences in knowledge between beginning teacher education majors and graduating student teachers. Similarly, if we wanted to determine whether students have the meaningful skills to perform in the real world, we would choose to examine performances of graduating student teachers. This is not a perfect design. There are concerns regarding group differences in age, experience and course work. However, some of these concerns have been addressed by the collection of demographic information from each respondent for the purpose of data disaggregation and statistical control.

#### Participants.

In addition to selecting a sampling strategy that answered our basic evaluation questions, it was also important to acquire samples that were easily accessible. Thus, four naturally formed groups were selected (i.e., quasi-experimental design) which were accessible and met the criteria to answer our evaluation questions. Detailed demographic breakdowns by group can be found on our web-site. These groups were:

Incoming Freshmen. This group consists of incoming Freshmen (enrolled in English Composition courses) who served as a pre-test (i.e., pre-program) group; this group has little or no university influence.

Beginning Teacher Education Majors. These are students who are enrolled in the Introduction to Teaching course. They serve as beginning project students (i.e., the pre-test

teacher education group).

Control Seniors. This group consists of non-teacher education graduates from a variety of disciplines (e.g., nursing, biology, business) on campus. These participants were assessed either through capstone courses or during their senior exit interview process.

Teacher Education Seniors. This group consists of seniors who are in their student teaching phase. They participate in the testing process three times during their senior year (before, during and after student teaching).

### Instrumentation

Assessments used in this study can be found on our website. The following assessments were created and used in this study:

Gender Questionnaire-Revised (GQR). The GQR is an 18 item, 5-point Likert Scale. It measures attitudes towards gender roles (e.g., “Some professions are better suited for men and some are better suited for women”). It has consistently passed tests of item analysis and reliability (with factor analysis confirmation). All reliability estimates have been above .84 and tests of validity are pending.

Educational Gender Questionnaire (EGQ). The EGQ is a 23 item, 5-point Likert scale. It measures gender attitudes in an educational setting (e.g., “boys tend to need more praise than girls”). It has consistently passed tests of item analysis and reliability (with factor analysis confirmation). Reliability estimates have consistently been in the high .80's and validity tests are pending.

Multicultural Questionnaire-Revised (MCR). The MCR is a 30 item, 5-point Likert scale. It measures attitudes toward national, international and educational diversity issues (e.g., “I

almost always try to understand customs of diverse cultures;” “World interests tend to be more important than the interest of my country;” “I feel that stressing different ethnic customs and traditions in public schools tends to reduce learning the basics”). The MCR has consistently achieved reliability estimates above .81. Validity tests are pending.

Multicultural Content Test-Educational (MCCT-E). The MCCT-E was designed to estimate a respondent’s level of MCD knowledge as related to teacher education. The MCCT-E contains 60 multiple choice items divided into 11 specific content area subscales (some items are used on more than one subscale). To meet test administration constraints, the test was bisected into two forms of relatively equal length and difficulty (Form A and Form B). Each form contains different sub-tests.

Multicultural Content Test-General (MCCT-G). The MCCT-G is still in an experimental state. Currently, it consists of two forms (A and B) each containing 27 multiple choice items measuring multicultural/diversity knowledge as related to a student’s general education experience. Although items on the MCCT-G have passed tests of item analysis (i.e., item difficulty, item discrimination and distractor analysis), the test is not a comprehensive assessment and is considered experimental at the time of this report.

Preparedness Survey. This is a 10-item survey designed to assess how well students feel they were prepared by ESU to deal with ten specific populations in a classroom.

MCD Lesson Plan. Students create “a lesson plan that delivers a multicultural/diversity objective while employing an inclusive teaching strategy or strategies for a diverse classroom.” Students must submit their lesson plan along with responses to five questions (e.g., “What makes this an example of an inclusive teaching strategy for a diverse classroom?”). Evaluation rubrics

have been constructed to examine student responses and are available on our website. Rubric factors include: Lesson Plan Objectives, Lesson Plan Mechanics, Lesson Plan Rationale, and Lesson Plan Inclusiveness.

Performance Vignettes. It is not feasible to observe all student teachers throughout their entire student teaching experience. Thus, simulated performance vignettes (in accordance with 6 of the 7 areas identified in Table 2) were constructed to estimate student reasoning ability. Focus groups consisting of various professionals from ESU and the local school district met to discuss the parameters of each of these assessments areas (i.e., What skills should graduates possess and how do we know they have attained these skills?). Small groups (one for each area) processed the focus group information to create vignettes and their corresponding rubrics. Several of the vignettes have been re-created on tape in a classroom setting consisting of middle school students.

Diversity Assessment. The Diversity Assessment is submitted as a part of the teacher education student's application to Phase 2 (their senior year of student-teaching). It is mainly used to ensure that the student has had some degree of experience in a diverse setting during prior placement (phase). If they have not, they will still have a chance to gain such an experience during Phase 2. This assessment not only gauges the quantity and quality of their diversity exposure, but also examines affective, knowledge and performance attributes. The Diversity Assessment requirement and accompanying rubrics are posted on our website.

Table 3 summarizes the characteristics, developmental stage and instrumentation issues for each assessment. Caution is exerted against using any single assessment as a sole means of evaluation for a program or an individual student.

Table 3. Instrument Summary Table.

<b>Instrument(s)</b>	<b>Status</b>	<b>Characteristics/Issues</b>
GQR/EGQ/ MCR	Revisions Complete	<ul style="list-style-type: none"> <li>• Achieved high reliability estimates.</li> <li>• Some items may represent knowledge of facts rather than attitudes.</li> <li>• All scales needs to undergo further validation studies.</li> <li>• Performance is susceptible to Halo and Ceiling Effects.</li> <li>• Interpretation of results require a reference point (e.g., norming sample, pre-post test gain, groups from other universities).</li> </ul>
MCCT-E	Revisions Complete	<ul style="list-style-type: none"> <li>• Several subtests lead to increased diagnostic specificity.</li> <li>• Test construct map was used for content validity.</li> <li>• Significant correlations exist between subtest scores which suggest construct validity.</li> <li>• Subtests are related to other assessments used in this evaluation (e.g., performance vignettes, MCD Lesson Plans).</li> <li>• Further forms of validity are needed.</li> <li>• Not all areas of importance to MCD are included or possible on a test of manageable length.</li> </ul>
MCCT-G	Incomplete- Further Development and Revisions Needed	<ul style="list-style-type: none"> <li>• A construct map is needed to produce greater content validity.</li> <li>• Additional faculty input is needed to create a more comprehensive assessment.</li> <li>• Items do not equally represent all disciplines of interest to a general education program.</li> </ul>
Preparedness Survey	Revisions Complete	<ul style="list-style-type: none"> <li>• Appears to have a moderate degree self-evaluation bias from respondents.</li> <li>• Identifies general populations rather than individual needs.</li> </ul>
MCD Lesson Plan	Revisions Complete	<ul style="list-style-type: none"> <li>• Task is a true performance assessment.</li> <li>• Distinguished inter-rater reliability.</li> <li>• Content validity based on multiple perspectives of personnel used in task construction and evaluation.</li> <li>• Related to some MCCT-E items.</li> <li>• Produces feedback that is "teachable."</li> </ul>
Performance Vignettes	Scenarios Complete, Rubrics need Revision	<ul style="list-style-type: none"> <li>• Realistic classroom scenarios are brief enough for use as an assessment, but contextual enough for potentially rich responses.</li> <li>• Results susceptible to motivation error; student accountability for responses required for use as a reliable, valid measure.</li> <li>• Rubrics need to be further developed based on data from "motivated" subjects. Current rubrics are designed to be used in classrooms where guided feedback from instructor is possible.</li> </ul>
Diversity Assessment	Revisions Complete	<ul style="list-style-type: none"> <li>• Designed for use in high-stakes placement situations.</li> <li>• Susceptible to motivation error if student is not aware of importance.</li> </ul>

## Design and Procedure

It would be impossible to give all the instruments to all the students considering our target sample is 100% of our student population. Thus, a stratified sampling approach was chosen. In this approach, instrumentation is randomized so that all instruments are given, but not all students complete all instruments (i.e., not all students within a group take the same set of instruments). Therefore, it is possible to administer many different instruments to any one group without exhausting student performance--or patience. No names are required on any of the assessments except the lesson plan and diversity assessment requirements. All pre-test to post-test comparisons were made by examining group means instead of tracking individual students from pre to post test time (i.e., a cross-sequential design). All instruments were counter-balanced prior to administration.

## **Evaluation Question #1**

*What is the effect of the Teacher Education program on student attitudes on MCD issues?*

This evaluation question involves the comparison between the Beginning Teacher Education Majors and the Teacher Education Senior groups. Respondents from both groups completed the EGQ, MCR and GQR attitude surveys. All 4 semesters of data were combined for each group for the analysis.

## Results

An analysis of variance (ANOVA) between gender indicated no significant differences on any of the attitude scales. Similarly, no statistical significance was found on any of the attitude scales between elementary and secondary education students or between traditional and non-traditional (i.e., >25 years of age) students.

An ANOVA showed statistically significant differences between Beginning Teacher

Education Majors and Teacher Education Seniors on the GQR [ $F(1, 415)=3.91, p<.05$ ], the MCR [ $F(1, 460)=9.58, p<.01$ ] and the EGQ [ $F(1,436)=12.45, p<.01$ ]. Teacher Education Seniors had statistically higher means on all scales (see Table 4). This was most likely due to the large sample sizes of these groups and, thus, the differences were not considered practically significant.

---

**Table 4.** Means, standard deviations and sample sizes for MCR, GQR, and EGQ total scores between Beginning Teacher Ed Majors and Teacher Education Seniors.

<b>Scale</b>	<b>Beginning Teacher Ed Majors</b>			<b>Teacher Education Seniors</b>		
	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>N</u></b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>N</u></b>
<b>MCR</b>	78.02	10.97	227	81.38	12.27	235
<b>GQR</b>	44.20	7.71	216	45.69	7.70	201
<b>EGQ</b>	59.96	8.35	221	63.27	11.13	217

*Note:* Total score ranges for each scale are as follows: Gender Issues Questionnaire Revised (GQR) (18-90) Multicultural Questionnaire Revised (MCR) (30-150) Educational Gender Questionnaire (EGQ) (23-115) Higher scores indicate more positive attitudes.

---

### Conclusions and Limitations

Teacher education students do appear to develop slightly more positive attitudes (as measured by the assessments used in this study) as they progress through their academic program. However, these differences were statistically significant only due to the large sample sizes in the study. The differences were not practically significant as indicated by the mean differences in Table 4. These results need to be interpreted with caution. There may be several confounding effects which obscure programmatic influences. First, attitudes typically remain

fairly stable over time unless there is a lengthy, positive experience between the respondent and the target groups. Second, programmatic effects compete with influences outside of the campus environment (e.g., media, personal experiences, family, friends). Third, attitude survey items of the Likert type may be vulnerable to the halo effect even in instances where respondents remain anonymous. This “halo effect” occurs when the respondent answers in what he or she believes is a socially desirable way. This tends to inflate attitude scores and can create a ceiling effect which may obfuscate pre-post score gains. This effect may have occurred with the MCR because the mean score fell significantly higher than the scale midpoint. This was not the case, however, for the gender based assessments (i.e., the GQR and EGQ) where mean scores fell close to the scale midpoint. This occurrence may suggest that gender issues are more susceptible to open stereotypical beliefs than are issues of race and ethnicity. However, further research is warranted before this conclusion can be asserted.

## **Evaluation Question #2**

*Do graduating student-teachers have the necessary knowledge to function in a diverse classroom/society?*

This evaluation question involves the comparison between the Beginning Teacher Education Majors and the Teacher Education Senior groups on the MCCT-E. Both forms of the MCCT-E have evolved over the course of the project. Test items have been revised based on item analysis statistics (i.e., item difficulty, item discrimination, and distractor analysis). The results presented below are from the final version of the MCCT-E administered in the Spring semester of 2000.

### **Results**

An ANOVA showed no statistically significant difference between groups on either forms

of the MCCT-E. Table 5 depicts the descriptive results.

---

**Table 5.** Descriptive statistics for the MCCT-E (forms A and B) between Intro Teacher Education Majors and Teacher Education Seniors.

	<b>Form A (35 items)</b>		<b>Form B (33 items)</b>	
	<i>Beginning Teacher Ed Majors</i>	<i>Teacher Education Seniors</i>	<i>Beginning Teacher Ed Majors</i>	<i>Teacher Education Seniors</i>
N	32	41	29	36
M	21.97	25.23	21.41	21.40
SD	4.96	4.19	3.78	4.11
SE	0.88	.65	0.70	0.72
Kurtosis	-0.67	.52	-0.82	3.46
Skewness	-0.43	-0.93	-0.01	-1.16

Note #1: Only students completing the entire test were included in the statistics.

Note #2: N=Number of Respondents

M=Mean Total Score

SD=Standard Deviation SE=Standard Error

---

### Conclusions and Limitations

Although the MCCT-E was based on a construct map which used a variety of professional sources (i.e., it has face content validity), the test may only partially reflect actual course content since several outside sources were used in the test construction process (as indicated on the construct map available on our web-site). This may have also been reflected in the normalized distributions found in total scores. Further examination of course content and method are required to improve course content/test content alignment. Another possible alternative explanation could be that covering content knowledge in a class may not be enough for

retainment. That is, the knowledge may have to be used in the development of teaching strategies and philosophies to improve retainment for summative evaluations. There may also have been sampling bias error. Seniors completed the test during a day-long series of events and often in hot classroom conditions within large groups. Beginning Teachers completed the test during their normal course time in a much more formal and intimate environment. This possible fatigue effect for seniors may have adversely affected their scores. Similarly, the more intimate environment for the Beginning Teachers may have reduced their perceived anonymity and these students may have been more motivated to think out their responses rather than engage in guessing. These testing conditions have implications for those wishing to replicate our research design at their institution.

### **Evaluation Question #3**

*Are student-teachers gaining the necessary skills to function in a diverse classroom/society?*

Teacher Education Senior performances on the MCD Lesson Plan, Performance Vignettes Diversity Assessment and Preparedness Survey were used to address this evaluation question.

#### **Results**

*Lesson Plan.* A four factor rubric (Objectives, Mechanics, Rationale, Inclusiveness) was constructed and tested on the lesson plans. A four point scale (Incomplete, Unsatisfactory, Developing, Proficient) was used for each factor. Inter-rater reliabilities for each rubric factor ranged from .83 to .89. The rubric as well as holistic and analytical exemplars for each rubric level, and for elementary and secondary education are provided on our web-site.

In order to better interpret the rubric and evaluation results, a brief narrative description of each rubric factor is included with the results. Factor 1 (Lesson Plan Objectives) evaluated a student teacher's ability to design a meaningful multicultural lesson. Low scores usually

indicated that the student omitted a multicultural objective or used a trivial objective (e.g., food day, or making beads without a context). Often the multicultural component appeared to be an add-on to an existing plan just to meet the requirement. Higher scores were awarded if MCD objectives were more central to the lesson, engaged attitudes and beliefs, and went beyond factual information. At the proficient level, the lesson plan synthesized personal reflection, addressed developmental adaptations, and included skills relevant outside the classroom. About half the student teachers addressed meaningful MCD objectives. Of those, about one-third used critical thinking strategies. The focus on facts and use of trivial objectives were main causes for low scores.

Factor 2 (Lesson Plan Mechanics) evaluated the ability to synthesize lesson plan objectives with classroom activities and assessments. Low scores represented weak or non-existent links between these components. This usually occurred because student teachers omitted assessment, assessed informally, evaluated by single means, or with a select group of their students. High scores represented stronger, multiple, and inclusive activities and assessments. About half of the students earned developing or proficient scores, mainly because they did not assess beyond the knowledge level.

Factor 3 (Lesson Plan Rationale) evaluated a student's ability to competently explain why their lesson plan was multicultural and inclusive. Low scores indicated the misconception that a MCD lesson employed culturally stereotyped activities (e.g., making teepees). While "developing" scores acknowledged diversity, "proficient" scores included societal perspectives. Again, results indicated that about half of the students earned scores in the developing and proficient range. The use of stereotypical activities was the major reason for low scores on this factor.

Factor 4 (Lesson Plan Inclusiveness) evaluated the ability to provide inclusive teaching strategies, especially to individuals whose primary language was not English. Low scores conveyed no adaptations or the single approach of: "Get an ESL tutor." Lesson plans with a "developing" score reflected more responsibility for addressing student needs, typically with varying instructional strategies. A "proficient" lesson plan promoted the classroom as a community of learners with multiple learning styles. Only a small proportion of lesson plans achieved proficiency.

Results indicated no practically significant differences on rubric factor mean scores (no differences exceeded .50) between testing times (e.g., Spring semester vs. Fall semester), or between emphasis areas (elementary vs. middle/secondary). Generally, rubric scores were approximately normally distributed for all rubric factors. However, since this is a criterion-referenced assessment (i.e., there is a benchmark score) it is expected that students would score in the developing or proficiency range on all rubric factors. In sum, approximately half of all respondents scored at or above this competency level on each rubric factor.

In summary, results indicated that half our student teachers demonstrated minimal skills (or above) in creating multicultural/diversity lesson plans. Common themes contributing to low scores included: approaching the lesson plan requirement by "adding on" a multicultural component as an afterthought, omitting assessment, limiting assessment options, targeting learning at knowledge (factual) levels, using trivial or stereotypical activities, and not personally addressing ESL needs. Sample scores differ slightly from those obtained from 40 "multicultural lesson plans" randomly selected from an internet search. While project lessons had a slightly higher mean score for "mechanics," internet samples showed slightly higher mean scores for "objectives" content and, especially, for "inclusiveness" factors (1.5 point mean difference).

Omitting or using limited assessments was the only deficit consistently found in both samples.

*Preparedness Survey*. The preparedness survey is a self-report measure of how well students feel they are prepared to deal with 10 specific classroom populations and should be related to the scores on the MCE Lesson Plan. Table 6 presents the means and standard deviations from 4 semesters of surveying Teacher Education Seniors. Participants were surveyed prior to and immediately after their student teaching experience.

---

**Table 6.** Item means and total score means and standard deviations for responses to the preparedness survey.

“Considering all my course work at ESU to this point, I feel \_\_\_\_\_prepared to deal with:”

	Well Prepared (WP) 1	Adequately Prepared (AP) 2	Ill-Prepared (IL) 3	Not Prepared (NP) 4
	<b>Pre/Post-Student Teaching</b>			
<b>Students...</b>	<b><u>M</u></b>	<b><u>SD</u></b>	<b><u>N</u></b>	
Diverse Classrooms	1.67/1.57	.58/.57	483/306	
Students with physical challenges	2.00/1.88	.71/.70	483/306	
Students with mental challenges	2.01/1.89	.70/.72	483/306	
Students with behavioral challenges	1.93/1.79	.73/.74	483/306	
Students of various ethnicities/cultures	1.69/1.59	.65/.61	483/306	
Students of opposite gender	1.29/1.23	.50/.47	483/306	
Students with different learning styles	1.33/1.29	.55/.51	483/306	
Students from single parent families	1.57/1.46	.71/.63	387/306	
Students with same gender parents	2.24/2.17	.98/.98	387/306	
Students with multi-racial parents	1.72/1.55	.78/.67	387/306	
<b>Scale Totals</b>	16.33/16.41	4.98/4.65	387/306	

Note 1: Scale range is 10-40.

Note 2: Sample numbers are smaller for the last three items because these were added after the first semester of testing. Scale totals include only those participants who received the full 10 item scale.

---

Results from the preparedness survey indicate that students feel they are more prepared to deal with diverse classroom populations than indicated on their performances to the MCD

Lesson Plan. There were no appreciable differences between group ratings before and after student teaching. The students felt least prepared to deal with students with same gender parents.

*Diversity Assessment.* The diversity assessment is collected every semester from students prior to entering their student teaching phase (although only one semester was analyzed for this report).

One purpose of the diversity assessment is to ensure that students who have not had adequate experience working professionally with diverse populations get student teaching appointments in diverse classroom settings. The diversity assessment also estimates a student's concept of diversity and their knowledge about teaching strategies that promote inclusive learning environments. Approximately 70% of the respondents (n=44) provided some evidence that they had a comprehensive definition of diversity, and some knowledge about employing a variety of inclusive strategies and creating a learning climate. Overall, students were less able to provide a connection between their philosophy and their suggested pedagogical strategies. Overall, the students who completed this assessment had contributed numerous hours towards fulfilling their diversity requirement. Some had worked three hours to as many as 40 hours per week with diverse populations. Prevailing themes in the data were scarce, and many students did not seem sure how their experiences related to their own philosophy of teaching. The essays did contain pedagogical examples, but in some instances, they were unrelated to the student's statement of philosophy. While several student were able to demonstrated a high level of proficiency, some essays still contained examples of stereotyped and "tourist" approaches to teaching. This result was consistent with the results obtained on the MCD Lesson Plans.

*Performance Vignettes.* The performance vignettes are 6 separate scenarios designed to match the areas listed in Table 2 (a vignette was not created for the seventh area--family). The vignettes

were administered to Teacher Education Seniors for 4 semesters. Each semester, revisions were made to the vignettes based on feedback from students, outside consultants and the ESU Project 30 committee. Final versions of the vignettes can be found on the web-site. Paid consultants were used to create a rubric for each vignette. Several approaches were attempted including looking at the number of problems and solutions identified (problem fluency), the number of different problem and solution areas identified (problem flexibility) and the quality of the problems and solutions to the conflicts depicted in the scenarios. Although some common themes were identified (e.g., respondents tended to focus on non MCD problems such as general classroom management issues instead of identifying the MCD conflict issues), the testing situation did not lend itself to producing good qualitative data. First, the anonymous nature of the assessment (which was a requirement of the grant) tended to reduce the motivation of the respondents. The result was laconic (written) responses. Second, to compound the brevity of response problem, we did not have an opportunity to query students about their responses. One essential component in using vignettes is that there is a dyadic relationship between the respondent and the administrator which allows the respondent to clarify their answers and provide additional rationale to follow-up questions. This shortcoming of written responses to cognitive tasks was first demonstrated by cognitive development studies. Children who score low on pencil/paper tests of cognitive development tend to score significantly higher when the researchers have a chance to probe and push the respondent to their upper potential limit. For this reason, no results from rubrics are presented in this report. To make this a viable programmatic evaluation tool, these vignettes are being used in our teacher preparation courses so that instructors can challenge students and provide guidance and feedback to their responses.

### Conclusions and Limitations

Overall, the students did fairly well given their experience level as student teachers and the resources available to them. It would be surprising if most new graduates performed at the "proficient" level. Graduating teachers need more experience to perfect the skills necessary to serve diverse populations and build inclusive curriculum. As with most teacher preparation programs, we are driven by state requirements for licensure and national guidelines for accreditation. Currently, there is no ESL/LEP requirement for certification in Kansas, nor is there a specific indicator in national guidelines. Given that reality, and the packed schedules of most students, it is unlikely for students to take an ESL course unless they are pursuing an ESL concentration. This could help explain the low scores on adapting lesson plans for children whose first language is not English or for children with disabilities. Most of the student teachers are assigned to settings with children from diverse backgrounds; but with the popularity of "pull out" programs, they may not observe classroom accommodations. Or, accommodations may come so naturally to a veteran teacher that they are not communicated to the student teacher. However, national accreditation guidelines do have requirements for inclusion and multicultural education. Because our program infuses multicultural education and inclusion strategies, no separate multicultural education course is required, and the required survey of exceptionalities course does not include methods. The results of this project may suggest that our students do not learn to use strategies for adaptation of instruction for children with special needs.

Some of these scores may be explained by the demands placed on the student teacher. In addition to all the classroom responsibilities inherent in this capstone semester, students must also complete forty assignments on public school policies and procedures, including the multicultural lesson plan assignment. Admittedly, most student teachers may focus their energies

on classroom success. Likewise, university supervisors may focus their attention on classroom performance, rather than on successful completion of forty assignments. Thus, lesson plans received varying attention by the student teachers. Some conscientious students did quite well, while others just made the attempt. We expect scores to improve once students have access to the rubric to use as a learning tool and as a benchmark to judge potential classroom resources (e.g., internet, resource books, etc.).

### **Evaluation Question #4**

What effect does the general education program have on student's attitudes and knowledge regarding Multicultural/Diversity issues?

This evaluation question involves the comparison between Incoming Freshmen and Control Seniors. Respondents from both groups completed the MCR and GQR attitude surveys as well as both forms of the MCCT-G. The MCCT-G has been revised over the course of this project and is still not considered a comprehensive content knowledge assessment at the time of this report. Only results from the most complete version are presented (i.e., the final semester of the project).

#### Results

Control Seniors significantly outscored Incoming Freshmen only on Form A of the MCCT-G [ $F(1, 57)=4.87, p<.05$ ]. Descriptive statistics on both forms of the MCCT-G are presented in Table 7. In contrast, there were no significant differences between Control Seniors and Incoming Freshmen on the GQR and MCR attitude assessments. Table 8 presents means and standard deviations by group on these two assessments.

Table 7. Descriptive statistics for the MCCT-G (forms A and B) between Control Freshmen and

Control Seniors.

	<b>Form A (27 items)</b>		<b>Form B (27 items)</b>	
	<i>Incoming Freshmen</i>	<i>Control Seniors</i>	<i>Incoming Freshmen</i>	<i>Control Seniors</i>
N	18	41	23	43
M	10.83	13.12	13.48	14.86
SD	4.00	3.52	3.57	3.32
SE	0.94	0.55	0.74	0.51
Kurtosis	-0.15	1.03	-0.47	1.22
Skewness	0.01	-0.05	0.02	-0.60

Note #1: Only students completing the entire test were included in the statistics.

Note #2: N=Number of Respondents

M=Mean Total Score

SD=Standard Deviation SE=Standard Error

**Table 8.** Means, standard deviations and sample sizes for MCR and GQR total scores between Incoming Freshmen and Control Seniors (4 semesters).

<b>Scale</b>	<b>Incoming Freshmen</b>			<b>Control Seniors</b>		
	<u>M</u>	<u>SD</u>	<u>N</u>	<u>M</u>	<u>SD</u>	<u>N</u>
<b>MCR</b>	78.42	10.89	165	76.97	11.74	60
<b>GQR</b>	44.18	6.94	181	43.28	6.50	71

*Note:* Total score ranges for each scale are as follows: Gender Issues Questionnaire Revised (GQR) (18-90)  
Multicultural Questionnaire Revised (MCR) (30-150). Higher scores indicate more positive attitudes.

### Conclusions and Limitations.

It is not possible to make any program evaluation conclusions or recommendations based on the results because the MCCT-G is in the very beginning stages of development at the time of this report. It should be noted that seniors did make slight gains in total scores on the MCCT-G. However, further analysis of course content and faculty involvement are necessary before the MCCT-G can become a useful tool for program feedback. Only slight gains were found between

pre and post assessment times for the attitude surveys. Again, the stability in attitudes over time without positive immersion experiences with target populations could have been a competing factor with classroom experiences and instruction. Furthermore, it is necessary to examine course content to identify practical immersion experiences and classroom activities that specifically address the formation of healthy, productive MCD attitudes. These assessments have indicated a need for such exploration and remediation as the next step in the development of MCD sensitive graduates at ESU.

## Summary

This evaluation study sought to examine the effect of course work on students' attitudes, knowledge and skills with respect to a set of identified multicultural/diversity related outcomes. Results, as measured by the instruments presented in this study, indicate modest changes in these outcomes. This is not to say students leave ESU with poor attitudes and lacking knowledge and skills. Rather, the attitudes, knowledge and skills adopted as essential for the purposes of the project appear to be relatively steady throughout a student's college experience. Much of these seemingly tenebrous results could be due to the lack of refinement of our assessments. Creating evaluation instruments and using them for actual evaluations is much like building a plane while in flight. In addition, the population of students should be considered when interpreting results. That is, students from more a homogeneous background, such as many of those who attend ESU, may be more resilient to change. Finally, if we had completely developed all our instruments and did so to a high degree of reliability, we must still recognize that evaluation is designed as an impetus for growth.

The benefits of this project are continually emerging. At the end of three years, there is a set of standard performance assessments (including rubrics) for knowledge, skills, and

dispositions in multicultural education available for public use. To the knowledge of these authors, scant work has been done on a programmatic (i.e., university-wide) level to examine the influence of pedagogy and curriculum on performance within a diverse setting. Furthermore, multicultural/diversity assessments are usually quantitative in nature (e.g., attitude scales, self-reports on knowledge/abilities), avoiding the tough question “can our graduates *do it*.” Thus, this project is unique in that it is looking at attitudes, knowledge, and skills at the programmatic level and creating meaningful performance-based assessments keyed to the needs identified by practitioners in the field. These performance assessments could be a catalyst for change in teacher education programs. They also could be applied as a self-study vehicle for teachers in the field. At Emporia State University, the results of these assessments will be used to analyze how well the courses have prepared our students to meet the eight outcomes. It is likely that both the outcomes as well as the teacher preparation program will be altered as indicated by the analysis. Any teacher preparation program, or public school system for that matter, undergoes fine tuning and redesign on an ongoing basis. Perhaps these standard performance assessments will help direct part of that process.

The end result, of course, is the improvement of the quality of classroom experiences in our elementary and secondary schools. New teachers will step into the situations like those we presented in our vignettes, armed with the confidence and ability to serve students and their families better because they have been in teacher education programs which use sophisticated and comprehensive diversity performance assessments for feedback and growth..

## **Recommendations**

At the conclusion of this project, the following recommendations are made:

1. University staff need to focus on the ecological validity of the assessments. The vehicle for this should be sub-committees used to refine the evaluations and maintain contemporarily valid assessments.
2. The emphasis of time and resources should dramatically shift to the creation of performance-based assessments including student-teacher observations. This is important to establishing construct validity as assessment scores can be correlated across instruments.
3. University faculty and staff should identify and pursue the mechanisms of changing courses to meet the program learning outcomes.
4. All endeavors described in the above recommendations should include additional faculty public school, and community involvement, participation and input.
5. All results from this evaluation should be interpreted with caution. Validation of instruments often occur later and take some time.
6. For the evaluation instruments to produce useful data, they must be made into program requirements. If students do not have a stake in the assessments (such as being anonymous in their responses), they will most likely produce laconic responses and incomplete products. Thus, respondent motivation is essential to the creation of instruments and requirements indicate the sincerity of our desires to promote MC/Diversity proficiencies.



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



## REPRODUCTION RELEASE

(Specific Document)

### I. DOCUMENT IDENTIFICATION:

Title: <i>Final Report on the Multicultural/Diversity Assessment Project</i>	
Author(s): <i>Anthony L. Ambrosio, Ph.D.</i>	
Corporate Source: <i>Emporia State University</i>	Publication Date: <i>2-6-01</i>

### II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

The sample sticker shown below will be affixed to all Level 2A documents

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

*Sample*

\_\_\_\_\_

\_\_\_\_\_

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

**1**

Level 1



Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

*Sample*

\_\_\_\_\_

\_\_\_\_\_

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

**2A**

Level 2A



Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

*Sample*

\_\_\_\_\_

\_\_\_\_\_

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

**2B**

Level 2B



Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.*

**Sign here, → please**

Signature: <i>Anthony Ambrosio</i>	Printed Name/Position/Title: <i>Anthony Ambrosio, Director</i>	
Organization/Address: <i>1200 Commercial St. Box 4036 Emporia, KS 66801</i>	Telephone: <i>341-5093 (620)</i>	FAX: <i>(620) 341-5185</i>
	E-Mail Address: <i>ambrosia@</i>	Date: <i>5-31-01</i>

*emporia.edu*

(over)

### III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

### IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

### V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Teaching  
and Teacher Education  
1307 New York Ave., NW  
Suite 300  
Washington, DC 20005-4701

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

**ERIC Processing and Reference Facility**  
4483-A Forbes Boulevard  
Lanham, Maryland 20706

Telephone: 301-552-4200

Toll Free: 800-799-3742

FAX: 301-552-4700

e-mail: [ericfac@inet.ed.gov](mailto:ericfac@inet.ed.gov)

WWW: <http://ericfac.piccard.csc.com>