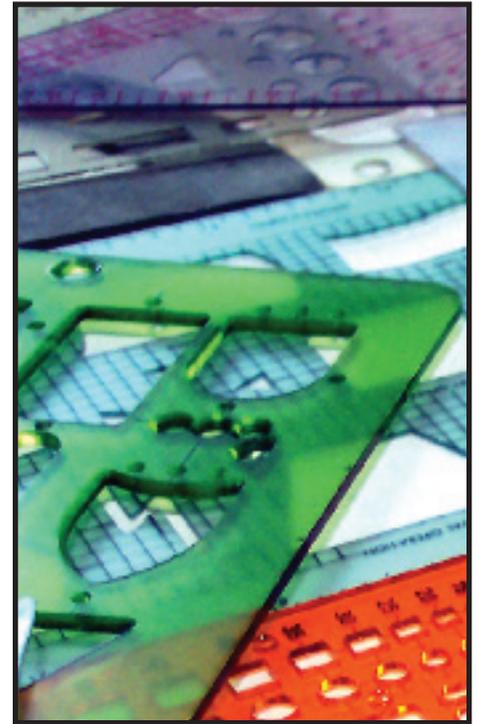# MEASUREMENT AND ASSESSMENT

CONVERSATIONS WITH PROFESSIONAL PEOPLE OF COLOR IN THE FIELD OF EDUCATION

ARETHA FAYE MARBLEY,
FRED BONNER
& RACHELLE BERG

Introduction

With the increased focus on issues of diversity and multiculturalism in the United States and abroad, educators and clinicians have become more and more concerned about the feasibility and validity of their assessment instruments and practices for children and adults from nonWhite backgrounds. Many researchers (e.g., Brown, Lipford-Sanders, & Shaw, 1995; Sue & Sue; 2003) emphasize that when working with racially diverse groups in multiple settings, issues relating to their cultures often differ from those of the dominant group and thus warrant different approaches and considerations. Therefore, the need to foster a multicultural perspective within educational and psychological assessment that embraces more than just test instruments is evident and paramount (D'Andrea, Daniels, & Heck, 1991; Ivey, Ivey, & Simek-Morgan, 1993; Johnson, Torres, Coleman, & Smith, 1995).

For people of color, the negative and harmful effects of assessment practices reaches far beyond mere measurement and focuses on more critical issues than just

Aretha Faye Marbley is an associate professor in the College of Education at Texas Tech University, Lubbock, Texas; Fred Arthur Bonner, II, is an associate professor in the College of Education at Texas A&M University, College Station, Texas; and Rachelle Berg is a doctoral student in counselor education at Texas Tech University.

testing bias. For people of color, a cultural approach to assessment must begin with a dialogue on multicultural standards, competencies, guidelines, training, and with the relevant profession for which the education and assessment is intended.

Thus, a multicultural approach necessitates a moral and ethical responsibility for higher education institutions to be proponents of diversity and to prepare culturally sensitive professional educators to work with the diverse groups in our society.

As a result of this call to professions to embrace a sense of ethical and moral responsibility to people of color, researchers from multiple disciplines in education have reiterated the importance of development of culturally appropriate and culturally sensitive assessment instruments as well as the development of ethical standards and training guidelines to prepare competent professionals who choose to work with culturally diverse people (e.g., Krishnamurthi, 2003; Lindsey, 2000; Suzuki, 2001; Wall & Walz, 2003).

The Impact of Assessment, Measurement, Testing, and Research on Marginalized People

This need for a cultural approach in educational assessment also reaches beyond a mere awareness and knowledge of racial and ethnic differences to encompass broader areas such as assessment, practice, training, ethics, and research related to working with people from diverse

backgrounds, specifically people of color (Miller-Jones, 1989; Ridley, Hill, & Wiese, 2001; Roysircar, Sandhu, & Bibbins, 2003; Suzuki, Ponterotto, & Meller, 2001; Valencia & Suzuki, 2001). A prime example is professionals acquiring competencies and culturally sensitive skills necessary to work with diverse groups in multiple settings, where issues relating to their race, ethnicity, and culture often differ from those of the dominant group, and thus may warrant different approaches and considerations than those of the dominant group (Roysircar et al., 2003).

A cultural assessment approach must also include the appropriate and responsible use of the dissemination of research and measurement instruments, including their selection, administration, scoring, interpretation, and communication, as well as recognition of the preponderance of bias existing within the practice of the field of educational measurement and other disciplines within the academy.

Therefore, a comprehensive understanding of the importance of the role culture plays in educational assessment requires examining assessment in light of its importance to race, culture, ethnicity, gender, and other diversity variables. In addition, there needs to be input on the bias and the negative impact of assessment on people of color from multiple sources and from a variety of disciplines within education, such as counselor education, higher education, educational psychology, teacher education, special education, educational technology, and educational leadership.

Cross-Cultural Assessment: A Phenomenon with Deep Pockets

In an overview of cross-cultural assessment, we offer a simple, but comprehensive explanation of various terms such as test, measurement, and assessment. This is a necessary prerequisite to a dialogue on multicultural assessment, as well as a starting point for a broader understanding of culture in assessment. According to Kaplan and Saccuzzi (1993), a test is a measurement tool that quantifies behavior, whereas measurement is the process by which things are differentiated (Hopkins & Stanley, 1990). In contrast, "assessment is used to evaluate an individual so that he or she can be described in terms of current functioning and also so that predictions can be made concerning future functioning" (Kaplan & Saccuzzi, p. 219). This means that tests are merely one method of assessment and one part of the assessment process.

Unfortunately, in American society, for people of color the process of measurement has often been used to discriminate against groups. In fact, members of marginalized groups are often assessed unfairly when their behavior, feelings, thoughts, or experiences deviate from those of the Anglo-Saxon, heterosexual, able-bodied, middle-class, White male standard (Feagin, Vera, & Batur, 2001; Robinson, 2005; Scheurich, & Young, 1997).

Historically, in different arenas of measurement and testing, culturally diverse groups have been unfairly and disproportionately labeled as dysfunctional, abnormal, of low intelligence, mentally ill, or dangerous to society. Such discrimination has been both costly and harmful to people of color and other marginalized groups, as well as to the well-being of society as a whole, and thus definitely warrants further consideration.

Another important area to examine is the damage caused by biased instruments and racist ideologies, and how this prejudice has encouraged the misuse and abuse of the entire testing process (Sabatier, 2003; Scheurich & Young, 1997; Wyche & Novich, 1986). The unfair treatment of marginalized groups due to biases toward individuals who are culturally different and the resultant biases of the measurement instruments we employ, leaves us—as professionals—in a precarious ethical quandary.

We know that if advancement is to be made in assessment, the research focus must be on developing measurement instruments that will be free of bias and will assure fairness for all cultural groups. Concomitantly, there is a charge to those disciplines and individuals involved in testing, measurement, and assessment to not only develop unbiased instruments, but also instruments that can accurately measure multicultural awareness, skills, sensitivity, and competency of the individuals involved in and in charge of assessment.

From the counseling and education literature we know that a necessary requisite of multicultural assessment in counseling and therapy is self-assessment (Dana, 2000; Ivey et al., 1993; Roysircar et al., 2003; Sue, Arredondo, & McDavis, 1992). Self-awareness and self-assessment co-exist and are not separate entities. According to Greeley, Garcia, Kessler, and Gilchrest, (1992), self-awareness starts with self-assessment, that is, knowledge of one's racial or cultural identity development.

A critical component of self-assessment is to acknowledge cultural differences and assess whether you as a practitioner feel comfortable discussing cultural, ethnic, or racial issues (Ivey et al., 1993; Sue & Sue, 2003). Pope-Davis and Ottawi (1994) concur that contrary to the experiences of most people of color, White Americans are rarely called on to assess their own attitudes about their White ethnicity. Therefore, to correct this historic pattern, all individuals involved in assessment should become self-aware, including an awareness of their own potential biases towards people from different cultures.

According to Greeley and colleagues (1992), a starting place for this assessment is with instruments created to assess stages of racial identity development. There is a developing awareness that some measure of cultural identity is essential for most counseling situations (Brown et al., 1995; Gainer, 1992; Harris, 2003). Rollock, Westman, and Johnson (1992) feel that White professionals who do not increase their cultural knowledge will have propensities toward over-pathologizing and thus undermining therapeutic effectiveness. Similarly, nonWhite professionals working with White clients in a counseling situation who do not possess and understand their client's cultural knowledge run the same risks of therapeutic ineffectiveness.

A Call to Action

A clear example of the continued disparate impact that educational measurement has exacted on people of color can be found in the work of Eeels, Davis, Havighurst, Herrick, and Tyler (1951) from over a half century ago. Although the educational reform and testing craze that we witness today didn't start until the early 1980s, these earlier findings hold relevancy in the area of cross-cultural measurement.

Eeels and colleagues found an apparent contradiction, that is, on the one hand, the instruments used to measure achievement and competence are being more urgently called for and more widely used than ever before whereas, on the other hand, tests are, at the same time, being more sharply criticized and strongly opposed. Unfortunately, this dilemma of balancing the validity of tests and minimizing bias and harm when assessing people of color is still relevant today.

This criticism of and opposition to tests has been fueled largely by multicultural literature identifying the unfair and sometimes harmful effects of racist ideologies, biases, and biased instruments and procedures on marginalized groups, specifically people of color (Burley, Butner, Marbley, & Bush, 2001; Drew, 1973; Gay & Abrahams, 1973; Sabatier, 2003; Scheurich, & Young, 1997). In our study, this criticism is widened to include a discussion of biases that exist with researchers and assessors. Therefore, for real strides to be made in the field of educational assessment and measurement, we must employ an effective, comprehensive approach that reaches beyond the areas of measurement and testing to encompass assessment, practice, training, and research (Miller-Jones, 1989; Roysircar et al., 2003; Suzuki, 2001).

The aforementioned opposition to more testing can be attributed in part to the bias, misuse, and abuse of the measurement instruments used to evaluate people of color (Sedlacek, 1994; Miller-Jones, 1989), but a great deal of the bias is inherent within the practice and the professionals. Therefore, established multicultural competencies and standards must advance the notion that the principal goals for culturally skilled counselors and for other professionals are to become aware of their own assumptions about human behavior, to seek an understanding of their diverse clientele's assumptions about human behavior, and to become active in developing appropriate intervention strategies (Marbley, 2004; Roysircar et al., 2003; Sue et al., 1992; Toporek & Reza, 2001).

In response to the urgent call for unbiased instruments as well as assessment procedures and instruments in the education and mental health fields that can adequately assess practitioners' multicultural competencies, numerous articles and books were published within the mental health arena regarding multicultural training, competencies, and guidelines (Ponterotto, Rieger, Barrett, & Sparks, 1994; Pope-Davis & Dings, 1994;

Roysircar et al., 2003; Sue, Arredondo, & McDavis, 1992).

In addition, professional associations and state licensure boards in some fields have implemented multicultural standards that provide guidelines and a framework for practitioners to become culturally skilled and responsible. For example, the Association for Assessment in Counseling and Education (AACE) Multicultural Assessment Standards stress the need to have assessment instruments that are fair and beneficial to members of all populations. Similar responses have occurred in other disciplines, such as teacher education, educational psychology, and higher education.

We believe that if a broader and more comprehensive understanding of multicultural assessment is to develop, it will take more than a call to the professions. As professionals ourselves, from different disciplines, we believe that a viable approach is to "go grassroots," that is, increase the dialogue with people of color within the relevant disciplines in order to get a better grasp of everyone's view of the key diversity issues within the field of educational measurement and assessment that are found to negatively impact people of color.

Therefore, the purpose of our study is to glean from professional people of color in the field of education their experiences and perceptions of contemporary assessment instruments and practices and the relative influence these practices exert on people of color.

Theoretical Framework

We used Sedlacek's (1994) model to organize the understanding of the cross-cultural assessment literature, as well as to serve as a theoretical framework. In essence, this model is used to: (a) frame the fragmented body of multicultural assessment research literature into a comprehensive body of information concerning the phenomenon of people of colors' experiences with measurement, testing and assessment; (b) identify the critical issues affecting the field; and (c) attempt to explain why such vast numbers of people of color are "falling through the cracks," that is, being adversely impacted by assessment.

In the areas of measurement and assessment, Sedlacek (1994) identified five diversity issues plaguing the field of counseling: (a) what groups should be included, and what terms should be applied to those groups; (b) the development of a single instrument or measure that would be valid for all; (c) studies and research that are not designed for people

of color; (d) the importance of missing or incomplete data and how the lack of that data may bias the sample and invalidate the instruments and the procedures; and (e) a shortage of well-trained professionals working in this area.

Method

The methodology we selected allowed us to pose and subsequently answer the following research questions:

(1) What factors in assessment and measurement are influenced or impacted by multiculturalism?

(2) What are the perceptions of professional educators of people of color's experiences in using assessment and testing services? This includes subquestions of: (a) What key factors are involved in the assessment and testing experiences considered unsuccessful? and (b) What factors are involved in the assessment and testing experiences considered successful?

(3) Through the lens of educators' of color unique academic disciplines, what have they learned from training, teaching, researching, and clinical and practical work with people of color in their perspective fields? Also, how can this collective data be honed in a manner that can be understood as both commonalities and discipline-specific competencies? Finally, how can this information foster and facilitate dialogue on diversity issues among disciplines and among colleges and universities?

Participants

The sample group for this study was purposefully selected using the maximum variation sampling strategy. According to Patton (2002), purposeful sampling aims at capturing and describing the main themes or major outcomes that cut across participant variation. The 14 participants chosen were professional educators: (a) from multiple disciplines in higher education; (b) self-identified as members of one of the four major groups of color: African American, Asian, Hispanic/Latino, and Native American Indian; and (c) had witnessed or experienced bias or biased instruments within the field of education.

This small purposeful sample of gender and ethnically diverse individuals had two advantages: (1) it yielded detailed descriptions of high quality; and (2) it uncovered important shared patterns among peoples of color of their experiences with

measurement, testing, and assessment emerging out of their heterogeneous cultural backgrounds.

For this study, we utilized professional contacts and gathered information through various data banks: State Board of Education information, alumni groups, and professional higher education organizations. These contacts provided 21 possible participants. Sixteen out of the 21 individuals agreed to participate in the study. Of the 14 selected who completed the study, there were six males and eight females and nine African Americans, one Hispanic/Latino, and two Native Americans.

The disciplines represented included counselor education, law, adult education, vocational education, health education, elementary education, higher education, education administration, curriculum and instruction, and history. Nine were students (eight full-time) with extensive professional backgrounds and part time jobs (eight in doctoral programs). Seven participants were administrators, four in higher education institutions and three in public schools. There were two attorneys who both had professional ties to education. The participants' level of education ranged from masters to postdoctorate.

The females ranged in age from 32 to 60, with a mean age of 36.5; the men ranged in age from 25 to 50, with a mean age of 42. In an effort to acknowledge the various cultural and ethnic differences existing within each group, participants were asked questions related to their level of acculturation and their ethnic and racial origin. Other demographic data gathered included educational backgrounds, professional identity, specialty areas, certifications, and employment status.

Design and Procedure

This study was conducted in three parts: (a) a standard open-ended interview; (b) informal conversational interviews; and (c) a focus on the participants' roles as a professional and as a layperson. The study is a qualitative investigation using a phenomenological theoretical orientation focusing on the participants' roles as both client and counselor.

Qualitative data methods consisted of three kinds: (1) in-depth, open-ended interviews; (2) informal conversational and follow-up interviews; and (3) collection and examination of written documents. According to Patton, (2002), the interviews provided people the opportunity to respond in their own words and to express their own personal perspectives about their experiences, opinions, feelings, and knowledge.

The standard open-ended interview

lasted approximately 90 minutes and was audio-recorded. Participants were asked to respond to each item as it related to their professional and personal experience as well as to their areas of study. After the taping, the participants were asked to review and critique their audiotaped interview and to review a transcript from a videotaped session. Some declined and preferred not to review the transcripts. Following the initial interviews, the first named author reviewed, critiqued, and transcribed the audiotape interviews.

Participants were also asked to either provide or review any written documentation regarding their experiences with assessment and measurement instruments (e.g., professional reports, standardized test scores, grades) and journal their reactions or comments to the materials. Lastly, participants were given a diskette and instructed to maintain a journal of their comments, feelings, or reactions to any parts of this study.

Data Analysis

The data were analyzed through the grounded theory approach in order to uncover relevant categories and the relationships among them. With grounded research, we wanted to uncover relevant categories and the relationships among them, that is, to merge categories in new, rather than standard ways (Strauss & Corbin, 1998). The initial codes were based on the broad questions and topical areas guiding the interview, but were expanded to include unexpected information that emerged from the discussion.

Initial coding yielded five pages of open codes (approximately one page per participant), which was further reduced to 12 axial codes using the paradigm model—a tool used to link subcategories to a category in a set of relationships (casual conditions, phenomenon, context, intervening conditions, actions, and consequences) and relate them in complex ways (Strauss & Corbin, 1998). Finally, the selective coding yielded three categories and three subcategories. From the qualitative inquiry, the findings are discussed by categories and compared and contrasted by theme.

Results

We identified three themes that emerged out of the data from participants' perceptions as professionals educators that may have hindered or facilitated some aspect of people of color's experiences with educational testing and measurement in academe. The first theme—*bias*—yielded four sub-topics: instruments, profession-als, practice, and language. The second theme discussed is *cultural congruency and sensitivity,* while the last theme discussed is *proponents of diversity.*

Bias

The *bias* theme is used to describe the participants' perceptions and experiences of discrimination and prejudice due to language, people, practice, and the use of certain instruments. This theme also includes the participants' personal and professional experiences. The first author developed a set of guidelines relating to various aspects of language, professional bias in assessment, administration, interpretation, and practice (which include the manifestation of bias within the various fields and disciplines related to education), and inappropriate abuse and misuse of assessment instruments in order to merge the data. The subthemes of language, people, practice, and instruments emerged from these guidelines and were subsumed under the bias theme

Language Bias

Under the *language bias* theme, the participants' advocated an awareness of how language impacted learning. Speaking and understanding a person's language is the key to eliminating some of the *language bias*. The participants defined language broadly to include bilingual and non-native English speakers, but also the cultural, intonation, semantic, or dialectal differences, and other language nuances of native English-speaking people of color (e.g., African-Americans and Native-Americans).

This includes matching the language and metaphors to those of the person, using simple language, communicating at the level of that person, and empowering him or her with words. It also includes a willingness to listen, ask questions, and share in their experiences as culturally different people. In addition, the participants' advocated for an awareness of how language impacted learning.

Some of the participants believed that language was the most pervasive form of bias, specifically as it related to students of color. For example, in one participant's experiences as an elementary language schoolteacher and teacher educator, she stated that, "Since I was a language teacher, it's everywhere." The participants also felt that those language barriers were further compounded by the lack of adequate language deficiency assessment tools. Too often students of color are unfairly assessed and discriminated against based on the way they speak without instruments to tease out culture from deficiency.

Another participant, a former teacher, gives further proof of this. From her experience as a teacher and an attorney for the State Board of Education, she noticed that students of color's learning and their academic performance were negatively impacted by not-so-easily measured factors such as teacher expectations and nonstandard, non-mainstreamed behavior. She also noticed that the big issue in these students' academic success relates to the role that language plays and "then you have the other intelligences, and these are not measurable." Meaning that, oftentimes, the students of color's potential for academic success is evident in other measures of intelligences.

Ultimately, these participants called for more researchers, clinicians, and educators in assessment who are bilingual, multilingual, and culturally diverse and from the same ethnic, racial and cultural backgrounds of the people they are serving. They felt that this was a critical and necessary step needed to help transform the measurement and assessment field to meet the needs of people of color. According to one participant, a former English teacher and teacher educator, in addressing language issues of students of color, "One needs to have a sense of how language bias impacts behavior."

Practice Bias

The area of *practice bias*, although closely related to professional bias, refers more to educational institutions and the inequity in the development and implementation of their policies and programs. In contrast to people bias, practice bias is institutional, similar to institutional racism. Similarly, *practice bias* may be covert and may appear neutral, but it is often detrimental to people of color in its core principles, applications, and practices. One clear example of practice bias is the notion of equality (the same for everyone) verses equity for all; meaning, having a set of assessment standards and tools that are universally applied to all ethnic racial groups regardless of the results.

According to one participant, there is an inherent bias to most assessment instruments being used in education—primarily that assessment is "done by the dominant culture and limited by a western, Eurocentric paradigm." Another example are schools that have curriculum designed for the one-size-fit-all students. In short, schools, colleges, and universities continue to apply policies that tend to benefit mainstream White students.

From their personal and professional experiences, the participants knew that

culture plays a major role in how children learn and that children of color may need diverse curricula. As one participant explains, "Teachers have seen the need to implement alternative assessment for kids that don't fit the typical kid mold." Further, she states, "Everybody benefits from multicultural assessment because we are all different and learn differently." Another participant, an attorney, adds that, "The discipline of law must interpret what a client has written and factor in culture."

The participants, although victims of bias and discrimination in many aspects of their personal lives, felt most affected by *practice bias* within their professions. For example, one participant, an African-American male, recalled a painful childhood memory relating to bias assessment procedures, saying, "Personally, I have been as greatly affected by multicultural assessment issues as well as others, which present themselves through my job."

Similarly, another participant, a Native American school counselor, was witness to the adverse effects of bias practices in her area of mental health. According to her experiences, "Media and journal articles on cultural biases involved in assessment, measurement, and testing have limited opportunities for input from individuals from minority groups."

On the other extreme, another participant, an African American and former school district superintendent, in her current position as a higher education administrator stated that multicultural assessment was not utilized. What she found instead was, "A great deal of stereotyping and assessment of individuals impacted by many external factors."

Professional Bias

Professionals bias, according to the participants, referred to the biases and discrimination perpetuated by the individuals within their professions. This entails the professionals' lack of awareness, knowledge, and skills to work with people from diverse backgrounds. This bias includes failure of professionals to consider or acknowledge (in research, teaching, and practice) the many cultures, races, and ethnicities that exist, and how these cultural variances impact their disciplines.

For example, in her role as lawyer, one participant felt that, "Attorneys often hold culture against their culturally different clients and think they are ignorant." Similarly, another participant, an African American, in his experience as an elementary teacher educator believes that because most test items are written by White authors, "Testing item writers suffer from bias which impacts the validity of the items they create."

History, according to another participant, is another discipline that, "Has historically accepted Caucasian/European race's history as the only true account." Further, he believes that historians have been biased in their assessment, "By giving biased points of views or by dismissing integral parts of viewpoints that relate to other races."

All of the participants agree that their academic areas are greatly impacted by diversity and therefore multicultural assessment needs to be utilized. Often, the issue is not the compilation of data and demographics. One participant points out that in higher education, "This information and subsequent assessment provide us with a perspective of what students we serve." Rather, the issue seems to be related to applying those data in an ethical, fair, unbiased, nonharmful, and multiculturally competent way.

Unfortunately, the difficulty of this is illustrated in one participant's (the Hispanic teacher educator) comment that there are still, "Lots of teachers that have a hard time understanding the need to assess children in an eclectic way." Even more disheartening is that, "There are many teachers who do not acknowledge or value other people's cultures." An authentic example, according to another participant, is that the "A, B, C grading system is still very arbitrary and unfair to students of color."

Instrument Bias

For these participants, *instrument bias* was the most prevalent, widespread, and discussed form of bias. In addition, from their perceptions, *instrument bias* was the most powerful and had the most potential for harm to people of color and was the one that had the most impact on their personal as well as their professional lives. According to one participant, an observant parent/grandparent and public school teacher, "Subjective testing denies children of color enrichment opportunities and the realization of their competence and fosters the belief these children can't learn."

According to the participants, many of the instruments used for assessment in their disciplines are not normed with and were found to be unfair to culturally diverse populations. As a Native American Indian, one participant reflects on her own experiences as a child. She recalls, "I learned very early that standardized tests were biased and therefore, I had to separate my personal cultural views from what I thought the majority answer would be."

Another participant, a former secondary English teacher and doctoral student, stated:

Multicultural issues in general have been very important. In the area of assessment, she felt that students on the secondary level were suffering tremendously because the tests utilized by the district did not make allowances for diversity. In her opinion, these tests were the primary determinants of the child's placement for the next year. If the student faired poorly on the English or reading sections of the test he or she could be denied entrance into gifted and talented classes and college preparatory classes. Teachers were also given the task of determining the placement of a given child. The teacher's referral was important; however, they were to keep in mind the child's test scores. If the scores did not reach a particular cut off score, teachers were instructed not to recommend them for certain services.

Another participant, the African American former public school teacher and administrator, recalls that one of the negative effects of biased instruments was, "Tracking of my son into lower level classes due to his performance on standardized tests." He also witnessed this occurrence "with a percentage of my students over the course of 15 years of teaching."

Overall, the participants believe that by understanding that all assessment instruments are not free from bias, professionals can be empowered to work more effectively with diverse populations. For example, one participant, a mental health professional, believes firmly that the, "Counseling programs are making great strides in training counselors to be more culturally aware." Most urgently, these participants call for more studies to be conducted to determine the accuracy of standardized testing in predicting levels of performance of people of color.

Cultural Congruency and Sensitivity

The participants believed that first, culture is not uniquely measurable, therefore, before assessment and testing can be effectively employed, the person's culture must be identified and then valued. In the emphatic words of one participant, "Culture incongruence" is the one single factor that can adversely affect the success of students of color. Another participant, a retired public school principal, observed, "When cultural needs are identified and met, the student is able to perform more closely to their level of ability than when simply treated as part of a homogeneous group."

As people of color, all of the partici-

pants shared their experiences of White people asking them to represent their race by responding to questions like, "How do you feel as an African American about X and what is the significance of Y to African Americans? It is this assumption that one person as a member of a racial/ethic group represents the culture of his or her entire racial ethnic group that shows an ignorance on the part of White people of the massive heterogeneity that exists within groups of color that pushed one of our participants, the higher education administrator, to advocate multicultural assessment to be "Used to sensitize others to the needs of those in a multicultural environment."

In fact, the participants believed that multicultural assessment should be a prerequisite to activities such as designing curricula, creating new programs, and the like. Multicultural assessment is an important element in determining the program needs of constituents. Without this awareness of cultural differences, and often because of institutional racism and bias, people of color have been victims of massive discrimination practices. One participant, in his roles as teacher and administrator, worked with culturally diverse student populations and observed that because no provisions were made for cultural differences, students of color were rarely rewarded for their work and achievements.

Participants have experienced or witnessed the positive benefit of multicultural assessment and the negative effects of not taking a multicultural approach when assessing people of color. One participant, in her role as a international student teacher, was sensitive to the fact that she was not from the American culture, and felt that she benefited from multicultural assessment mainly because the professors she encountered were open-minded to a diverse student body.

These participants talked about the pain associated with abandoning their cultural values in order to become a part of mainstream academe. One participant, though not visibly a minority, comes from a multicultural background—Jewish and Native American ancestry. He states that he "often finds it necessary to rethink answers and responses to minimize my cultural background influences."

The participants' conversations about multicultural assessment range from the more concrete and specific professional and personal experiences to the more abstract theoretical explanations of the critical issues affecting the assessment. To begin with, multiculturalism has caused a great deal of fear among individuals in the educational system. A lot of the fear,

our participants believe, is because those individuals who are frightened truly do not understand the essence of the mission of multiculturalism. The bottom line is that multicultural education, which as a theoretical construct (nontangible, a thought process), is all about assessing and meeting individual needs. Student success is what we seek in the educational realm.

According to one participant, in order to ensure the success of students of color, "educators must assess and take students where they are, help them to use what they have, to learn and do what they have to do." Another benefit of multicultural assessment and related issues, as expressed by our participants, is helping people become more sensitive to other cultures. For example, becoming more cognizant of how language affects people, in contrast to their prior belief that language was universally valid when working with diverse culture, something that most of them unknowingly had taken for granted.

One participant articulates that multicultural assessment, "Is more like an awareness or an insight." Further, according to another participant, "It has provided me with a better understanding of the needs of a multicultural society and better planning for programming."

The participants also believe that an awareness of differences due to cultural diversity can allow the practitioner to more accurately interpret results generated through multicultural assessment and to more closely meet the needs of the individuals being served. For one participant, another benefit of multicultural assessment has been the creation of programs and projects geared towards cultural diversity.

Further, according to several of the participants, multicultural assessment is an invaluable tool in working in a multicultural environment when measuring knowledge levels as well as determining tolerance levels or programming needs for institutions. According to one participant, "It helps to offset environmental constraints."

According to the participants, some educators feel that the focus on multiculturalism is a fad, here today and gone tomorrow. On the contrary, in the opinion of our participrants, a real focus on diversity training and competencies can assist us in ensuring academic success for all people and not just a limited few.

Lastly, other benefits of cross-cultural assessment procedures and practices, in many cases, are the assurance of equity and fairness in our schools for students from marginalized backgrounds and the recognition of the importance of the multiple talents and gifts that people of color have

given to our nation. In fact, the resource pool in the academy and the workforce is vast when you include people of color, and the sooner assessment practices find a way to make sure all segments of the population are represented, the greater resources this country can expect to have.

Proponents of Diversity

Having educators serve as *proponents of diversity* is critical to institutions of higher education and assists in leveling the playing field for people of color. Several researchers, such as Sue and Sue (2003), in the field of counseling and psychotherapy have for decades advocated increased representation of people of color in academic disciplines. Therefore, there is a need for all of us to actively campaign for the representation of all groups in every aspect of assessment and measurement. To these participants of color, being a strong *proponent of diversity*, in part, means to advocate more representation of people of color.

Most importantly, being a *proponent of diversity* means that one has adapted a willingness to implement a multicultural approach within his or her respective discipline. The participants believed that they, like other professionals in the field, are strong advocates for diversity programs.

As an African-American educational administrator, one participant concluded:

> Research and practical implications of multicultural assessment to the educational discipline are not at the level that it should be. There are many educators, such as myself, who will be advocates of implementing a multicultural approach into the field of education. Similarly, we serve a community of students who are very diverse and this diversity must be addressed.

It means also to know the important contribution of professionals who are culturally, ethnically, and racially diverse.

As proponents of diversity, the participants shared an urgency and necessity to address ethnic and racial diversity within their own professional and academic fields. For example, one participant stated that, "My area of expertise has impacted multiculturalism by creating laws to ensure that measures are taken to create more diversity in our nation's companies, universities, and government-affiliated organizations."

Likewise, another participant acknowledges that, "An enormous amount of research is needed in our area in order to find different types of assessment that will focus on those groups that are underrepresented in gifted and talented education programs." Unfortunately, the participants realize that too often, in their own disciplines, assess-

ment determines who, what, and how resources are allocated. In public schools, one participant discovered that, "If assessment means the varying ways we determine those students who will receive services, multiculturalism has greatly impacted the area of gifted and talented education."

Therefore, the participants believe that those who are proponents of diversity seek ways that the field of assessment and measurement can objectively address diversity issues as well as the needs of underrepresented groups in a less biased way. For example, in one participant's area, new initiatives were developed to address the increase in diversity, "In my workplace, they have had a great impact on the development of new programs designed to increase the graduate minority enrollment."

In some of the participants' fields, in order to determine the applicability to and effectiveness on the targeted groups, certain diversity initiatives were assessed and evaluated. According to one participant, "Yearly state mandates must be evaluated to assess whether or not the criteria used in their assessment met the needs of their population." In another participant's profession, "Data is often compiled to get an accurate view of the campus demographics. This information and subsequent assessment provides us with a perspective of what students we serve."

Unfortunately, not all disciplines are as committed to or as successful at addressing diversity issues. According to one participant, in her field multicultural assessment is being more frequently used, but is not yet widely used. In her opinion, people have attempted to implement other methods in order to avoid assessment, "I've known people to use learning styles in place of the assessment because they are not proponents of diversity."

The bottom line is that participants first believe that is it is imperative to have researchers, clinicians, and educators within all academic areas who are not just visibly persons of color, but who are grounded and acculturated within their own ethnic and racial referent groups, rather than in the dominant culture.

Second, to these participants, as people of color, they have discovered that being proponents of diversity is personal and unavoidable. As an African-American attorney and an administrator in a large research university, one participant captures this sentiment, "In my personal life, it has caused me to be more consciously aware of others and their surroundings." In other words, to these participants, being a proponent of diversity transcends work and professionalism; it is survival.

Discussion

Trustworthiness of the Design

In order to assure trustworthiness of the data, as researchers, we utilized multiple data collection and triangulation methods. Data triangulation included data sources, such as observations, interviews, and available documents. Theory triangulation was also used to gain multiple perspectives to interpret the data. And lastly, we employed methodological triangulation that consisted of combining qualitative inquiry with quantitative instruments and the use of multiple methods, primary and secondary data sources.

The primary sources consisted of SOIG interviews, critiques of transcripts and audiotape interviews, interactive and informal conversational interviews, field notes, and researchers' journals. The secondary sources consisted of personal accounts (i.e., diaries, journals, personal notes) and clinical documents.

In order to further assure trustworthiness, we used prolonged engagement, persistent engagement, member checking, and audit trail. This means that we continued to collect data until all the holes in the data were filled and all questions had been answered (prolonged engagement). We constantly looked for inconsistencies and pressing anomalies (persistent engagement). We reviewed data with participants to ensure accuracy or check inconsistencies (member checks). We discussed the findings with fellow students, dissertation committee members, qualitative research group, and faculty members in order to clarify and validate assumptions (peer debriefing).

We examined all pertinent documents that were available to us. We maintained an account of all the data collected and the inferences made from these data and kept accurate, detailed notes of the data collected from the beginning of my research (audit trail). We tried to determine if the inferences are warranted or justified and flow from the data. We conducted follow-up observations and interviews and interactive interviews.

Ethical Considerations

Patton (2002) states that it is important to deal with ethical issues of qualitative inquiry because: (1) qualitative methods are highly personal and interpersonal; (2) naturalistic inquiry takes the researcher into the real world where people live and work; (3) in-depth interviewing opens up what is inside people; and (4) qualitative inquiry may be more intrusive and involve greater reactivity than surveys, tests, and other quantitative approaches (p. 407). Ac-

cording to Patton, there are several ethical issues that a qualitative researcher should deal with in designing and implementing qualitative studies; mainly, reciprocity, risk assessment (causing no harm to the participants), confidentiality, informed consent, and data access and ownership.

In terms of reciprocity, the benefits for the participants in this study included the personal and professional satisfaction of contributing to research concerning ethnic/racial diverse groups' experiences of assessment, measurement, and testing. We believe that the patterns and commonalties that we uncovered, as well as the research itself, was helpful to the participants personally and in their professional roles as educators, practitioners, researchers, and educators.

In assessing for risk factors and maintaining confidentiality, risks were minimal since we kept the participants and locations confidential and all taped materials as well as interviews, observations, and document data in a personal file. At the end of study, all video and audio taped material was destroyed. All recorded and transcribed interview data were kept confidential. Information was stored on diskettes for necessary updates and only we, the researchers, analyzed all data.

Anonymity was assured by protecting the identity of the participants, their clients, students, instructors, supervisees, and institutions, making certain that the names and other identifying labels were not disclosed. Names and identifying information were coded initially, and only we, the researchers, had access to the data. Participants were asked to choose a pseudonym of personal or ethnic significance to them. That pseudonym was used throughout the study.

Next, informed consent was addressed via a form. Participants were given a copy of the informed consent form to read. It included the description of the study, the nature and purposes of this study, the procedures to be used, and the potential risks, as well as the option to withdraw at any time. Participants were told that participation was voluntary and that they were free to refuse to participate in this study or to withdraw from this study at any time with no negative consequences or penalty.

Closely aligned with informed consent is data access and ownership. The participants were informed (verbally and on the informed consent form) that by signing the informed consent form they were giving us, the researchers, permission to make audio/video tape recordings and proceedings of interview and test interpretation sessions and granting us the right to use

the original audio tapes made of them for research purposes only.

They were also apprised that they had the following rights: to request that video or audio recording be turned off at any time; to request erasure of any part of a recording at the time of its creation or within seven days thereafter; and to review the audio tapes. We told them that all audiotapes would be destroyed and that audiotape materials would be handled in accordance with professional standards and the code of ethics of the American Counseling Association (1995).

Summary of Results

The participants in this study, as multicultural practitioners, witnessed unfairness in assessment in their respective disciplines and they see a need for fair assessment devices. As one administrator stated, "Multicultural assessment is an invaluable tool in working in a multicultural environment." To summarize, the findings in this study uncovered the following themes: (a) bias in terms of instruments, professionals, practice, and language; (b) cultural congruency and sensitivity; and (c) proponents of diversity as key issues in educational assessment and measurement.

These findings substantiate issues associated with the stress process described in Sedlacek's (1994) theoretical framework and similar findings in the literature. That is, these findings are consistent with diversity issues identified in the theoretical framework of Sedlacek's model.

First, consistent with the model, there remains a need for culturally sensitive methods and approaches for including people of color. Second, we continue to employ assessment techniques and instruments that put people of color at a disadvantage, because we have yet to discover a valid single instrument or measure for all.

Next, major efforts need to continue in conducting research, and most importantly, recruiting researchers of color to conduct research that is inclusive of the culturally different. Also, professionals continue to witness bias in situations of incomplete data and lack of data.

Lastly, there remains a shortage of well-trained professionals working in this area, and an even greater shortage of professionals of color, as is evident in the statistics on the small number of teachers, professors, and administrators of color (NCES, 2000; Turner & Myers, 2000).

The findings in this study identified some key factors in educational and psychological assessment and measurement, such as bias, cultural congruency, and the

need for institutions and professionals to be proponents of diversity. These findings also revealed a lack of literature representing the perspectives of people of color within training, teaching, researching, and clinical and practical work, even in cases where such work is being undertaken by people of color.

In addition, the findings reflect the participants' views of the existing commonalities and the discipline-specific competencies in educational and psychological testing and gave suggestions on how we can create meaningful dialogue on diversity issues among and across disciplines and among colleges and universities. This dialogue, according to the participants, must begin at home, with professional people of color, and move out into the arena of educational reform in the areas of assessment and testing for people of color.

The participants' thoughts on needed competencies in assessment of people of color are consistent with the multicultural education and mental health research literature on ethical considerations, training guidelines, and competencies in relating to people from marginalized backgrounds in general (Krishnamurthi, 2003; Toporek & Reza, 2001; Wall & Walz, 2003).

Therefore, if multiculturalism is to fully strengthen the education field, then a challenge to practitioners and researchers from all disciplines impacted by assessment, measurement, and testing must be consistent with Sedlacek's (1994) vision. That is, as professionals, we must look beyond the obvious. This includes strengthening our ethical standards, and scouting for ways to advance diversity through assessment.

As a first step, we must address *professionals bias*, by conducting an internal check, a self-examination of our own values and practices that too often tend to be biased. Next, we must address *instrument bias* to see if we are designing and administering instruments from beginning to end with people of color in mind, instead of just as an afterthought throwing in diverse groups.

We must ask ourselves the ethical questions, such as, are we discarding important data about underrepresented groups? And finally, in terms of *practice bias*, are we committed to developing policies that encourage those in the fields of mental health and education, which have an interest and commitment to the issues impacting diverse groups, to become trained assessors?

However, if any understanding is to develop, we in the field of education must include people of color's voices, perceptions,

and experiences about cultural bias and culturally biased assessment practices and their effects on disciplines and students' cultural awareness and knowledge of cultural issues. It is equally critical that assessors of all races be aware of the importance of the impact of the cultural self on assessment, diagnosis, and intervention, especially as it relates to the bias of White professionals when assessing people of color (Dana, 2000; Suzuki, 2001; Valencia & Suzuki, 2001). These are necessary steps for implementing cultural congruency in the field of educational assessment and measurement, and in becoming proponents of diversity.

It is also our hope that researchers, statisticians, counselor educators, and practitioners be sufficiently challenged by the findings of this study to assess their own cultural awareness, abilities, skills, and knowledge as well as the group under their microscope before attempting to assessment people of color. Researchers and test developers must realize that some biases, whether bias of instruments, people, or practice, are an integral part of most fields, and what is needed is a conscientious effort to avoid inadvertently perpetuating harm.

Concluding Remarks

In this study, the participants have shared their personal and professional experiences of assessment instruments and practices with people of color. The voices of these participants of color speak in unison; that is, assessment has been a harmful tool in its application to people of marginalized backgrounds, specifically people of color. In viewing the plurality of our society, the resource pool is vast and so are the invaluable gifts and talents brought to the table by people of color.

Because of this vast resource pool of people of color, assessment procedures and practices should be used in a manner that demonstrates and maximize the success of these diverse populations in our schools, workplaces, and thus in society. Therefore, the sooner, assessment practices that are fair and unbiased are used with all segments of the population, the greater resources we can expect to have as a nation.

However, until norms are set with the various cultural groups, cross-cultural assessment should be approached from a well-rounded perspective, one that includes input via interviews, observations, and self-reports with multiple assessment instruments and multiple sources, including the voices of people of color.

References

American Counseling Association. (1995). *ACA code of ethics and standards of practice*. Alexandria, VA: Author.

Brown, S. P., Lipford-Sanders, J., & Shaw, M. (1995). Kujichagulia—Uncovering the secrets of the heart: Group work with African American women on predominantly White campuses. *Journal for Specialists in Group Work, 20,* 151-158.

Burley, H., Butner, B., Marbley, A., & Bush, L. (2001). Standardized testing: For richer or poorer, for democracy or meritocracy? *Connections: Journal of Principal Development and Preparation, 3*(5), 5-19.

Dana, R. H. (2000). The cultural self as locus for assessment and intervention with America Indians/Alaska natives. *Journal of Multicultural Counseling and Development, 28,* 66-82.

D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural counseling training. *Journal of Counseling and Development, 70,* 143-150.

Drew. C. J. (1973). Criterion-referenced and norm-referenced assessment of minority group children. *Journal of School Psychology, 11,* 323-329.

Eeels, K., Davis, A., & Havighurst, R. J., Herrick, V.E., & Tyler, R.W. (Eds). (1951). *Intelligence and cultural differences.* Chicago: University of Chicago Press.

Feagin, J. R., Vera, H., & Batur, P. (2001). *White racism: The basics* (2nd ed.). New York: Routledge

Gainer, K. A., (1992). Internalized oppression as a barrier to effective group work with Black women. *Journal for Specialists in Group Work, 17,* 235-242.

Gay, G., & Abrahams, R. D. (1973). Does the pot melt, boil, or brew? Black children and White assessment procedures. *Journal of School Psychology, 11,* 330-340.

Greeley, A. T., Garcia, V. L., Kessler, B. L., & Gilchrest, G. (1992). Training effective multicultural group counselors: Issues for a group training course. *Journal for Specialists in Group Work, 17,* 196-209.

Harris, J. L. (2003). Toward an understanding of literacy issues in multicultural school-age populations. *Language, Speech, and Hearing Services in Schools, 34* (1), 17-19.

Hopkins, K. D., & Stanley, J. C. (1990). *Educational and psychological measurement and evaluation,* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ivey, A. E., Ivey, M. B., & Simek-Morgan, L. (1993). *Counseling and psychotherapy: A multicultural perspective,* (3rd ed.). Boston: Allyn & Bacon.

Johnson, I. H., Torres, J. S., Coleman, V. D. & Smith, M. C. (1995). Issues and strategies in leading culturally diverse counseling groups. *Journal for Specialists in Group Work, 20,* 143-150.

Kaplan, R. M., & Saccuzzo, D. P. (1993). *Psychological testing: Principles, applications, and issues.* (3rd ed). Pacific Grove, CA: Brooks/Cole.

Krishnamurthi, M. (2003). Assessing multicultural initiative in higher education institutions. *Assessment and Evaluation in Higher Education, 28*, 263-77.

Lindsey, M. L. (2000). Mental health assessment and treatment of African Americans: A multicultural perspective (pp. 159-208). In Cuellar, I., & Paniagua F. A. (Eds.). *Handbook of multicultural mental health.* San Diego, CA: Academic Press.

Marbley, A. F. (2004). His eye is on the sparrow: A counselor of color's perception of facilitating predominantly White groups. *Journal for Specialists in Group Work, 29,* 247-258.

Miller-Jones D. (1989). Culture and testing. *American Psychologist, 44,* 360-366.

National Center for Educational Statistics (NCES). (2000). *Digest of educational statistics.* Washington, DC: U. S. Government Printing Office.

Padilla, A. M. (2001). Issues in culturally appropriate assessment. In L. A. Suzuki, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical psychological, and educational applications* (2nd ed.) San Francisco: Jossey-Bass/Pfeiffer.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education. Higher and Adult Education Series.* San Francisco: Jossey-Bass.

Patton, M. Q. (2002). *Qualitative research and evaluation methods.* (3rd ed.). Thousand Oaks, CA: Sage.

Pennington, H. J. (2000). Issues in mathematics education with African-American students. *Multicultural Education, 7(3),* 36-41.

Ponterotto, J. G., Rieger, B. P., Barrett, A., & Sparks, R. (1994). Assessing multicultural counseling competence: A review of instrumentation. *Journal of Counseling and Development, 72,* 323-328.

Pope-Davis, D. B., & Dings, J. G. (1994). An empirical comparison of two self-report multicultural counseling competency inventories. *Measurement and Evaluation in Counseling and Development, 27,* 93-102.

Pope-Davis, D. B. & Ottavi, T. M. (1994). The relationship between racism and racial identity among White Americans: A replication and extension. *Journal of Counseling and Development, 72,* 293-297.

Ridley, C. R., Hill, C. L., & Wiese, D. L. (2001). Ethics in multicultural assessment: A model of seasoned application. In L. A. Suzuki, J. G. Ponterotto, et al. (2001). *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed.). San Francisco: Jossey-Bass/Pfeiffer.

Robinson, T. L. (2005). *The convergence of race, ethnicity, and gender: Multiple identities in counseling.* Upper Saddle River, NJ: Prentice Hall.

Rollock, D. A., Westman, J. S., & Johnson, C. (1992). A Black student support group on a predominantly white university campus: Issues for counselors and therapists. *Journal for Specialists in Group Work, 17,* 243-252.

Roysircar, G., Sandhu, D. S., & Bibbins, V. E. Sr. (Eds.), (2003). *Multicultural competencies: A guidebook of practices.* Alexandria, VA: Association for Multicultural Counseling and Development.

Sabatier, C. (2003). Assessment and culture: Psychology tests with minority Populations. *Journal of Cross-Cultural Psychology. 34,* 484-495.

Sedlacek, W. E. (1994). Issues in advancing diversity through assessment. *Journal of Counseling and Development, 72,* 549-553.

Scheurich, J. J., & Young, M. D. (1997). Coloring epistemologies: Are our research epistemologies racially biased? *Educational Researcher, 26,* 4-16.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed) Thousands Oaks, CA: Sage.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling andDevelopment, 70,* 477-486.

Sue, D. W., & Sue, D. (2003). *Counseling the culturally different: Theory and practice* (4th ed.). New York: Wiley.

Suzuki, L. A. (2001). Multicultural assessment: Trends and directions revisited. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical psychological, and educational applications* (2nd ed.) San Francisco: Jossey-Bass/Pfeiffer.

Suzuki, L. A., Ponterotto, J. G., & Meller, P. J. (2001). *Handbook of multicultural assessment: Clinical psychological and educational applications* (2nd ed.) San Francisco: Jossey-Bass/Pfeiffer.

Toporek, R. L., & Reza, J. V. (2001). Context as a critical dimension of multicultural counseling: Articulating personal, professional, and institutional competence. *Journal of Counseling and Development, 29,* 13-30.

Turner, C. V. S., & Myers, S. L. (2000). *Faculty of color in academe: Bittersweet success.* Boston: Allyn & Bacon.

Valencia, R. R., & Suzuki, L. A. (2001). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues. Racial and ethnic minority psychology series.* Thousand Oaks, CA: Sage.

Wall, J. E., & Walz, G. R. (Eds). (2003). *Measuring up: Assessment issues for teachers, counselors, and administrators. Ann Arbor, MI:* ERIC Clearinghouse on Counseling and Student Services.

Wyche, L. G., & Novick, M. R. (1986). Standards for educational and psychological testing: The issue of test bias from the perspective of school psychology and psychometrics. *Journal of Black Studies, 11,* 43-48.