

# Behind closed doors: Assessing individuals from diverse backgrounds

Elizabeth V. Horin<sup>a,\*</sup>, Brigida Hernandez<sup>b</sup> and Oscar A. Donoso<sup>c</sup>

<sup>a</sup>VA Long Beach Healthcare System, Long Beach, CA, USA

<sup>b</sup>YAI Network, New York, NY, USA

<sup>c</sup>Children's Hospital Los Angeles, Los Angeles, CA, USA

**Abstract.** This study surveyed 150 professionals on testing considerations made for clients' race/ethnicity and disability when (a) selecting tests, (b) administering tests, and (c) interpreting and writing results. Participants, who tested for the vocational rehabilitation system, completed a semi-structured survey online. Results indicated that they were more likely to consider clients' disability when selecting tests than race/ethnicity. During test administration, the majority reported not making adaptations or modifications (thus adhering to standardized instructions). However, participants were likely to factor clients' race/ethnicity and disability when interpreting results and writing reports. Content coding of open-ended responses revealed concerns when testing individuals who did not speak English fluently. An important area of research considering demographic shifts occurring in the United States, findings suggest the need to develop evidence-based practices when assessing culturally diverse populations.

**Keywords:** Psychological assessment, racial and ethnic minorities, disabilities

## 1. Introduction

Although much valuable information can be garnered from psychological evaluations, the use of standardized tests with individuals from diverse backgrounds (including racial and ethnic minorities and people with disabilities) has been a source of debate. Some argue that bias can occur during test development, administration, and interpretation [1–5]. To address potential bias, the American Psychological Association (APA) has called for assessors to recognize the limitations of tests and to strengthen cultural knowledge, awareness, and skills. Considering the changing demographics of the U.S. [6–11], research related to testing individuals from diverse backgrounds is critical and scant. Thus, the overarching purpose of this study was to examine

the testing practices of professionals when evaluating racial and ethnic minorities and individuals with disabilities.

### 1.1. Assessing individuals from diverse backgrounds

Historically, the constructs of intelligence and psychological disorders have been conceptualized using a Western perspective and such conceptualizations may not transfer easily to other cultures. For instance, in China, the Western conceptualization of depression is uncommon and therefore presents a challenge for valid diagnosing [12, 13]. In a similar vein, psychological tests have been for the most part developed using a Western framework and normed with members of the mainstream population. Consequently, particular groups may have an unintended advantage over others [14]. A case in point, the standardization sample of the Minnesota Multiphasic Personality Inventory – Second Edition does not adequately represent Latinos,

---

\*Address for correspondence: Elizabeth V. Horin, VA Long Beach Healthcare System, Long Beach, CA, USA. E-mail: elizabeth.horin@va.gov.

Asian Americans, and Native Americans when current population demographics are considered [2].

For people with disabilities, bias may also be evident when test items reflect the sequelae of disability more than the symptoms of psychological disorders [15]. Some have questioned whether the somatic items of depression screening tools (e.g., Beck Depression Inventory) retain their diagnostic utility when administered to individuals with physical disabilities [16, 17]. Compounding the problem of test utility, the majority of standardized tests have not been normed with individuals with disabilities, and research addressing the impact of disability on the testing process is limited [3]. As a result, psychological evaluations may inaccurately characterize the strengths and weaknesses of this group [18]. Olkin [3] noted that three types of generalizations can be harmful for individuals with disabilities: (1) generalizing from a small group of individuals with a specific disorder to the general population of individuals with the same disorder, (2) generalizing from a specific type of disability to all individuals with disabilities, and (3) generalizing either of these across races and ethnicities.

In addition, there may be subsidiary factors that contribute to test bias including the mode of test administration or examiners' personality style [4]. For example, examiners are trained to administer tests in a standardized manner, which may result in instructions being ambiguous and not easily understood by individuals from diverse backgrounds. In other words, technical verbiage may lead diverse people to respond to standardized measures in ways that differ from convention [1, 19].

### 1.2. Multicultural assessments

Multicultural assessment refers to the use of standardized tests, in conjunction with a clinical interview and other sources, to obtain information relevant for multicultural competence [20]. Competence in multicultural assessment is the ability and committed intention to account for cultural factors in order to develop accurate, comprehensive, and impartial conceptualizations [21]. Moreover, understanding of cultural factors (i.e., race, ethnicity, gender, socioeconomic status, disability) should occur throughout all assessment phases including test selection, test administration, scoring of protocols, interpretation of results, and written reports.

Over the years, numerous guidelines have been established to address the psychological assessment of

individuals from diverse groups [22]. The latest version of the APA's Ethical Principles of Psychologists and Code of Conduct [23] includes standards that address testing bias. Specifically, Standard 9.02a states: *Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques*. The inclusion of the word "adapt" in this standard suggests that departures from standard administration procedures may be considered if adaptations and modifications do not impact the test's construct. For instance, a person with Cerebral Palsy may receive assistance with circling items on a written test or a person with Chronic Fatigue Syndrome may receive breaks as long as these adaptations and modifications do not affect the construct validity of the test.

In selecting tests, Standard 9.02b stipulates that it is incumbent on psychologists to determine if a particular test can be used reliably and validly given clients' population characteristics such as race, ethnicity, culture, language, gender, age, or disability. If reliability or validity data do not exist (or if psychologists use tests without established norms for the individual being assessed), psychologists are encouraged to describe the strengths and limitations of the results and interpretations. Lastly, Standard 9.02c states that language preference and proficiency be taken into account when selecting an assessment method, if and when the alternative language is not the variable being tested. Professionals are encouraged to discuss any personal, situational, linguistic, and cultural differences that may impact test scores [23].

Furthermore, the *Standards for Educational and Psychological Testing (Standards)*, established by the American Educational Research Association (AERA), APA, and National Council on Measurement in Education (NCME) [24] detail a number of items to minimize test bias. The *Standards* recognize that the psychological testing process is not infallible and involves the participation of multiple stakeholders including (a) test developers, (b) tests publishers and marketers, (c) test administrators and interpreters, (d) decision-makers using test results, (e) test sponsors (whether institutional or governmental agencies), (f) test reviewers and evaluators, and (g) test-takers. All of these stakeholders have responsibility in promoting the sound and ethical use of tests to ensure the fair treatment of test-takers.

The 1985 version of the *Standards* [25] included a chapter on issues concerning the testing of individuals

with disabilities [26]. This early version encouraged assessors to note non-standard administrations in order to avoid misleading test users or “even harm handicapped test takers whose scores do not accurately reflect their abilities”. The 1999 *Standards* [24] recommend that assessors raise their awareness of testing bias and base testing decisions and practices on empirical research. It also encourages professionals to pay particular attention to the validity of inferences gleaned from results when adaptations and modifications are made.

### 1.3. Strategies for multicultural assessments

Numerous strategies have been proposed to evaluate individuals whose cultural background may impact test performance. One particular approach is *testing of limits*, which provides opportunities to explore an individual's performance after completion of standardized testing. Examples include reviewing items again, removing time limits, presenting correct responses and determining if examinees understand them, inquiring why specific items or subtests were difficult, and presenting items in examinees' native language. Although testing of limits is used to determine whether performance may be enhanced with adaptations and modifications, it is important to consider the effects of learning from the original administration [27].

The AERA, APA, and NCME [24] have also suggested strategies when assessing individuals with physical, mental, and developmental disabilities. Specifically, Standard 10 identifies six types of adaptations and modifications that may be appropriate: (1) alterations to the test format (e.g., reading sections of a test aloud for individuals with visual impairments), (2) alterations to the response format (e.g., short answers in lieu of multiple choice), (3) additional time or breaks to complete testing, (4) omission of subtests, (5) use of alternate test forms (e.g., Braille or large print versions), and (6) change to the testing setting (e.g., making test sites accessible to individuals who use wheelchairs). Standard 10 also lists instances when the use of modifications would not be appropriate (i.e., presence and/or degree of a disability is being assessed). Other adaptations and modifications that have been suggested include population-specific norms, test translations, and culture-specific interpretations [1, 3, 26].

To date, there has been discussion about the need to account for cultural factors when testing individuals of diverse backgrounds. Whether such steps are actually taking place is unclear. To the authors' knowledge, no study has examined the testing practices of

professionals when assessing individuals from diverse backgrounds.

### 1.4. The vocational rehabilitation system and need for multicultural assessments

Obtaining employment is one of the biggest challenges experienced by the disability community. When compared to the national unemployment rate (8.3%), the rate for those with disabilities (13.0%) is notably larger. Furthermore, data from the overall labor market indicate that only 34.9% of men and 30.6% of women with disabilities aged 16 to 64 are working [28]. For racial and ethnic minorities with disabilities, employment struggles are even more pronounced [29, 30]. In 2009, Blacks (28.7%), Black Hispanics (29.6%), and Native Americans/Alaskan Natives (32.1%) aged 21 to 64 with disabilities had lower employment rates than non-Hispanic Whites with disabilities (37.4%). Interestingly, White Hispanics (37.9%) and Asians (39.3%) had the highest employment rates [31].

The vocational rehabilitation (VR) system was established to improve employment outcomes of individuals with disabilities deemed eligible for services. Funded on both the federal and state level, VR offers a variety of services including psychological and vocational assessments; academic, business, or vocational training; personal or vocational adjustment training; employment counseling; and job placement [32, 33]. During the 2005 fiscal year, over 1.4 million adults were served by VR programs [34, 35].

Psychological assessments play a critical role in the rehabilitation process with regards to eligibility determination and services [36]. In a longitudinal study of 8,500 VR consumers, approximately 35% received psychological evaluations for eligibility determination while 39% had prior psychological assessments retrieved for this purpose [32]. Although the goals of the VR system are significant, racial and ethnic minorities tend to experience disparities related to VR access, acceptance, and closure [37–45]. Of note, Hayward and Schmidt-Davis [30] found that employment outcomes were poorer for VR consumers of color when compared to White consumers.

### 1.5. Purposes of this study

The current study examined the testing practices of professionals conducting psychological evaluations for the VR system, with particular attention paid to the assessment of clients who are racial and ethnic

minorities and have disabilities. First, it examined the extent that professionals considered examinees' race/ethnicity and disability during test selection. Second, it explored the use of adaptations and modifications during test administration. Third, it examined the consideration of race/ethnicity and disability when interpreting test data and writing reports. Fourth, this study examined the impact of participants' demographic and professional variables on overall findings. It is important to note that the terms adaptation, modification, accommodation, and translation are often used interchangeably in the literature [46]. For this particular study, all these terms were used to capture as much information as possible.

## 2. Method

### 2.1. Participants

Participants were professionals who conducted psychological assessments for the VR system during a 12-month period and practiced in the states of California, Florida, Illinois, New York, or Texas. These states were selected because of their size, regional representation, and cities with great cultural diversity [47, 11]. Of 198 individuals who completed the survey, 150 were retained (147 online and 3 by phone). The remaining 48 were discarded because of numerous missing items, haphazard responding, or failure to meet eligibility criteria (e.g., participant did not endorse conducting psychological assessments for people with disabilities seeking VR services within the past year). To minimize random responding, participants were required to answer key items; however, because of skip patterns related to item applicability, not every participant was required to complete all survey items. Although survey completion was confidential, participants entered their names and contact information into a separate database in order to receive their incentive. An overall response rate could not be calculated as it was unclear how many individuals received information about the study through the various recruitment efforts.

### 2.2. Instrument

Consisting of 78 items, a semi-structured survey (Psychological Assessment Survey, PAS) was developed by the authors to assess the testing practices of professionals conducting psychological evaluations for the VR system. The study reported herein was part of

a larger study that examined participants' test usage patterns across 8 testing domains: (1) achievement, (2) adaptive behavior, (3) cognitive/intelligence, (4) neuropsychological, (5) objective personality, (6) projective personality, (7) symptom checklists, and (8) vocational aptitude and interests [48]. The development of the PAS was based on a review of the literature related to psychological testing, racial and ethnic minorities, and individuals with disabilities [2, 32, 49]. The PAS was pilot-tested with three psychologists from Illinois who met criteria to participate in this study. They completed the online survey and were asked to comment on the clarity of items, thoroughness of the survey, and completion time. The psychologists provided minimal feedback and no changes were made to the survey.

For the current study, there was a focus on PAS items related to considerations of race/ethnicity and disability during the test selection, test administration, and test interpretation and report writing phases of psychological evaluations. Specifically, test selection items were asked after each of the eight testing domains and included: *During the past 12 months, did you select a particular test because of a client's racial/ethnic background? During the past 12 months, did you select a particular test because of a client's type of disability?* If either were endorsed, participants were asked to specify the test and reason(s) for selecting the test.

For test administration, participants were asked the following questions after each testing domain: *During the past 12 months, did you make any adaptations or modifications (e.g., paraphrasing standardized instructions, using norms specific to a racial/ethnic group) to any test for members of a particular racial/ethnic group? During the past 12 months, did you make any adaptations or modifications (e.g., providing materials in Braille, using norms specific to a disability group) for members of a particular disability group?* If either item was endorsed, participants were asked to specify the test and the type of adaptation or modification made.

The interpretation/report writing items were presented only once and included: *During the past 12 months, did you make considerations of clients' racial/ethnic background when interpreting test data and writing reports. During the past 12 months, did you make considerations of clients' type of disability when interpreting test data and writing reports?* If either were endorsed, participants were asked to elaborate upon their endorsement.

In addition, the PAS gathered information pertaining to participants' demographic and professional characteristics. Demographic information included age, sex,

and race/ethnicity. Professional information included academic attainment, theoretical orientation, number of years conducting VR assessments, number of VR assessments conducted in the past 12 months, clients' race/ethnicity, clients' disability, and state of practice. In addition, the PAS inquired about common reasons for referral as well as degree of freedom participants had in test selection.

### 2.3. Procedure

During the course of this study, a number of recruitment strategies were employed including posting information on the APA's listserv for the Division of Rehabilitation Psychology; e-mailing and mailing information to vendors contracted to conduct psychological assessments for each target state; posting information on each state's Psychological Association's website; and utilizing a snowball sampling approach. Participants had the option of completing the survey online or over the phone with a research assistant. The estimated time to complete the survey averaged 45 minutes. Initially, participants were compensated for their time and effort with a \$25 gift card, which later was increased to \$50 in order to improve the participation rate.

## 3. Results

Results of this study are presented under five sub-headings: (1) description of participants' demographic and professional characteristics, (2) racial/ethnic and disability considerations during test selection, (3) racial/ethnic and disability considerations during test administration, (4) racial/ethnic and disability considerations during test interpretation and report writing, and (5) impact of participants' demographic and professional characteristics on test selection, test administration, and test interpretation/report writing. For open-ended responses (where participants elaborated upon items pertaining to test selection, test administration, and test interpretation/report writing), content analysis was used to code main categories [50]. Specifically, written responses were read and categorized independently by two of the authors. As an example, administering a test that was available in Braille was categorized as an alternate format of a test. After initial coding was completed, the two authors met to discuss and agree upon a list of main categories for test selection, test administration, and test

interpretation/report writing. The third author was used when consensus was not achieved by the two researchers.

### 3.1. Participants' demographic and professional characteristics

Table 1 displays the demographic and professional characteristics of the sample. In brief, participants tended to be male (64.7%), White (80.7%), and over the age of 49 (52.0%). The majority held a Ph.D. (60.7%) or Psy.D. (22.7%) in clinical psychology and endorsed a cognitive-behavioral (54.0%) theoretical orientation. Of the sample, 55.4% had been testing for over nine years, 51.4% had completed over 25 test batteries during a 12-month period, and 67.4% reported "a lot" or "some" freedom in selecting tests. Participants reported having experience with testing African Americans (82.7%), Latinos (79.3%), and Asian Americans (45.3%) as well as individuals with a wide range of disabilities. Reasons for testing referrals included assessing cognitive difficulties (92.7%), depressive disorders (76.7%), anxiety (68.0%), personality disorders (59.3%), and psychotic symptoms (54.7%). Geographically, participants represented Texas (38.7%), Florida (24.7%), California (12.7%), New York (11.3%), and Illinois (10.7%).

### 3.2. Test selection: Racial/ethnic and disability considerations

Table 2 presents the frequency and percentage of participants who reported selecting tests based on clients' race/ethnicity and disability: *During the past 12 months, did you select a particular test because of a client's racial/ethnic background...disability?* When the eight testing domains were collapsed, participants reported being more inclined to select tests because of disability (62.7%) rather than race/ethnicity (28.7%). When data for each of the eight testing domains were examined, test selection considerations were most evident with cognitive, intelligence, and neuropsychological tests. Participants' open-ended responses revealed that nonverbal and Spanish versions of tests (such as the Wechsler Adult Intelligence Scale (WAIS) and Woodcock Johnson) were used to address language barriers particularly with Latino clients. For clients with disabilities, participants reported selecting tests that addressed specific referral concerns (i.e., learning disabilities, traumatic brain injury, intellectual disabilities).

Table 1  
Participants' demographic and professional characteristics

Variable	<i>n</i>	%	Variable	<i>n</i>	%
Sex			Freedom in test selection		
Male	97	64.7	A lot of freedom	73	48.7
Female	53	35.3	Some freedom	28	18.7
			A little freedom	27	18.0
			No freedom	17	11.3
			Missing	5	3.3
Race/ethnicity			Clients' race/ethnicity		
White/Caucasian	121	80.7	White/Caucasian	138	92.0
Other	11	7.3	African American/Black	124	82.7
Latino/Hispanic	10	6.7	Latino/Hispanic	119	79.3
African American/Black	4	2.7	Asian/Pacific Islander	68	45.3
Asian/Pacific Islander	1	0.7	Native American	37	24.7
Missing	3	2.0			
Age			Clients' disability		
<40 years	38	25.3	Cognitive/developmental	139	92.7
40–49 years	34	22.7	Psychiatric/emotional	131	87.3
50–59 years	40	26.7	Medical/chronic illness	116	77.3
>60 years	38	25.3	Physical	102	68.0
Academic degree			Hearing	66	44.0
PhD clinical psychology	91	60.7	Visual	64	42.7
PsyD clinical psychology	34	22.7			
Other degree	25	16.7	Reason for referral		
Theoretical orientation			Cognitive difficulties	139	92.7
Cognitive behavioral	81	54.0	Depressive disorders	115	76.7
Eclectic	30	20.0	Anxiety disorders	102	68.0
Psychodynamic	27	18.0	Personality disorders	89	59.3
Other	12	8.0	Psychotic symptoms	82	54.7
			Other	36	24.0
Years testing for the VR system			State testing for VR system		
<4 years	37	24.7	Texas	58	38.7
5–9 years	30	20.0	Florida	37	24.7
10–19 years	37	24.7	California	19	12.7
20–29 years	24	16.0	New York	17	11.3
30 years and >	22	14.7	Illinois	16	10.7
Test batteries in last 12 months			Multiple states	3	2.0
<10	34	22.7			
10–25	39	26.0			
26–50	28	18.7			
51–100	24	16.0			
101 and >	25	16.7			

### 3.3. Test administration: Racial/ethnic and disability considerations

Table 3 displays the frequency and percentage of participants who endorsed making an adaptation or modification during test administration given clients' race/ethnicity and disability: *During the past 12 months, did you make any adaptations or modifications to any test for members of a particular racial/ethnic group... disability group?* When the eight testing domains were combined, participants endorsed making more adaptations and modifications for disability

Table 2  
Test selection based on clients' race/ethnicity and disability

Testing domain	Race/ethnicity		Disability	
	<i>n</i>	%	<i>n</i>	%
Achievement	8	5.3	30	20.0
Adaptive behavior	–	–	28	18.7
Cognitive/intelligence	29	19.3	44	29.3
Neuropsychological	9	6.0	44	29.3
Objective personality	7	4.7	42	28.0
Projective personality	3	2.0	25	16.7
Symptom checklist	8	5.3	30	20.0
Vocational interest/aptitude	6	4.0	17	11.3

Note: Participants may have endorsed selecting tests from more than one testing domain.

Table 3  
Adaptations or modifications during test administration based on clients' race/ethnicity and disability

Testing domain	Race/ethnicity		Disability	
	<i>n</i>	%	<i>n</i>	%
Achievement	7	4.7	16	10.7
Adaptive behavior	10	6.7	5	3.3
Cognitive/intelligence	13	8.7	13	8.7
Neuropsychological	13	8.7	13	8.7
Objective personality	7	4.7	19	12.7
Projective personality	7	4.7	10	6.7
Symptom checklist	14	9.3	18	12.0
Vocational interest/aptitude	4	2.7	8	5.3

Note: Participants may have endorsed making adaptations/modifications from more than one testing domain.

(33.3%) than race/ethnicity (22.7%), although it was clear that the majority of participants reported that they were not adapting or modifying.

When open-ended responses were coded [50], eight types of adaptations and modifications were identified: (1) paraphrased or clarified instructions or items; (2) used group-specific norms; (3) translated or interpreted items (including American Sign Language); (4) administered alternate formats of tests (e.g., large print, Braille, verbal administration); (5) provided physical assistance (e.g., completing pencil-and-paper measures); (6) modified or eliminated subtests or items; (7) provided other type of adaptation or modification (e.g., testing of limits); and (8) endorsed making an adaptation or modification but did not provide clear information about it. Table 4 displays the number of

Table 4  
Types of adaptations or modifications endorsed based on clients' race/ethnicity and disability

Adaptation/modification	Race/ethnicity		Disability	
	<i>n</i>	%	<i>n</i>	%
Paraphrased or clarified instructions/items	28	18.7	14	9.3
Used group specific norms	4	2.7	3	2.0
Translated or interpreted tests	14	9.3	22	14.7
Administered alternate formats of test	5	3.3	35	23.3
Provided physical assistance	—	—	12	8.0
Modified or eliminated subtests/items	6	4.0	16	10.7
Provided other adaptation/modification	3	2.0	8	5.3

Note: Participants may have endorsed more than one adaptation/modification.

participants who reported making each aforementioned adaptation and modification. When testing racial and ethnic minorities, "paraphrasing or clarifying instructions or items" was most frequently reported (18.7%). For clients with disabilities, "administering alternate formats of tests" was the most common adaptation or modification (23.3%) endorsed.

### 3.4. Test interpretation and report writing: Racial/ethnic and disability considerations

To examine considerations made during test interpretation and report writing, participants were asked, *During the past 12 months, did you make considerations of clients' racial/ethnic background (... type of disability) when interpreting test data and writing reports?* A high number of participants endorsed making considerations during these phases for both groups (70.0% for racial/ethnic minorities and 73.3% for individuals with disabilities). Content coding of their responses indicated an appreciation that cultural background (whether racial/ethnic background or disability) may impact test results and need to be addressed when interpreting test results and writing reports.

### 3.5. Impact of participants' demographic/ professional characteristics on test selection, test administration, and test interpretation/ report writing

Chi square analyses were conducted to assess group differences in (a) test selection, (b) test administration, and (c) test interpretation/report writing given participants' demographic and professional characteristics (specifically, sex, age, theoretical orientation, years testing for the VR system, number of VR assessments conducted in past 12 months, freedom in test selection, and state of practice). Other variables (i.e., race/ethnicity of participants and clients, academic degree, clients' type of disability, reason for referral) were not included in these analyses because cell sizes were either small or not meaningfully different. In addition, to reduce the likelihood of Type 1 error, *p* values were set at 0.01.

Significant group differences were found with only two variables: (1) Number of VR assessments conducted in the past 12 months - Participants who reported conducting over 100 batteries were more likely to select tests based on clients' race/ethnicity when compared to participants who reported conducting fewer batteries,  $\chi^2(4, N = 150) = 23.80, p < 0.001$ . (2) Freedom in test

selection – Participants endorsing “some” and “a lot” of freedom in test selection reported more adaptations and modifications based on disability when compared to participants endorsing “a little” and “no” freedom,  $\chi^2(3, N = 145) = 11.88, p < 0.01$ .

#### 4. Discussion

The main purpose of this study was to examine the testing practices of professionals who conduct psychological evaluations for the vocational rehabilitation (VR) system. In particular, it focused on the extent that race/ethnicity and disability were considered during the phases of test selection, test administration, and test interpretation/report writing. This under-investigated area of research is of great importance given the demographic shifts this country is experiencing [6–11] and recognition that standardized tests may not accurately reflect the abilities of diverse populations [2, 3].

With regards to test selection, the majority of participants (62.7%) reported that they considered a client's disability when selecting assessment tools; much fewer (28.7%) considered their race/ethnicity. Of note, test selection decisions were often seen in the domains of cognitive, intelligence, and neuropsychological functioning. Participants' open-ended responses suggested a need to select tests that were appropriate to address highly specific referral concerns (i.e., learning disabilities, traumatic brain injury, intellectual disabilities) when testing clients with a wide range of disabilities. When race and ethnicity was considered during test selection, these decisions seemed driven by concerns with English language fluency among Latino clients. In such cases, participants reported that they relied on standardized tests that were either non-verbal or available in Spanish. With the Latino and Asian populations projected to grow in the U.S. [9, 10], it is imperative that standardized tests continue to be both culturally and linguistically appropriate [2].

During test administration, the majority of participants reported not making adaptations or modifications during the assessment process. This particular finding suggested that participants adhered to standardization instructions in order to obtain reliable and valid results. When adaptations and modifications did occur, they were more likely considered for clients' disability (33.3%) rather than race/ethnicity (22.7%). Administering alternate formats of a test was the most common adaptation for clients' with disabilities (including large print, Braille, verbal administrations), whereas

paraphrasing and clarifying instructions and items was most frequently reported for clients who were racial/ethnic minorities.

When compared to race/ethnicity, it is not surprising that more participants reported making adaptations and modifications based on clients' disabilities. Most test manuals include instructions for assessing individuals with a variety of disabilities [51]. As an example, the WAIS-IV manual discusses the omission of certain subtests that require motor skills when testing individuals with physical impairments. It also suggests placing greater emphasis on certain subtests when estimating abilities of individuals with severe language difficulties [52]. To date, researchers and practitioners have paid attention to providing accommodations to people with disabilities as there are clear concerns with the reliability and validity of certain standardized tests [51, 53, 54].

There are a number of factors that may account for the limited use of adaptations and modifications with racial and ethnic minorities. Issues of cultural diversity have only recently begun to receive attention and be integrated into psychological research, practice, and training [55, 56]. As a result, our knowledge base about appropriate practices for diverse populations is still evolving. As such, the impact of adaptations and modifications on the validity of scores is unclear [52, 57, 58]. Moreover, there is tremendous heterogeneity among racial and ethnic minorities, making it difficult to establish appropriate adaptations and modifications that address all potential subgroups.

Participants reported greater willingness to factor in both race/ethnicity (70.0%) and disability (73.3%) during the interpretation and report writing phases of an assessment. These high percentages suggest awareness that cultural background may impact standardized test scores and need to be considered when interpreting results and writing reports. This practice is in line with the multicultural assessment guidelines proposed by APA [2, 24] and Allen [59], which encourage professionals to interpret results in a contextual manner rather than stray away from standardization procedures.

Finally, this study examined the impact of participants' demographic and professional characteristics on (a) test selection, (b) test administration, and (c) test interpretation/report writing. Significant group differences were found in only two instances: number of VR assessments conducted in the past 12 months and freedom in test selection. When examining these group differences more closely, they are not particularly surprising. For example, participants who conducted the largest number of VR assessments (over 100) were



more likely to endorse considering race/ethnicity during test selection when compared to those conducting fewer assessments. One would anticipate this finding given that the high-assessment group probably had more opportunities to test individuals of culturally diverse backgrounds. Similarly, it would have been expected that participants who reported “some” and “a lot” of freedom in test selection would report more adaptations and modifications based on clients’ disability than those endorsing less freedom.

#### 4.1. Implications

To our knowledge, this study breaks ground in its examination of professionals’ testing practices with clients from diverse backgrounds. Although the APA emphasizes the need to account for cultural factors during the psychological assessment process, research related to this domain is limited. In other words, are professionals considering race, ethnicity, and disability when selecting tests, administering tests, interpreting tests, and writing reports? Although this study provides preliminary answers, findings are based on self-report and not actual behaviors. It would behoove future researchers to investigate testing practices using methods that are more behavioral (e.g., observer ratings of testing sessions; archival review of actual test reports).

Moreover, the psychological assessment field lacks empirically-based knowledge on appropriate adaptations and modifications when testing diverse populations. This is particularly true for racial and ethnic minority groups. Establishing evidence-based practices related to test selection and test administration are critical areas for future research. When these practices are established, they should be incorporated in graduate training programs and continuing education courses.

Lastly, test developers need to account for cultural diversity when standardization samples are determined. By having diverse groups represented, test developers are better positioned to examine differential test performance and scores. When appropriate and grounded by empirical research, it is recommended that test manuals address the assessment of racial and ethnic minorities and individuals with disabilities.

#### 4.2. Limitations

Although findings from this study provide much needed information on testing practices, there are a

number of limitations to keep in mind. First, participants represented a convenience sample and results could have differed if non-responders had participated. Second, this study did not examine reasons for not making adaptations and modifications. Third, data were self-report and may have under- or over-estimated actual assessment behaviors. Fourth, the important variable of acculturation was not considered in this study and would have likely provided richer information on participants’ testing practices. Fifth, the sample was drawn from populous states and may not be reflective of the demographics of other states.

### 5. Conclusion

When conducting psychological evaluations with diverse populations, it is critical to keep in mind cultural factors that may impact test results. Currently, the APA [23] and the *Standards* [24] provide general guidelines to assist with multicultural assessments. However, the broadness of these guidelines may prevent their applicability and appropriate use in practice. Clearly, evidence-based guidelines are needed to help inform assessments with culturally diverse groups and these guidelines need to be included in test manuals. In addition, as standardized tests are developed, it is essential that they adequately address the cultural and linguistic differences of our diverse population.

### Acknowledgments

We wish to thank Andrea Saul, Carrie Kaufman, and Arthur Wang for their assistance during the course of this project as well as our participants for their willingness to share their experiences. This work was funded by the National Institute on Disability and Rehabilitation Research (NIDRR), U.S. Department of Education, through a grant to the Center for Capacity Building on Minorities with Disabilities Research (Grant# H133A040007). The opinions expressed herein are those of the authors and not necessarily those of NIDRR or the Center. This study was conducted by the authors while they were affiliated with DePaul University. We wish to thank the Long Beach VA Healthcare System, Edward Hines, Jr. VA Hospital, YAI Network, and Children’s Hospital Los Angeles for their support in completing this work.

## References

- [1] Dana RH. Multicultural assessment practices for professional psychology. Needham Heights: Allyn & Bacon; 1993. pp. 1-130.
- [2] Hernandez B, Horin EV, Donoso OA, Saul A. Psychological testing and multicultural populations. In: Balcazar F, Suarez-Balcazar Y, Taylor-Ritzler T, Keys C, editors. *Race, culture and disability: Rehabilitation, science and practice*. Boston: Jones and Bartlett; 2010. pp. 55-79.
- [3] Olkin R. What psychotherapists should know about disability. New York: The Guilford Press; 1999. pp. 1-358.
- [4] Padilla AM. Issues in culturally appropriate assessment. In: Suzuki LA, Ponterotto JG, Meller PJ, editors. *Handbook of multicultural assessment: Clinical, psychological, and educational applications*. 2nd ed. San Francisco: Jossey-Bass; 1995. pp. 5-27.
- [5] Suzuki LA, Kugler JF (1995). Intelligence and personality assessment: Multicultural perspectives. In: Ponterotto JG, Casas JM, Suzuki LA, Alexander CM, editors. *Handbook of multicultural counseling*. Thousand Oaks, CA: Sage Publications; 1995. pp. 493-515.
- [6] Sue DW, Bingham RP, Porche-Burke L, Vasquez M. The diversification of psychology: A multicultural revolution. *Am Psychol*. 1999; 54:1061-1069.
- [7] Ory MG, Lipman PD, Barr R, Harden JT, Stahl SM. A national program to enhance research on minority aging and health promotion. *Journal of Mental Health & Aging*. 2000;6:9-18.
- [8] U.S. Census Bureau [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2010a. The Black population: 2010; 2010 September 29 [cited 2011 Sept 1]. Available from: <http://www.census.gov/prod/cen2010/briefs/c2010br-06.pdf>
- [9] U.S. Census Bureau [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2010b. The Hispanic population: 2010; 2011 March 24 [cited 2011 Sept 1]. Available from: <http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>
- [10] U.S. Census Bureau. [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2010c. Overview of race and Hispanic Origin: 2010; 2011 March 24 [cited 2011 Sept 1]. Available from: <http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>
- [11] U.S. Census Bureau [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2010d. Population distribution and change: 2000 to 2010; 2011 March 24 [cited 2011 Sept 1]. Available from: <http://www.census.gov/prod/cen2010/briefs/c2010br-01.pdf>
- [12] Kleinman A. What is a psychiatric diagnosis? In: Kleinman A, editor. *Rethinking psychiatry: From cultural category to personal experience*. New York: The Free Press; 1998. pp. 5-17.
- [13] Kleinman A. Do psychiatric disorders differ in different cultures?: The methodological questions and the findings. In: Kleinman A, editor. *Rethinking psychiatry: From cultural category to personal experience*. New York: The Free Press; 1998. pp. 18-33.
- [14] Sedlacek WE, Kim SH. Multicultural assessment. Greensboro: ERIC Clearinghouse on Counseling and Student Services. (ERIC Document Reproduction Service No. ED 391112); 1995. pp. 1-2.
- [15] Olkin R, Pledger C. Can disability studies and psychology join hands? *Am Psychol*. 2003;58:296-304.
- [16] Hughes RB, Robinson-Whelen S, Taylor HB, Peterson NJ, Nosek MA. Characteristics of depressed and nondepressed women with physical disabilities. *Arch Phys Med Rehab*. 2005;86:473-9.
- [17] Radnitz CL, Bockian N, Moran AI. Assessment of psychopathology and personality in people with physical disabilities. In: Frank RG, Elliott TR, editors. *Handbook of rehabilitation psychology*. Washington, D.C.: American Psychological Association; 2000. pp. 287-309.
- [18] Mensh E, Mensh H. *The IQ mythology: Class, race, gender, and inequality*. Carbondale: Southern University Press; 1991.
- [19] Alston RJ, McCowan CJ. African American women with disabilities: Rehabilitation issues and concerns. *J Rehabil*. 1994;60:36-40.
- [20] Dana RH. Standard and multicultural assessment. In: Dana RH, editor. *Multicultural assessment: Principles, applications, and examples*. Mahwah: Lawrence Erlbaum Associates; 2005. pp. 1-34.
- [21] Ridley CR, Hill CL, Wiese DL. Ethics in multicultural assessment: A model of reasoned application. In: L. A. Suzuki LA, Ponterotto JG, Meller PJ, editors. *Handbook of multicultural assessment: Clinical, psychological, and educational applications*. 2nd ed. Hoboken: John Wiley & Sons, Inc; 2001. pp. 29-45.
- [22] Hansen ND, Pepitone-Arreola-Rockwell F, Greene AF. Multicultural competence: Criteria and case examples. *Prof Psychol – Res*. 2000;31:652-660.
- [23] American Psychological Association [Internet]. Washington, D.C.: American Psychological Association; c 2002. Ethical principles of psychologists and code of conduct; 2002 August 21 [cited 2011 Sept 1]. Available from: <http://www.apa.org/ethics/code/index.aspx>
- [24] American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Testing individuals with disabilities*. In: *Standards for Education and Psychological Testing*. Washington, D.C.: American Educational Research Association; 1999. pp. 101-8.
- [25] American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, D.C.: American Psychological Association; 1985. pp. 1-98.
- [26] Pullin D. Testing individuals with disabilities: Reconciling social science and social policy. In: Ekstrom RB, Smith D, editors. *Assessing individuals with disabilities in educational, employment, and counseling settings*. Washington, DC: American Psychological Association; 2002. pp. 11-31.
- [27] Sattler JM. *Assessment of children: Cognitive applications*. San Diego: Jerome M. Sattler, Publisher, Inc; 2001.
- [28] U.S. Department of Labor Office of Disability Statistics Policy [Internet]. Washington, DC: U.S. Department of Labor; c 2011 [cited 2011 Sept 1]. Available from: <http://www.dol.gov/odep/topics/DisabilityEmploymentStatistics.htm>
- [29] Gardecki RM. Racial differences in youth employment. *Mon Labor Rev*. 2001;124:51-67.
- [30] National Council on Disability. [Internet]. Washington, DC: National Council on Disability; c 1999. Lift every voice: Modernizing disability program to serve a diverse nation; 1999 December 1 [cited 2011 Sept 1]. Available from: <http://www.ncd.gov/publications/1999/Dec11999>
- [31] U.S. Census Bureau [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2010. American Community Survey: Disability Statistics; 2010 [cited 2011 Sept 1]. Available from: <http://www.disabilitystatistics.org/reports/acs.cfm?statistic=2>

- [32] Hayward BJ, Schmidt-Davis H. Final report 1: How consumer characteristics affect access to, receipt of, and outcomes of VR services. Report prepared under Contract Number HR92-022-001 for the Rehabilitation Services Administration, U.S. Department of Education. Research Triangle Park: Research Triangle; 2003. pp. ES-1-6-4.
- [33] Kaye HS. *Vocational rehabilitation in the United States*. In: Disability Statistics Abstract 20. Washington, D.C.: U.S. Department of Education, National Institute on Disability and Rehabilitation Research; 1998. pp. 2-5.
- [34] U.S. Department of Education [Internet]. Washington, D.C.: U.S. Government Printing Office; c2006a. VR expenditures by category and percentage change from a year ago: Table 16. Rehabilitation Services Administration Fiscal Year 2005; 2006 [cited 2011 Sept 1]. Available from: <http://www.ed.gov/rschstat/eval/rehab/statistics.html>
- [35] U.S. Department of Education [Internet]. Washington, D.C.: U.S. Government Printing Office; c 2006b. Eligibility Rate: Table 15. Rehabilitation Services Administration Fiscal Year 2005; 2006 [cited 2011 Sept 1]. Available from: <http://www.ed.gov/rschstat/eval/rehab/statistics.html>
- [36] Leahy M, Chan F, Saunders J. Job functions and knowledge requirements of certified rehabilitation counselors in the 21st century. *Rehab Couns Bull*. 2003;46:66-81.
- [37] Bellini J. Counselors' multicultural competencies and vocational rehabilitation outcomes in the context of counselor-client racial similarity and difference. *Rehab Couns Bull*. 2003;46:164-73.
- [38] Capella ME. Inequities in the VR system: Do they still exist? *Rehab Couns Bull*. 2002;45:143-153.
- [39] Dziekan KI, Okocha AAG. Accessibility of rehabilitation services: Comparison by racial-ethnic status. *Rehab Couns Bull*. 1993;36:183-90.
- [40] Kundu MM, Dutta A, Walker S. Participation of ethnically diverse personnel in state-federal vocational rehabilitation agencies. *J Appl Rehabil Couns*. 2006;37:30-7.
- [41] Moore CL. Racial and ethnic members of under-represented groups with hearing loss and VR services: Explaining the disparity in closure success rate. *J App Rehabil Couns*. 2001;32:15-23.
- [42] Moore CL, Feist-Price S, Alston RJ. VR services for persons with severe/profound mental retardation: Does race matter? *J Appl Rehabil Couns*. 2002;45:162-7.
- [43] Moore, CL, Giesen JM, Cavanaugh BS. Latino VR access rates by disability type and proportions in the general population with disabilities. *J Appl Rehabil Couns*. 2005;36:25-32.
- [44] Thomas DF, Rosenthal DA, Banks ME, Schroeder M. Diversity in vocational rehabilitation: People, practice, and outcomes. Community-based rehabilitation: Research for improving employment outcome conference 2002 State of the science conference proceedings; 2002. pp. 165-187.
- [45] Wilson KB. Predicting vocational rehabilitation eligibility based on race, education, work status, and source of support at application. *Rehab Couns Bull*. 2000;43:97-105.
- [46] Hambleton RK. The rise and fall of criterion referenced measurement. *Ed Measu*. 2005;13:21-26.
- [47] Ryan CP. Work isn't what it used to be: Implications, recommendations, and strategies for vocational rehabilitation. *J Vocat Rehabil*. 1995;61:7-11.
- [48] Donoso OA, Hernandez B, Horin EV. Use of psychological tests within vocational rehabilitation. *J Vocat Rehabil*. 2010; 32:191-200.
- [49] Camara, WJ, Nathan JS, Puente AE. Psychological test usage: Implications in professional psychology. *Prof Psychol – Res Pr*. 2000;31:141-154.
- [50] Miles MB, Huberman AM. *Qualitative data analysis*. 2nd ed. Thousand Oaks: Sage; 1994. pp. 1-352.
- [51] Hill-Briggs F, Dial JG, Moreere DA, Joyce A. Neuropsychological assessment of persons with physical disability, visual impairment or blindness, and hearing impairment of deafness. *Arch Clin Neuropsych*. 2007; 22:389-404.
- [52] Pearson Education, Inc. Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV). San Antonio: Psychological Corporation; 2008.
- [53] Abedi J, Hofstetter CH, Lord C. Assessment accommodations for English language learners: Implications for policy-based empirical research. *Rev Educ Res*. 2004;74:1-28.
- [54] Sireci SG, Li S, Scarpatti S. The effects of test accommodation on test performance: A review of literature. Center for Educational Assessment Research Rep. No. 485. Amherst: University of Massachusetts; 2003. pp. 1-100.
- [55] Bernal G, Saez-Santiago E. Culturally centered psychosocial interventions. *J Commun Psychol*, 2006;34:121-31.
- [56] Sue S. In search of cultural competence in psychotherapy and counseling. *Am Psychol*. 1998;53:440-8.
- [57] Ysseldyke JE, Thurlow ML, Bielinski J, House A, Moody M, Haigh J. The relationship between instructional and assessment accommodations in an inclusive state accountability system. *J Learn Disabil*. 2001;34:212-20.
- [58] Tindal G, Fuchs L. A summary of research on test changes: An empirical basis for defining accommodations. Lexington: University of Kentucky, Mid-South Regional Resource Center; 1999. pp. 1-125.
- [59] Allen J. A multicultural assessment supervision model to guide research and practice *Prof Psychol – Res Pr*. 2007;38: 248-58.

Copyright of Journal of Vocational Rehabilitation is the property of IOS Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.