Psychological Assessment
1999, Vol. 11, No. 2, 186–197

1040-3590/99/$3.00

Ethnicity and Psychopathology: A Meta-Analytic Review of 31 Years of Comparative MMPI/MMPI-2 Research

Gordon C. Nagayama Hall, Anita Bansal, and Irene R. Lopez
Kent State University

Meta-analyses were performed on 25 comparative Minnesota Multiphasic Personality Inventory (MMPI) and MMPI-2 studies of 1,428 male African Americans versus 2,837 male European Americans, 12 studies of 1,053 female African Americans versus 1,470 female European Americans, and 13 studies of 500 male Latino Americans and 1,345 male European Americans. Aggregate effect sizes suggest higher scores for ethnic minority groups than for European Americans on some MMPI/MMPI-2 scales and lower scores on others. However, none of the aggregate effect sizes suggest substantive differences from either a statistical or clinical perspective. The MMPI and MMPI-2 apparently do not unfairly portray African Americans and Latinos as pathological. Effect sizes across studies generally did not vary as a function of sociodemographic variables, research setting, or use of the MMPI versus MMPI-2. It is recommended that additional between- and within-ethnic groups psychopathology research continue.

In an increasingly multicultural society, ethnic differences in psychopathology could have far reaching implications. Perhaps the most benevolent implication of ethnic differences would be the need for culture-specific models of psychopathology, assessment, and treatment (Florsheim, Tolan, & Gorman-Smith, 1996; Okazaki, 1997). However, cultural differences have traditionally been regarded in our society as deficiencies (Jones, 1988). Ethnic differences in psychopathology could result in ethnic minority persons receiving different, and possibly negative, treatment in educational, employment, legal, mental health, and other settings in which measures of psychopathology are used to determine one's status. Thus, when ethnic differences in psychopathology are reported, the validity of such differences has been questioned (Okazaki & Sue, 1995b).

Epidemiological data on psychopathology suggest few ethnic differences in rates of psychopathology. In the Epidemiological Catchment Area Project (ECAP), African Americans had significantly greater lifetime prevalence than European Americans only of simple phobia, agoraphobia, and cognitive impairment (Robins et al., 1984). In the same project, European Americans had a significantly greater prevalence of drug abuse and major depressive disorders than did Mexican Americans, but there were not significant differences for other disorders (Karno et al., 1987). More recent data also suggest comparable rates of psychopathology across ethnic groups (Huertin-Roberts, Snowden, & Miller, 1997; Kessler et al., 1994; Roberts & Sobhan, 1992). In contrast to the ECAP, the prevalence of *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; *DSM–III–R;* American Psychiatric Association, 1987) disorders in the National Comorbidity Survey (NCS) among African Americans versus European Americans was not higher for any disorder, and it was significantly lower for affective, substance-use disorders, and comorbidity (Kessler et al., 1994). Latinos in the NCS had a higher prevalence than European Americans only of affective disorders over the past year and comorbidity (Kessler et al.). Thus, the epidemiological findings appear to suggest that ethnicity is not a significant source of variance in psychopathology.

Sources of Potential Ethnic Differences

Why should differences in psychopathology be expected among ethnic groups in the United States? Cultural differences may persist across ethnic groups even among persons who are not recent immigrants and who are highly acculturated. One general dimension along which ethnic groups in the United States differ is individualism–collectivism. Compared to European Americans, ethnic minority persons in the United States have a relatively collectivist orientation (Greenfield, 1994; Hill, Soriano, Chen, & LaFromboise, 1994; Phinney, 1996; D. W. Sue & D. Sue, 1990). Collectivist cultures emphasize attending to others, fitting in, and interpersonal harmony (Markus & Kitayama, 1991). Such cultural differences may be associated with differences in the expression of psychopathology.

Minority status may also influence psychopathology among persons of color in the United States. African Americans are more likely to perceive events as racist than are European Americans (Inman & Baron, 1996). Perceived discrimination may be minimized in some situations to enhance self-esteem and perceptions of control (Ruggiero & Taylor, 1997). Nevertheless, discrimination based on minority status is a stressor that may be more prominent for persons of color than for European Americans. It is possible that the stress associated with discrimination may have negative health and mental health consequences for ethnic minority persons (cf. Myers et al., 1995).

Gordon C. Nagayama Hall, Anita Bansal, and Irene R. Lopez, Department of Psychology, Kent State University.

We thank John McNulty for reviewing an earlier version of this article.

Correspondence concerning this article should be addressed to Gordon C. Nagayama Hall, who is now at the Department of Psychology, 333 Moore Building, Pennsylvania State University, University Park, Pennsylvania 16802. Electronic mail may be sent to gch3@psu.edu.

Definition of Ethnicity

A primary issue when attempting to assess ethnic differences in psychopathology is the definition of ethnicity or race. Ethnicity has been recently defined as referring to broad groupings on the basis of both race and ethnicity (Phinney, 1996). In keeping with Phinney's (1996) definition, *ethnicity* in this article will encompass *race,* a term on which there is disagreement as to its meaning, and it will focus on members of nondominant groups in the United States. In most research, it is assumed that persons of the same ethnicity or race share psychological characteristics associated with culture and that these characteristics are associated with psychopathology (Okazaki & Sue, 1995b).

Not all persons share the same degree of identification with their ethnic group (Dana, 1988; Greene, 1987; Suzuki & Kugler, 1995). Moreover, not all individuals in an ethnic group are similar, and within-ethnic group heterogeneity may be as important as between-groups heterogeneity (Greene, 1987; Jones, 1991; Okazaki & Sue, 1995b; Suzuki & Valencia, 1997). One method of addressing within-group heterogeneity is to match or statistically control for demographic characteristics, such as age, gender, and socioeconomic status. However, such matching cannot control for all potentially relevant variables, such as degree of societal discrimination experienced, and many demographic variables are confounded with ethnicity (Phinney, 1996). Moreover, it is possible that some of these demographic variables are causal variables and that ethnicity obscures such causality because it is confounded with these variables (cf. Meehl, 1971).

The expected effects of ethnic identification on psychological functioning are not clear. There is some evidence that a stronger ethnic identification is associated with better psychological functioning, such as increased self-esteem (Phinney, 1989). However, there is other evidence that being ethnic identified may be maladaptive when there are negative stereotypes associated with the ethnic domain, such as in academic achievement (Fordham, 1988; Fordham & Ogbu, 1986; Steele, 1997). However, the benefits of ethnic disidentification (e.g., possible academic achievement) may incur psychological costs, such as depression (Arroyo & Zigler, 1995). Similarly, the adaptiveness of ethnic identification may be associated with social context (Tanaka, Ebreo, Linn, & Morera, 1998). For example, a strong ethnic identification may serve as a buffer against discrimination for situations in which a person is a numerical minority, yet it also could interfere with one's functioning within the majority social context.

Although it is an important construct, there is no generally accepted method of measuring ethnic identity. Several different ethnic-identity scales have recently been developed. However, most of the scales do not have well-established psychometric properties, and those that do are not necessarily useful in clinical samples (Kohatsu & Richardson, 1996; Sabnani & Ponterotto, 1992). Moreover, few of the scales have been designed to measure ethnic identity across multiple ethnic groups (see Phinney, 1992, for an exception).

Measurement Issues

A major issue is how to measure psychopathology across ethnic groups. *Emic* or culture-specific measures may capture unique cultural characteristics, but may not allow between-ethnic groups comparisons (Dana, 1988, 1993; Okazaki & Sue, 1995a). *Etic* approaches are based on the assumption that the same constructs exist across cultures, which allows between-groups comparisons. Diagnostic judgments by clinicians have been used to identify ethnic differences in psychopathology, but clinical judgments for some diagnoses tend to be biased against ethnic minority persons (Garb, 1997; Snowden & Cheung, 1990). Presumably, self-reports on standardized measures of psychopathology might be less negatively biased than the judgments of others. However, measures to assess psychopathology are not necessarily equivalent across cultural groups, even when attempts are made to establish equivalence, such as including ethnic minority groups in standardization samples and language translation. Tests tend to favor persons from the culture in which the test was developed (Dana, 1988; Mackenzie, 1984). Despite these methodological problems, however, the assessment of ethnic differences is important in determining the applicability and limits of theories and tests of psychopathology (Ben-Porath, 1990; Okazaki & Sue, 1995b; Triandis & Brislin, 1984).

Although there is no gold standard of psychopathology, the Minnesota Multiphasic Personality Inventory (MMPI) is one of the most widely used measures of psychopathology. It is also one of the most psychometrically reliable and valid tests, and it is able to effectively discriminate between pathological and nonpathological groups (Parker, Hanson, & Hunsley, 1988; Zalewski & Gottesman, 1991). There also have been more studies of ethnic differences using the MMPI and the MMPI-2 than any other measure. Thus, the criterion measures of psychopathology in the current review will be the MMPI and MMPI-2.

Some have contended that the MMPI-2 is culturally biased (Dana & Whatley, 1991; Park, Upshaw, & Koh, 1988), whereas others have contended that it is not (Butcher & Williams, 1992; Weiner, 1995). An absence of ethnic-group differences is not sufficient evidence that a test is not biased. Test scores may or may not be associated with the same extra-test behavior across ethnic groups (Butcher, Graham, & Ben-Porath, 1995; Greene, 1987; McNulty et al., 1997; Pritchard & Rosenblatt, 1980; Timbrook & Graham, 1994). However, the same measures of extra-test behaviors are not consistently used across MMPI/MMPI-2 ethnic differences studies, and often these measures are less reliable and valid than the MMPI/MMPI-2 itself. Moreover, the measurement of such extra-test behavioral criteria (e.g., diagnosis, clinical ratings, other test scores) typically is not culturally sensitive, and culturally unbiased criteria for psychopathology are difficult to identify (Adebimpe, 1994; Garb, 1997; Gynther, 1989; Snowden & Cheung, 1990). Test bias will not be directly addressed in this review.

A related measurement issue involves the functional equivalence of behavior measured by a test, or whether the same behavior exhibited by two different ethnic groups is interpreted similarly or has the same impact in the individual's context. One method of assessing functional equivalence is to determine if persons from different ethnic groups in particular settings have similar scores. For example, if the MMPI/MMPI-2 scores of two ethnic groups were not significantly different across studies in psychiatric inpatient settings, it could be contended that the same levels of psychopathology are required for hospitalization, and that the same types of psychopathological behaviors are not interpreted differently. Conversely, if the groups exhibited significantly different

scores, then it could be contended that the same behaviors are interpreted differently, insofar as the different scores reflect differing hospitalization criteria. Thus, lower levels of psychopathology might be interpreted as warranting hospitalization for one group than for another. Granted, the criteria for psychiatric hospitalization or admission to other clinical settings (e.g., substance-abuse programs) may also be culturally biased (Adebimpe, 1994; Garb, 1997; Gynther, 1989; Snowden & Cheung, 1990).

Ethnicity and the MMPI/MMPI-2

The original MMPI did not include ethnic minorities in its standardization sample. Consequently, some have contended that the test is biased against ethnic minority groups. Some reviews have reported differences, whereas others have not (e.g., Dana, 1988; Greene, 1987; Gynther, 1989; Pritchard & Rosenblatt, 1980). However, ethnic minority persons were included in the standardization sample of the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), and some studies of ethnic differences have been conducted with this revised measure. However, most MMPI/MMPI-2 studies of ethnic groups have not included adequately large, representative samples of ethnic minority groups (Okazaki & Sue, 1995b). One solution to this issue would be to aggregate samples across individual studies through meta-analytic methods. Another reason for a meta-analytic review is that existing reviews have been qualitative (e.g., Greene, 1987; Gynther, 1989) and there have not been recent comprehensive reviews of the MMPI ethnic-differences literature. Moreover, since the earlier reviews of MMPI ethnic differences, the MMPI-2, which included ethnic minority persons in its normative sample, has been developed.

The purpose of this article is to review the MMPI and MMPI-2 research that has compared ethnic groups over the past 31 years. We considered all studies since 1967 that included direct comparisons of the MMPI/MMPI-2 Validity and Clinical scale scores of male and female European Americans and of male and female ethnic minorities. Because we included only comparisons for which there were at least 10 different published studies, meta-analyses of MMPI/MMPI-2 studies only for male and female European Americans, male and female African Americans, and male Latino Americans are included. Where it was possible, we created subgroups of studies conducted in particular settings (i.e., clinical, substance-abuse treatment, forensic, inpatient) to determine if MMPI/MMPI-2 scores across ethnic groups in relatively homogeneous settings are more similar than those in more heterogeneous settings. This subgrouping was conducted in part to provide an estimate of the functional equivalence of psychopathology in these particular settings.

Method

Literature Search and Criteria for Inclusion

Studies included in the PsycInfo database from 1967–1998 that directly compared ethnic minority groups with European Americans on the MMPI or MMPI-2 were considered. Studies that did not report between-groups comparisons were excluded, as were studies that did not report on all Validity and Clinical subscales of the MMPI. Male and female participants' data were analyzed separately because combined analyses may obscure potential gender differences. Studies that did not report separate analyses for male and female participants were excluded. Also excluded were studies that did not report between-groups means and standard deviations or statistical comparisons for each of the three MMPI/MMPI-2 Validity and 10 MMPI/MMPI-2 Clinical scales. Studies in which adolescents were administered the MMPI were considered for inclusion. At least 10 between-groups studies comparing European Americans versus another male or female ethnic group were required to qualify for inclusion in these analyses. This resulted in the identification of 25 usable studies of male African Americans versus European Americans, 12 usable studies of female African Americans, and 13 usable studies of male Latinos versus male European Americans. There were not at least 10 usable studies of Latinas ($N = 4$), or of male or female Asian Americans ($N = 4$) or American Indians ($N = 7$).

Procedure

The individual study was the unit of analysis. Any overlapping samples across studies from the same studies were identified. If more than one study had sample overlap, then the study that analyzed the most inclusive sample was included in the current meta-analyses. For individual studies that compared more than one sample (e.g., African American vs. European American psychiatric inpatients and African American vs. European American forensic patients), each sample comparison was considered as an individual effect size. In studies in which the MMPI and MMPI-2 were both included, MMPI-2 data were used in the meta-analyses. The three Validity and 10 Clinical scales were analyzed because these were consistently analyzed scales across studies and are included on both the MMPI and MMPI-2. We used *K*-corrected *T* scores for each MMPI/MMPI-2 scale unless only raw scores were reported. There were too few studies for analyses that reported non-*K*-corrected *T* scores.

Several studies were identified that statistically controlled for demographic characteristics (e.g., education, intellectual functioning). If these studies also included analyses in which demographic characteristics were not statistically controlled, then the statistically controlled analyses were included in the current meta-analysis. Other studies matched participants on demographic characteristics. In studies in which matched and unmatched analyses were presented, the matched analyses were included in the current meta-analyses. Studies in which participants were generally identified as similar on a demographic characteristic (e.g., socioeconomic status), but in which the participants were not matched or in which the demographic characteristic was not directly analyzed were not considered to be demographically matched studies (e.g., Moore & Handal, 1980).

Calculation of Effect Size

The Pearson product–moment correlation between ethnicity and MMPI/MMPI-2 scales scores was used as an estimate of effect size, using the formulas provided by Rosenthal (1991). Pearson's *r* is more accurate than Cohen's *d* (1988) in estimating effect sizes in samples with unequal sizes (Rosenthal, 1991), which were common among the studies reviewed. For studies that reported no between-groups differences for certain MMPI/MMPI-2 scales and did not provide additional information to compute an exact effect size (e.g., test statistic, mean, standard deviation), a *p* value of .50 was assigned, which is an effect size of 0 (Rosenthal, 1991). After correlational estimates were computed for each study, the correlation coefficients were transformed to Fisher's *z*s to normalize the distribution. Fisher's *z*s were weighted for each study in the current meta-analyses as a function of sample size (i.e., weighting = $N - 3$). Weighted Fisher's *z*s were combined across studies and divided by the sum of the sample size weightings ($N - 3$) for each study for a mean weighted Fisher's *z*. The mean weighted Fisher's *z* was transformed to a correlation coefficient. The correlation coefficient was also transformed to Cohen's *d* (1988) for descriptive purposes using Rosenthal's (1991) formula: $d = 2r/\sqrt{1 - r^2}$.

The statistical significance of the heterogeneity of the combined effect sizes in each meta-analysis was then determined in a chi-square analysis

Table 1
Description of Samples in Studies of Male African Americans and Male European Americans

Study	Age	Ethnicity	Match	Population
Ben-Porath et al. (1995)	Adult	47 AA, 137 EA	No	Outpatient forensic
Boone & Green (1991)	Adolescent	105 AA, 331 EA	No	Incarcerated forensic
Butcher et al. (1983)	Adult	60 AA, 60 EA	Yes	Inpatient psychiatric
Costello et al. (1973)	Adult	49 AA, 49 EA	Yes	Inpatient/outpatient psychiatric
Costello et al. (1973)	Adult	37 AA, 37 EA	Yes	Incarcerated forensic
Davis (1975)	Adult	40 AA, 40 EA	Yes	Inpatient psychiatric
Goldman et al. (1995)	Adult	48 AA, 65 EA	No	College
Holcomb et al. (1984)	Adult	49 AA, 111 EA	Yes	Incarcerated forensic
Holland (1979)	Adult	208 AA, 396 EA	No	Incarcerated forensic
King et al. (1977)	Adult	56 AA, 56 EA	Yes	Nonclinical employment
McCreary & Padilla (1977)	Adult	40 AA, 40 EA	Yes	Outpatient forensic
McNulty et al. (1997)	Adult	42 AA, 225 EA	No	Outpatient mental health
Moore & Handal (1980)	Adolescent	19 AA, 19 EA	No	High school
Nelson et al. (1996)	Adult	14 AA, 101 EA	Yes	Outpatient pain center
Patalano (1978)	Adult	40 AA, 40 EA	No	Inpatient substance abuse
Patterson et al. (1981)	Adult	52 AA, 248 EA	Yes	Inpatient substance abuse
Penk et al. (1982)	Adult	159 AA, 494 EA	No	Inpatient substance abuse
Robyak & Byers (1990)	Adult	57 AA, 57 EA	Yes	Inpatient substance abuse
Sutker & Kilpatrick (1973)	Adult	15 AA, 11 EA	No	College
Timbrook & Graham (1994)	Adult	116 AA, 116 EA	Yes	National sample
Velasquez & Callahan (1990b)	Adult	70 AA, 66 EA	No	Inpatient psychiatric
Walters (1986)	Adult	25 AA, 26 EA	No	Incarcerated forensic, at least one MMPI scale > 89
Walters (1986)	Adult	26 AA, 20 EA	No	Incarcerated forensic, MMPI scales > 71
Walters et al. (1984)	Adult	27 AA, 46 EA	No	Inpatient alcohol abuse
Walters et al. (1984)	Adult	27 AA, 46 EA	No	Inpatient psychiatric

Note. AA = African American; EA = European American. Match = Matching between ethnic groups on demographic variables, or statistical control of such variables. The Minnesota Multiphasic Personality Inventory (MMPI) was used in all studies except Ben-Porath et al. (1995), Goldman et al. (1995), McNulty et al. (1997), and Timbrook & Graham (1994) in which the MMPI-2 was used.

based on the sum of weighted squared deviation Fisher's *z*s. The formula used was χ^2 with $K - 1$ *df*s $= \Sigma(N - 3)/(Z_j - \bar{Z})$, where Z_j is the Z for any single study, $\bar{Z}$ is the mean of all the Zs obtained, K is the number of studies being compared, and N is the number of participants per study (Rosenthal, 1991). Because comparisons were conducted for the 3 MMPI/MMPI-2 Validity and 10 Clinical scales within each of the three sets of studies (male African Americans vs. European Americans, female African Americans vs. European Americans, male Latino Americans vs. European Americans) a p value of .004 (.5/13) was used as the criterion for statistical significance. This correction procedure was implemented for the post hoc comparisons because there were no a priori hypotheses about MMPI/MMPI-2 ethnic differences. To determine the source of effect size heterogeneity, we categorized each study in terms of (a) age of participants (inclusion of adolescents vs. exclusively adults); (b) whether there was matching or statistical control for demographic variables (as discussed above); (c) use of the MMPI versus MMPI-2; (d) whether the participants were in clinical (e.g., psychiatric, forensic, substance abuse) versus nonclinical settings (e.g., school); and (e) the specific setting of the study (i.e., substance abuse vs. other setting; forensic vs. other setting; inpatient, including incarceration vs. other setting).

To determine the number of studies averaging null results required to change the overall significance level of the meta-analyses from statistical significance to nonsignificance (or the converse), we used the formula $X = 19\,s - n$, where X is the number of other studies required to bring the overall p to .50, 19 is the ratio of the total number of nonsignificant (at $p >$ .05) results to the number of significant (at $p < .05$) results expected when the null hypothesis is true, s is the number of studies summarized in the meta-analysis significant at $p < .05$, and n is the number of summarized studies not significant at $p < .05$ (Rosenthal, 1991). This number constitutes the fail safe statistic.

Results

Effect Size and Date of Publication

Pearson correlations between effect size and date of publication of each of the studies were examined. Only MMPI studies were included in these analyses, as the comparison of MMPI/MMPI-2 studies is examined below. None of these correlations was statistically significant (all ps $> .06$). For the MMPI studies involving African American versus European American men, the absolute value of the mean correlation across MMPI scales was .21 (absolute value range = .01 to .54). For the MMPI studies involving African American versus European American women, the absolute value of the mean correlation across MMPI scales was .28 (absolute value range = .05 to .49). For the MMPI studies involving Latino versus European American men, the absolute value of the mean correlation across MMPI scales was .17 (absolute value range = .01 to .41).

Male African Americans Versus Male European Americans

The samples of male African Americans versus European Americans are described in Table 1. Across studies, there were

Table 2
Minnesota Multiphasic Personality Inventory Scale Effect Sizes for Male African Americans Versus Male European Americans

Study	Scale												
	L	*F*	*K*	1	2	3	4	5	6	7	8	9	0
Ben-Porath et al. (1995)	.24	.09	.01	−.12	−.21	−.13	−.14	−.21	−.14	−.17	.02	.40	−.42
Boone & Green (1991)	.10	.45	.00	.33	.25	−.08	−.27	.26	.14	.36	.42	.28	.15
Butcher et al. (1983)	.01	.29	.05	.12	−.40	−.16	−.23	−.26	.23	−.15	.18	.41	−.40
Costello et al. (1973)	.00	.43	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Costello et al. (1973)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Davis (1975)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Goldman et al. (1995)	.33	.28	−.12	.50	.11	.10	.32	−.11	.15	.09	.46	−.05	.01
Holcomb et al. (1984)	.27	.07	.43	−.21	−.18	−.15	−.11	.13	−.13	−.28	−.16	.23	−.46
Holland (1979)	.12	.23	.04	.10	−.10	−.07	.08	.01	−.05	.15	.29	.28	−.14
King et al. (1977)	.32	.25	−.29	.06	.29	−.11	.21	.19	−.51	.11	.25	.72	.29
McCreary & Padilla (1977)	−.04	.23	−.39	−.03	−.02	−.29	−.06	.01	.08	−.06	−.03	.59	−.02
McNulty et al. (1997)	.35	.02	.10	.10	−.03	.04	.04	−.04	−.02	−.15	.05	.30	−.21
Moore & Handal (1980)	.55	.75	.15	.31	.05	.10	−.05	.22	.46	.44	.68	.80	.18
Nelson et al. (1996)	.00	.37	.00	.00	.00	.00	.00	.00	.00	.00	.37	.37	.00
Patalano (1978)	.58	.12	.35	−.15	−.36	−.42	−.08	−.65	−.16	−.37	−.18	.02	−.43
Patterson et al. (1981)	.00	.00	.00	.00	.00	−.26	.00	.00	.00	.00	.00	.00	.00
Penk et al. (1982)	.26	.08	.11	−.52	.08	−.21	.03	.04	−.10	.09	−.03	−.62	.00
Robyak & Byers (1990)	.03	.42	−.02	.06	−.19	−.19	−.19	−.12	−.02	−.20	.00	.08	−.09
Sutker & Kilpatrick (1973)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.40	.00	.00
Timbrook & Graham (1994)	.28	.21	.10	.31	.31	−.03	.33	−.04	−.07	.26	.63	.36	−.05
Velasquez & Callahan (1990b)	.31	.59	.07	.27	.20	.10	.46	.05	.19	.61	.33	.72	.08
Walters (1986)	.36	.43	.12	.23	.04	.25	−.76	.18	.57	.00	.34	−.08	.00
Walters (1986)	−.18	−.06	.28	.26	−.04	−.10	−.34	.17	−.72	.22	.32	.37	−.78
Walters et al. (1984)	.04	.33	−.21	−.13	.09	.00	−.09	−.41	.19	−.06	.11	.02	.14
Walters et al. (1984)	.31	.28	.42	−.34	−.23	−.31	.21	−.14	.14	−.30	.02	.11	−.42
Aggregate	.18	.21	.03	.08	−.02	−.10	−.03	−.02	.00	.03	.17	.21	−.09
Fail safe *N* for aggregate	216	373	12	3	24	35	22	20	25	20	217	339	49

Note. N = 1,428 African American and 2,837 European American participants. Effect sizes are expressed as Cohen's d statistic. Negative values mean that European American participants had higher scores than did African American participants. The aggregate effect size for each scale is weighted by the N of each individual study. The fail safe statistic is the number of studies averaging null results required to change the overall significance level meta-analyses from statistical significance to nonsignificance (or the converse). L = Lie; F = Frequency; K = Correction.

only 12 individual scale differences that were of medium effect size (d = .60; Cohen, 1988) among 325 individual scale comparisons (Table 2). Aggregate effect sizes indicate that African Americans exhibited higher scores than European Americans on Scales Lie (L), Frequency (F), Correction (K), 1, 7, 8, and 9. European Americans exhibited higher scores than African Americans on Scales 2, 3, 4, 5, and 0. There were no between-groups differences across studies on Scale 6. However, all aggregate effect sizes for male African Americans versus European Americans are considered small at best (small effect size, d = .20; medium effect size, d = .60, large effect size, d = 1.00; Cohen, 1988). The fail-safe statistic suggests that the aggregate effect sizes for Scales L, F, 8, and 9 were robust, using a tolerance level of 5 K + 10, where K = the number of studies in the meta-analysis (Rosenthal, 1991).

The aggregate effect size for Scale 4 was heterogeneous, $\chi^2(1, N = 24) = 43.30$, $p < .009$. Effect sizes were significantly different in studies (a) that used the MMPI (−.06) versus the MMPI-2 (.12), $Z = 2.16$, $p < .02$; (b) of nonclinical (.26) versus clinical (−.07) settings, $Z = 1.83$, $p < .04$; and (c) of forensic settings (−.11) versus other settings (.02), $Z = 2.50$, $p < .006$.

The aggregate effect size for Scale 7 was heterogeneous, $\chi^2(1, N = 24) = 52.32$, $p < .0007$. Effect sizes were significantly different in studies of (a) adolescents (.36) versus adults (−.01), $Z = 2.23$, $p < .01$; and (b) substance-abuse settings (−.13) versus other settings (.09), $Z = 1.72$, $p < .04$.

The aggregate effect size for Scale 8 was heterogeneous, $\chi^2(1, N = 24) = 46.24$, $p < .004$. Effect sizes were significantly different in studies (a) of adolescents (.44) versus adults (.14),

Table 3
Description of Studies of Female African Americans and Female European Americans

Study	Age	Ethnicity	Match	Population
Boone & Green (1991)	Adolescent	37 AA, 76 EA	No	Incarcerated forensic
Butcher et al. (1983)	Adult	37 AA, 37 EA	Yes	Inpatient psychiatric
Costello et al. (1973)	Adult	117 AA, 117 EA	Yes	Inpatient/outpatient psychiatric
Goldman et al. (1995)	Adult	58 AA, 66 EA	No	Nonclinical college
Harrison & Kass (1967)	Adult	383 AA, 389 EA	No	Outpatient prenatal clinic
McGill (1980)	Adult	50 AA, 78 EA	No	Nonclinical welfare recipients
McNulty et al. (1997)	Adult	81 AA, 336 EA	No	Outpatient mental health
Moore & Handal (1980)	Adolescent	19 AA, 19 EA	No	Nonclinical high school
Nelson et al. (1996)	Adult	41 AA, 103 EA	Yes	Outpatient pain center
Patalano (1978)	Adult	40 AA, 40 EA	No	Inpatient substance abuse
Sutker & Kilpatrick (1973)	Adult	14 AA, 33 EA	No	Nonclinical college
Timbrook & Graham (1994)	Adult	176 AA, 176 EA	Yes	Nonclinical national sample

Note. AA = African American; EA = European American. Match = Matching between ethnic groups on demographic variables, or statistical control of such variables. The Minnesota Multiphasic Personality Inventory (MMPI) was used in all studies except Goldman et al. (1995), McNulty et al. (1997), and Timbrook & Graham (1994) in which the MMPI-2 was used.

$Z = 2.19$, $p < .01$; (b) in which African American and European American participants were matched (.08) versus those in which participants were not matched (.27), $Z = 1.83$, $p < .03$; (c) of nonclinical (.38) versus clinical (.14) settings, $Z = 1.69$, $p < .05$; (d) of substance-abuse settings (−.03) versus other settings (.24), $Z = 2.08$, $p < .02$; and (e) of inpatient settings (.13) versus other settings (.27), $Z = 2.41$, $p < .008$.

The aggregate effect size for Scale 9 was heterogeneous, $\chi^2(1, N = 24) = 59.48$, $p < .00008$. Effect sizes were significantly different in studies of (a) adolescents (.32) versus adults (.20), $Z = 1.91$, $p < .03$; (b) substance-abuse settings (−.01) versus other settings (.30), $Z = 4.05$, $p < .00003$; and (c) of inpatient settings (.15) versus other settings (.37), $Z = 2.89$, $p < .002$.

Female African Americans Versus Female European Americans

The samples of female African Americans versus European Americans are described in Table 3. Across studies, only 7 of 156 individual scale differences constituted medium effect sizes and there were no large effect sizes (Table 4). Aggregate effect sizes indicate that African Americans exhibited higher scores than European Americans on Scales *L*, *F*, 1, 2, 4, 5, 6, 7, and 8, and lower scores on Scales *K*, 2, 3, and 9. There was no difference across studies on Scale 0. As with the studies of male participants, all aggregate effect sizes in the female African American versus European American studies are considered small. The fail-safe statistic suggests that aggregate effect sizes for Scales 5 and 9 were robust.

The aggregate effect size for Scale 4 was heterogeneous, $\chi^2(1, N = 11) = 40.73$, $p < .000004$. Because only one study was in a substance-abuse setting and only one study was in a forensic setting, these two variables were not examined as potential sources of variance in heterogeneous effect sizes. Effects sizes were significantly different in studies (a) of adolescents (−.40) versus adults (.09), $Z = 1.87$, $p < .03$; (b) of matched (.27) versus nonmatched (−.04) samples, $Z = 2.34$, $p < .009$; (c) in nonclinical (.13) versus clinical (−.04) settings, $Z = 2.89$, $p < .002$; and (d) in studies of inpatient (−.02) versus other (.08) settings, $Z = 2.51$, $p < .006$.

Male Latino Americans Versus Male European Americans

Most of the studies of Latinos were of Mexican Americans (Table 5). Twenty-one of the 169 effect sizes across individual studies constituted medium effects, and one constituted a large effect (Table 6). Aggregate effect sizes suggest that Latino Americans had higher scores than European Americans on the three Validity scales (*L*, *F*, and *K*) and lower scores on all Clinical scales (1, 2, 3, 4, 5, 6, 7, 8, 9, and 0), although all these effect sizes were small. The fail-safe statistic suggests that the aggregate effect sizes were robust only for Scales *L* and 5.

The aggregate effect size for Scale 5 was heterogenous, $\chi^2(1, N = 12) = 49.53$, $p < .0000001$. However, none of the variables considered significantly accounted for this variance.

Discussion

A general conclusion of this study is that MMPI/MMPI-2 differences among European Americans, African Americans, and Latino Americans are trivial. The MMPI and MMPI-2 apparently do not unfairly portray African Americans and Latino Americans as pathological. The aggregate effect sizes for each MMPI/MMPI-2 scale revealed that ethnic minority groups exhibited greater scores than European Americans on some scales and lower scores on others in published studies conducted over the last 31 years. Across studies, male African Americans exhibited higher scores than male European Americans on 7 MMPI/MMPI-2 scales, lower scores on 5 scales, and no difference on 1 scale. Female African Americans exhibited higher scores than female European Americans on 8 scales, lower scores on 4 scales, and no difference on one scale across studies. Latino Americans exhibited higher scores across studies than male European Americans on 3 scales and lower scores on 10 scales. However, most of the aggregate

Table 4
Minnesota Multiphasic Personality Inventory Scale Effect Sizes for Female African Americans Versus Female European Americans

Study	Scale												
	L	*F*	*K*	1	2	3	4	5	6	7	8	9	0
Boone & Green (1991)	.60	.30	−.05	.07	.06	−.26	−.54	.17	.33	.11	.27	.16	.38
Butcher et al. (1983)	.09	.09	−.23	.18	−.18	−.18	.20	.64	.23	−.06	.27	.67	.10
Costello et al. (1973)	.00	.22	−.35	.00	.00	.00	.38	.00	.30	.00	.26	.44	.35
Goldman et al. (1995)	.23	.00	−.31	.00	−.04	−.13	.17	.33	.30	.00	.22	.18	.09
Harrison & Kass (1967)	−.04	.16	−.04	.17	.01	−.11	.02	−.01	.06	.06	.17	.24	−.13
McGill (1980)	−.41	.00	−.50	.00	.00	.00	.00	.43	.00	.00	.00	−.02	.00
McNulty et al. (1997)	.19	−.02	.05	−.14	−.19	−.18	−.08	.13	.00	−.22	−.05	.28	−.18
Moore & Handal (1980)	.51	.25	−.03	.14	.34	−.42	−.01	.30	.38	−.27	.14	.03	.08
Nelson et al. (1996)	.00	.33	.00	.00	.00	.00	.00	.33	.00	.00	.00	.00	.33
Patalano (1978)	.31	.06	−.02	−.30	−.29	−.67	−.69	.71	.05	−.39	−.24	.08	−.38
Sutker & Kilpatrick (1973)	.00	.00	.00	.00	.00	.00	.78	.00	.00	.00	.00	.00	.00
Timbrook & Graham (1994)	.20	.20	−.22	.22	.12	.03	.33	.43	−.02	−.21	.20	.31	.01
Aggregate	.10	.13	−.12	.06	−.02	−.11	.06	.19	.09	−.06	.12	.24	.00
Fail safe *N* for aggregate	23	28	30	8	11	25	6	120	15	1	17	98	9

Note. *N* = 1,053 African American and 1,470 European American participants. Effect sizes are expressed as Cohen's *d* statistic. Negative values mean that European American participants had higher scores than did African American participants. The aggregate effect size for each scale is weighted by the *N* of each individual study. The fail-safe statistic is the number of studies averaging null results required to change the overall significance level meta-analyses from statistical significance to nonsignificance (or the converse). *L* = Lie; *F* = Frequency; *K* = Correction.

effect sizes were not statistically robust. Robust aggregate effect sizes were found only for Scales *L*, *F*, 8, and 9 in studies that compared male African Americans and European Americans, Scales 5 and 9 in studies that compared female African Americans and European Americans, and Scales *L* and 5 in studies that compared male Latino Americans and European Americans. Studies that are published are typically biased toward statistically significant differences (Rosenthal, 1991). Thus, many of the nonrobust aggregate differences are suspect. Although the actual between-ethnic groups differences may be attenuated somewhat because some studies did not report sufficient information to compute exact effect sizes when between-groups differences were not statistically significant, even when differences were found, they were not substantive. None of the aggregate effect sizes, including the robust ones, in the current meta-analyses qualify as a medium effect size. The small aggregate sizes that were found

Table 5
Description of Samples in Studies of Male Latin Americans and Male European Americans

Study	Age	Ethnicity	Match	Population
Hibbs et al. (1979)	Adolescent/adult	38 MA, 51 EA	No	Inpatient psychiatric
Holland (1979)	Adult	114 MA, 396 EA	No	Incarcerated forensic
McCreary & Padilla (1977)	Adult	36 MA, 36 EA	Yes	Outpatient forensic
Nelson et al. (1996)	Adult	18 L, 101 EA	Yes	Outpatient pain center
Page & Bozlee (1982)	Adult	11 L, 11 EA	No	Inpatient substance abuse
Penk et al. (1989)	Adult	54 L, 476 EA	No	Inpatient substance abuse
Plemons (1977)	Adult	18 MA, 44 EA	No	Outpatient psychiatric
Velasquez & Callahan (1990a)	Adult	29 L, 46 EA	No	Inpatient alcohol abuse
Velasquez et al. (1993)	Adult	30 L, 30 EA	Yes	Inpatient schizophrenics
Velasquez et al. (1993)	Adult	14 L, 14 EA	Yes	Inpatient depressives
Velasquez et al. (1993)	Adult	10 L, 10 EA	Yes	Inpatient antisocials
Venn (1988)	Adult	16 MA, 16 EA	Yes	Inpatient alcohol abuse
Weisman et al. (1989)	Adult	112 MA, 114 EA	No	Inpatient substance abuse

Note. MA = Mexican American; EA = European American; L = Latino. Match = Matching between ethnic groups on demographic variables, or statistical control of such variables. All studies used the Minnesota Multiphasic Personality Inventory.

Table 6
Minnesota Multiphasic Personality Inventory Scale Effect Sizes for Male Latinos Versus Male European Americans

Study	Scale												
	L	*F*	*K*	1	2	3	4	5	6	7	8	9	0
Hibbs et al. (1979)	.61	.21	.18	−.32	.35	.10	−.51	−.12	.01	−.03	.21	−.05	.16
Holland (1979)	.19	.05	−.15	−.11	−.10	−.29	−.27	−.42	−.27	−.11	−.08	−.04	−.14
McCreary & Padilla (1977)	.51	−.17	.79	.32	.00	.24	−.36	−.28	−.16	.09	.08	−.39	−.44
Nelson et al. (1996)	.51	.03	.30	.35	.02	.01	.01	−.40	−.06	.13	.09	−.04	−.02
Page & Bozlee (1982)	.00	.00	.00	.00	.93	.00	.00	.00	.00	.00	.00	.00	.00
Penk et al. (1989)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Plemons (1977)	.62	.00	.52	.00	.00	.00	.00	−.62	.00	.00	.00	.00	.00
Velasquez & Callahan (1990a)	.15	−.27	.30	.06	−.35	−.21	−.70	−.90	−.45	−.43	−.37	−.05	−.47
Velasquez et al. (1993)	.61	.73	−.13	.91	.36	.40	.47	.06	.44	.32	.69	.78	.26
Velasquez et al. (1993)	−.09	.66	−.18	.29	.04	−.12	.55	−1.45	.90	.07	.51	.23	.14
Velasquez et al. (1993)	.05	.58	−.26	.64	.39	.44	−.07	.00	.42	.24	.66	.88	−.19
Venn (1988)	.47	.01	−.16	−.17	−.38	−.60	−.60	−.51	.03	−.63	−.42	−.12	.07
Weisman et al. (1989)	.16	−.26	−.23	−.33	−.13	−.46	−.41	−.87	−.39	−.35	−.35	−.33	−.07
Aggregate	.21	.02	.01	−.01	−.02	−.12	−.16	−.34	−.12	−.08	−.03	−.04	−.07
Fail safe *N* for aggregate	98	9	9	9	12	2	29	221	7	5	12	13	7

Note. $N = 500$ Latino American and 1345 European American participants. Effect sizes are expressed as Cohen's *d* statistic. Negative values mean that European American participants had higher scores than did Latino participants. The aggregate effect size for each scale is weighted by the *N* of each individual study. The fail-safe statistic is the number of studies averaging null results required to change the overall significance level meta-analyses from statistical significance to nonsignificance (or the converse). *L* = Lie; *F* = Frequency; *K* = Correction.

for all MMPI/MMPI-2 scales in this study constitute less than 5 *T*-score points on any particular MMPI scale, which is not clinically meaningful (Greene, 1987).

The largest aggregate effect size in this study was the difference between male Latino Americans and European Americans on Scale 5, with male Latinos scoring lower on the scale. Although Scale 5 is heterogeneous in content, lower scores may suggest stereotypically masculine preferences in work, hobbies, and other activities (Graham, 1993). This was a robust finding, in that 221 studies would be required to change this significant difference to a nonsignificant one. There may exist subgroups of Latino men having a traditionally masculine gender identity, but it is unknown if such subgroups of men are any larger in Latino cultures than in other cultures (Casas, Wagenheim, Banchero, & Mendoza-Romero, 1995). Moreover, there may be positive aspects of some patriarchal cultural groups, including family nurturance and a sense of community (Sorenson & Siegel, 1992). It is unknown from these analyses whether different extra-test behaviors are associated with Scale 5 for Latino versus European American men. Although the aggregate effect size for Scale 5 in studies of male Latino versus European Americans was heterogeneous, effect sizes did not differ across settings. Moreover, effect sizes in studies for which there was matching based on sociodemographic characteristics did not differ from those in which there was not. Perhaps the failure in these analyses to detect a specific source of the heterogeneity of the aggregate Scale 5 effect size was that most of the sources of heterogeneity that were tested in this study involved *different clinical settings*. Scale 5 is not strictly a measure of psychopathology (Graham, 1993) and may not differ across clinical settings. Although the effect size for Scale 5 for male Latinos versus European Americans was the largest aggregate effect size in the current meta-analyses, it should be kept in mind that this is considered a small effect (i.e., less than 5 MMPI/MMPI-2 *T*-score points).

Sociodemographic influences had remarkably little influence on the findings. Twelve of the studies considered in the current meta-analyses included some form of matching of participants on sociodemographic variables. Effect sizes did not differ between studies that matched and did not match participants, with the exceptions of MMPI/MMPI-2 Scale 4 in the comparisons between male and female African Americans and European Americans, and MMPI/MMPI-2 Scale 8 for male African Americans and European Americans. In the male comparisons, ethnic differences were smaller in matched studies, whereas ethnic differences were larger in the matched studies for the female comparisons. Effect sizes in studies that included adolescents were not significantly greater than the effect sizes in studies that included only adults, with the exceptions of Scales 7, 8, and 9 for male African Americans and European Americans, and Scale 4 for female African Americans and European Americans. There is some evidence that elevations on Scales 4, 7, 8, and 9 are generally greater among adolescents than among adults (Graham, 1990). Thus, the larger effect sizes on these scales among adolescents in this study may be a function of a less restricted range of scores (i.e., more variance) relative to adults.

Effect sizes generally were not different in studies that used the MMPI versus the MMPI-2. However, MMPI-2 studies yielded larger effects sizes than the MMPI for Scales 1 and 4 for male African Americans versus European Americans and for Scale 4 for female African Americans versus European Americans. Unlike the MMPI, ethnic minority samples were included in the MMPI-2 standardization sample. The current results suggest that inclusion of ethnic minorities in a standardization sample does not necessarily eliminate all ethnic differences. Nevertheless, none of the aggregate individual scale effect size differences in the MMPI-2 studies constituted a medium effect size or greater. The primary conclusion is that differences between European Americans and African Americans, and between male European Americans and male Latinos on both the MMPI and the MMPI-2 are minimal.

These findings provide some evidence that the behavior measured by the MMPI/MMPI-2 is functionally equivalent. Most of the MMPI/MMPI-2 scale effect sizes in the current meta-analyses were not heterogeneous, meaning that effect sizes did not significantly vary across studies or settings. In no instance did any effect size for a subset of studies (MMPI vs. MMPI-2, adolescent vs. adult samples, clinical vs. nonclinical samples, substance-abuse vs. other settings, forensic vs. other settings, inpatient vs. other settings) reach or exceed a medium effect size. These findings may mean that self-reported psychopathology is generally equivalent across ethnic groups in the settings that were examined in this study, which included educational, employment, medical, and psychiatric settings. In most cases in which aggregate effect sizes were significantly heterogeneous, studies from homogeneous settings (i.e., clinical, inpatient) yielded smaller between-ethnic groups differences than did studies from other settings. Thus, it does not appear that there are major ethnic differences in self-reported psychopathology in most clinical and psychiatric settings. Exceptions were for male African Americans and European Americans for which there were somewhat larger effect sizes on Scale 4 in forensic samples versus other samples (African Americans had lower scores in forensic samples) and for Scale 7 in substance-abuse treatment versus other samples (African Americans had lower scores in substance-abuse samples). These exceptions may be a function of a generally greater range of psychopathology in forensic and substance-abuse populations. Alternatively, it could be contended that a lower criterion level of psychopathology was sufficient for African American than for European American men to qualify for involvement in forensic and substance-abuse programs. Nevertheless, these ethnic differences in psychopathology associated with clinical setting occurred for only 2 of the 39 aggregate MMPI/MMPI-2 scale comparisons in this study.

Although the behaviors measured by the MMPI/MMPI-2 appear functionally equivalent in most of the clinical settings assessed in the current meta-analyses, this does not necessarily mean that the behaviors are functionally equivalent in ethnic community contexts. A limitation of the current study is that no extra-test measures were consistently available across studies. It is possible that the behaviors measured by the MMPI/MMPI-2 could have a very different impact in European American versus non-European American communities. Such community impact may not be adequately assessed by the mental health professionals who usually diagnose clients, and who are highly unlikely to be ethnic minority persons (Bernal & Castro, 1994). In fact, there is evidence of clinical judgment bias against ethnic minority persons for certain diagnoses (Garb, 1997; Snowden & Cheung, 1990). One method of reducing bias might be to include ratings of ethnic minority persons by ethnic minority mental health professionals as extra-test measures (cf. S. Sue, Fujino, Hu, Takeuchi, & Zane, 1991). A more portable approach might be to have ethnic minority mental health professionals develop culturally sensitive measures that could be used across settings.

Another culturally sensitive extra-test measure might involve peer ratings by members of the same ethnic group as the person being rated. There is evidence that the correlations between partner ratings and MMPI-2 scores generally do not significantly differ between African Americans and European Americans (Timbrook & Graham, 1994). However, the ethnicity of the partners was not specified in that study and the peer ratings were in response to an existing adjustment scale that was not necessarily designed for use with ethnic minority populations. Peer ratings may also be biased. For example, a friend or spouse could have a positive bias. Less bias might occur in same ethnicity peer ratings from colleagues from employment or educational settings. In general, efforts should be made to gather culturally sensitive extra-test data to determine the contextual impact of the behavior measured by personality tests.

The current results appear consistent with epidemiological data that suggest few ethnic differences in psychopathology (Huertin-Roberts et al., 1997; Karno et al., 1987; Roberts & Sobhan, 1992; Robins et al., 1984). Recent evidence also suggests the consistency of personality structure across cultures (McCrae & Costa, 1997). However, it is possible that there are ethnic differences in psychopathology and the commonly used tests and diagnostic classifications are insensitive to such differences. One possible reason for such insensitivity may be the etic nature of many personality tests and diagnostic classifications, which may exclude important personality characteristics and aspects of psychopathology specific to non-European American cultures. The content of the MMPI and MMPI-2 was not developed with specific attention to the multiple cultural groups in the United States to which the tests would be administered (Dana & Whatley, 1991; Park et al., 1988; Suzuki & Kugler, 1995).

Even if the current lack of ethnic differences in psychopathology was replicated with culturally sensitive test and extra-test measures, there may still exist important ethnic personality differences. It has been contended that a lack of difference between advantaged and disadvantaged groups may mean that the disadvantaged group had made efforts to overcome the disadvantage (Hall & Barongan, 1997). Equal levels of psychopathology between American majority and minority groups may imply that the minority groups are engaging in additional coping with stressors (e.g., societal discrimination) that the majority group experiences less of the time. Thus, another important extra-test construct to measure would be any impediments to psychological health that may differentially affect certain groups.

A general limitation of meta-analysis is that a common database is necessary for inclusion. Thus, some MMPI/MMPI-2 studies of ethnic minority populations were excluded because data were reported that were not comparable with the data from studies that were included (e.g., male and female data combined, data from a single ethnic group reported, selected MMPI/MMPI-2 scales included). A weakness of the current meta-analyses is the limited number of studies that qualified for inclusion. Nevertheless, rela-

tively large numbers of male African Americans (N = 1,428), female Americans (N = 1,053), and male Latino Americans (N = 500) were included in the current meta-analyses. However, only four of the studies involving African Americans and none involving Latinos in the current meta-analyses used the MMPI-2. There were fewer than 10 available comparative studies per group to conduct meta-analyses involving Latinas, Asian Americans, or American Indians.

The MMPI-2 is the most widely used personality instrument and its applicability and limitations with ethnic minority populations need to be more fully investigated. Despite the general absence of ethnic differences in this study, most of the effect sizes were not robust, and more research is necessary to determine whether there are ethnic differences on the MMPI-2. Missing from most of the studies in the current meta-analyses was a consideration of within-group variability that might be associated with psychopathology. For example, most of the studies did not report how participant ethnicity was determined. Presumably, participants designated their ethnicity. Only one study examined the degree of acculturation of the participants. In this study, most participants were identified with Mexican culture (Velasquez & Callahan, 1990a). Alternatively, it is possible that many of the ethnic minority participants in the studies reviewed were relatively disidentified with their ethnic group, insofar as they were involved in mainstream European American institutions (e.g., colleges, mental health centers, psychiatric hospitals, Veterans Affairs hospitals).

Perhaps the emphasis in the currently reviewed studies on between-ethnic groups differences resulted in a lack of attention to within-group differences. However, there is certainly a great degree of variability within all ethnic groups, including European Americans. Such variables as generation in the United States, language fluency, acculturation, ethnic identity, perceived minority status, and discrimination are within-group variables that should be considered. The assessment of ethnicity has become much more sophisticated than simple self-report of one's ethnic group, as was used in most of the currently reviewed studies (Kohatsu & Richardson, 1996; Sabnani & Ponterotto, 1992). Because of this degree of within-group variability and the difficulties in controlling for relevant between-groups demographic differences, Phinney (1996) had contended that between-groups studies should not be conducted and that the sole focus should be on within-group variability. Nevertheless, ethnic differences research is important in determining the applicability and limits of psychological theories and tests (Ben-Porath, 1990; Okazaki & Sue, 1995b; Triandis & Brislin, 1984). Future personality and psychopathology research should examine both between- and within-groups variability.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Adebimpe, V. R. (1994). Race, racism, and epidemiological surveys. *Hospital and Community Psychiatry, 45,* 27–31.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Arroyo, C. G., & Zigler, E. (1995). Racial identity, academic achievement, and the psychological well-being of economically disadvantaged adolescents. *Journal of Personality and Social Psychology, 69,* 903–914.

Ben-Porath, Y. S. (1990). Cross-cultural assessment of personality: The case for replicatory factor analysis. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 8, pp. 27–48). Hillsdale, NJ: Erlbaum.

* Ben-Porath, Y. S., Shondrick, D. D., & Stafford, K. P. (1995). MMPI-2 and race in a forensic diagnostic sample. *Criminal Justice and Behavior, 22,* 19–32.

Bernal, M. E., & Castro, F. G. (1994). Are clinical psychologists prepared for service and research with ethnic minorities? Report of a decade of progress. *American Psychologist, 49,* 797–805.

* Boone, D., & Green, S. B. (1991). Predicting with the MMPI the adjustment of juvenile delinquents to institutionalization: Does gender make a difference? *Journal of Personality Assessment, 57,* 61–76.

* Butcher, J. N., Braswell, L., & Raney, D. (1983). A cross-cultural comparison of American Indian, Black, and White inpatients on the MMPI and presenting symptoms. *Journal of Consulting and Clinical Psychology, 51,* 587–594.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Butcher, J. N., Graham, J. R., & Ben-Porath, Y. S. (1995). Methodological problems and issues in MMPI, MMPI-2, and MMPI-A research. *Psychological Assessment, 7,* 320–329.

Butcher, J. N., & Williams, C. L. (1992). *Essentials of MMPI-2 and MMPI-A interpretation.* Minneapolis: University of Minnesota Press.

Casas, J. M., Wagenheim, B. R., Banchero, R., & Mendoza-Romero, J. (1995). Hispanic masculinity: Myth or psychological schema meriting consideration. In A. M. Padilla (Ed.), *Hispanic psychology: Critical issues in theory and research* (pp. 231–244). Thousand Oaks, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

* Costello, R. M., Fine, H. J., & Blau, B. I. (1973). Racial comparisons on the Minnesota Multiphasic Personality Inventory. *Journal of Clinical Psychology, 29,* 63–65.

Dana, R. H. (1988). Culturally diverse groups and MMPI interpretation. *Professional Psychology: Research and Practice, 19,* 490–495.

Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology.* Boston: Allyn & Bacon.

Dana, R. H., & Whatley, P. R. (1991). When does a difference make a difference? MMPI scores and African-Americans. *Journal of Clinical Psychology, 47,* 400–406.

* Davis, W. E. (1975). Race and the differential "power" of the MMPI. *Journal of Personality Assessment, 39,* 138–140.

Florsheim, P., Tolan, P. H., & Gorman-Smith, D. (1996). Family processes and risk for externalizing behavior problems among African American and Hispanic boys. *Journal of Consulting and Clinical Psychology, 64,* 1222–1230.

Fordham, S. (1988). Racelessness as a factor in Black students' school success: Pragmatic strategy or pyrrhic victory? *Harvard Educational Review, 58,* 54–84.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of acting White." *Urban Review, 18,* 176–206.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice, 4,* 99–120.

* Goldman, V. J., Cooke, A., & Dahlstrom, W. G. (1995). Black–White differences among college students: A comparison of MMPI and MMPI-2 norms. *Assessment, 2,* 293–299.

Graham, J. R. (1990). *MMPI-2: Assessing personality and psychopathology.* New York: Oxford University Press.

Graham, J. R. (1993). *MMPI-2: Assessing personality and psychopathology* (2nd ed.). New York: Oxford University Press.

Greene, R. L. (1987). Ethnicity and MMPI performance: A review. *Journal of Consulting and Clinical Psychology, 55,* 497–512.

Greenfield, P. M. (1994). Independence and interdependence as develop-

mental scripts: Implications for theory, research, and practice. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 1–37). Hillsdale, NJ: Erlbaum.

Gynther, M. D. (1989). MMPI comparisons of Blacks and Whites: A review and commentary. *Journal of Clinical Psychology, 45,* 878–883.

Hall, G. C. N., & Barongan, C. (1997). Prevention of sexual aggression: Sociocultural risk and protective factors. *American Psychologist, 52,* 5–14.

* Harrison, R. H., & Kass, E. H. (1967). Differences between Negro and White pregnant women on the MMPI. *Journal of Consulting and Clinical Psychology, 31,* 454–463.

* Hibbs, B. J., Kobos, J. C., & Gonzalez, J. (1979). Effects of ethnicity, sex and age on MMPI profiles. *Psychological Reports, 45,* 591–597.

Hill, H. M., Soriano, F. I., Chen, S. A., & LaFromboise, T. D. (1994). Sociocultural factors in the etiology and prevention of violence among ethnic minority youth. In L. D. Eron, J. H. Gentry, & P. Schegel (Eds.), *Reason to hope: A psychosocial perspective on violence and youth* (pp. 59–97). Washington, DC: American Psychological Association.

* Holcomb, W. R., Adams, N. A., & Ponder, H. M. (1984). Are separate Black and White MMPI norms needed? An IQ-controlled comparison of accused murderers. *Journal of Clinical Psychology, 40,* 189–193.

* Holland, T. R. (1979). Ethnic group differences in MMPI profile pattern and factorial structure among adult offenders. *Journal of Personality Assessment, 43,* 72–77.

Huertin-Roberts, S., Snowden, L., & Miller, L. (1997). Expressions of anxiety in African Americans: Ethnography and the epidemiological catchment area studies. *Culture, Medicine and Psychiatry, 21,* 337–363.

Inman, M. L., & Baron, R. S. (1996). Influence of prototypes on perceptions of prejudice. *Journal of Personality and Social Psychology, 70,* 727–739.

Jones, J. M. (1988). Racism in black and white: A bicultural model of reaction and evolution. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 117–135). New York: Plenum.

Jones, J. M. (1991). Psychological models of race: What have they been and what should they be? In J. Goodchilds (Ed.), *Psychological perspectives on human diversity in America* (pp. 3–46). Washington, DC: American Psychological Association.

Karno, M., Hough, R. L., Burnam, A., Escobar, J. I., Timbers, D. M., Santana, F., & Boyd, J. H. (1987). Lifetime prevalence of specific psychiatric disorders among Mexican Americans and non-Hispanic Whites in Los Angeles. *Archives of General Psychiatry, 44,* 695–701.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States. *Archives of General Psychiatry, 51,* 8–19.

* King, H. F., Carroll, J. L., & Fuller, G. B. (1977). Comparison of nonpsychiatric Blacks and Whites on the MMPI. *Journal of Clinical Psychology, 33,* 725–728.

Kohatsu, E. L., & Richardson, T. Q. (1996). Racial and ethnic identity assessment. In L. A. Suzuki, P. J. Meller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 611–650). San Francisco: Jossey-Bass.

Mackenzie, B. (1984). Explaining race differences in IQ: The logic, the methodology, and the evidence. *American Psychologist, 39,* 1214–1233.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist, 52,* 509–516.

* McCreary, C., & Padilla, E. (1977). MMPI differences among Black, Mexican-American, and White male offenders. *Journal of Clinical Psychology, 33,* 171–177.

* McGill, J. C. (1980). MMPI score differences among Anglo, Black, and Mexican American welfare recipients. *Journal of Clinical Psychology, 36,* 147–151.

* McNulty, J. L., Graham, J. R., Ben-Porath, Y. S., & Stein, L. A. R. (1997). Comparative validity of MMPI-2 scores of African American and Caucasian mental health center clients. *Psychological Assessment, 9,* 464–470.

Meehl, P. E. (1971). High school yearbooks: A reply to Schwartz. *Journal of Abnormal Psychology, 77,* 143–148.

* Moore, C. D., & Handal, P. J. (1980). Adolescents' MMPI performance, cynicism, estrangement, and personal adjustment as a function of race and sex. *Journal of Clinical Psychology, 36,* 932–936.

Myers, H. F., Kagawa-Singer, M., Kumanyika, S. K., Lex, B. W., Markides, K. S., et al. (1995). Panel III: Behavioral risk factors related to chronic diseases in ethnic minorities. *Health Psychology, 14,* 613–621.

* Nelson, D. V., Novy, D. M., Averill, P. M., & Berry, L. A. (1996). Ethnic comparability of the MMPI in pain patients. *Journal of Clinical Psychology, 52,* 485–497.

Okazaki, S. (1997). Sources of ethnic differences between Asian Americans and White American college students on measures of depression and social anxiety. *Journal of Abnormal Psychology, 106,* 52–60.

Okazaki, S., & Sue, S. (1995a). Cultural considerations in psychological assessment of Asian-Americans. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 107–119). New York: Oxford University Press.

Okazaki, S., & Sue, S. (1995b). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment, 7,* 367–375.

* Page, R. D., & Bozlee, S. (1982). A cross-cultural MMPI comparison of alcoholics. *Psychological Reports, 50,* 639–646.

Park, K. B., Upshaw, H. S., & Koh, S. D. (1988). East Asians' responses to Western health items. *Journal of Cross-Cultural Psychology, 19,* 51–64.

Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103,* 367–373.

* Patalano, F. (1978). Personality dimensions of drug abusers who enter a drug-free therapeutic community. *Psychological Reports, 42,* 1063–1069.

* Patterson, E. T., Charles, H. L., Woodward, W. A., Roberts, W. R., & Penk, W. E. (1981). Differences in measures of personality and family environment among Black and White alcoholics. *Journal of Consulting and Clinical Psychology, 49,* 1–9.

* Penk, W. E., Roberts, W. R., Robinowitz, R., Dolan, M. P., Atkins, H. G., & Woodward, W. A. (1982). MMPI differences of Black and White male polydrug abusers seeking treatment. *Journal of Consulting and Clinical Psychology, 50,* 463–465.

* Penk, W. E., Robinowitz, R., Black, J., Dolan, M., Bell, W., Dorsett, D., Ames, M., & Noriega, L. (1989). Ethnicity: Post-traumatic stress disorder (PTSD) differences among Black, White, and Hispanic veterans who differ in degrees of exposure to combat in Vietnam. *Journal of Clinical Psychology, 45,* 729–735.

Phinney, J. (1989). Stages of ethnic identity development in minority group adolescents. *Journal of Early Adolescence, 9,* 34–49.

Phinney, J. (1992). The Multigroup Ethnic Identity Measure: A new scale for use with diverse groups. *Journal of Adolescent Research, 7,* 156–176.

Phinney, J. S. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51,* 918–927.

* Plemons, G. (1977). A comparison of MMPI scores of Anglo- and Mexican-American psychiatric patients. *Journal of Consulting and Clinical Psychology, 45,* 149–150.

Pritchard, D. A., & Rosenblatt, A. (1980). Racial bias in the MMPI: A methodological review. *Journal of Consulting and Clinical Psychology, 48,* 263–267.

Roberts, R. E., & Sobhan, M. (1992). Symptoms of depression in adolescence: A comparison of Anglo, African, and Hispanic Americans. *Journal of Youth and Adolescence, 21,* 639–651.

Robins, L. N., Helzer, J. E., Weissman, M. M., Oraschel, H., Gruenberg, E., Burke, J. D., & Regier, D. A. (1984). Lifetime prevalence of specific psychiatric disorders in three sites. *Archives of General Psychiatry, 41,* 949–958.

* Robyak, J. E., & Byers, P. H. (1990). Effects of race and severity of alcoholism on MMPIs of male alcoholics. *Journal of Social Behavior and Personality, 5,* 409–416.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Ruggiero, K. M., & Taylor, D. M. (1997). Why minority group members perceive or do not perceive the discrimination that confronts them: The role of self-esteem and perceived control. *Journal of Personality and Social Psychology, 72,* 373–389.

Sabnani, H. B., & Ponterotto, J. G. (1992). Racial/ethnic minority-specific instrumentation in counseling research: A review, critique, and recommendations. *Measurement and Evaluation in Counseling and Development, 24,* 161–187.

Snowden, L. R., & Cheung, F. K. (1990). Use of inpatient mental health services by members of ethnic minority groups. *American Psychologist, 45,* 347–355.

Sorenson, S. B., & Siegel, J. M. (1992). Gender, ethnicity, and sexual assault: Findings from a Los Angeles study. *Journal of Social Issues, 48,* 93–104.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different: Theory and practice* (2nd ed.). New York: Wiley.

Sue, S., Fujino, D. C., Hu, L., Takeuchi, D., & Zane, N. W. S. (1991). Community mental health services for ethnic minority groups: A test of the cultural responsiveness hypothesis. *Journal of Consulting and Clinical Psychology, 59,* 533–540.

* Sutker, P. B., & Kilpatrick, D. G. (1973). Personality, biographical, and racial correlates of sexual attitudes and behavior. *Proceedings of the 81st Annual American Psychological Association Convention, 8,* 262–262.

Suzuki, L. A., & Kugler, J. F. (1995). Intelligence and personality assessment: Multicultural perspectives. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 493–515). Thousand Oaks, CA: Sage.

Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence: Educational implications. *American Psychologist, 52,* 1103–1114.

Tanaka, J. S., Ebreo, A., Linn, N., & Morera, O. F. (1998). Research methods: The construct validity of self-identity and its psychological implications. In L. C. Lee & N. W. S. Zane (Eds.), *Handbook of Asian American psychology* (pp. 21–79). Thousand Oaks, CA: Sage.

* Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6,* 212–217.

Triandis, H. C., & Brislin, R. W. (1984). Cross-cultural psychology. *American Psychologist, 39,* 1006–1016.

* Velasquez, R. J., & Callahan, W. J. (1990a). MMPI comparisons of Hispanic- and White-American veterans seeking treatment for alcoholism. *Psychological Reports, 67,* 95–98.

* Velasquez, R. J., & Callahan, W. J. (1990b). MMPIs of Hispanic, Black, and White *DSM-III* schizophrenics. *Psychological Reports, 66,* 819–822.

* Velasquez, R. J., Callahan, W. J., & Young, R. (1993). Hispanic-White MMPI comparisons: Does psychiatric diagnosis make a difference? *Journal of Clinical Psychology, 49,* 528–534.

* Venn, J. (1988). MMPI profiles of Native-, Mexican-, and Caucasian-American male alcoholics. *Psychological Reports, 62,* 427–432.

Walters, G. D. (1986). Screening for psychopathology in groups of Black and White prison inmates by means of the MMPI. *Journal of Personality Assessment, 5,* 257–264.

Walters, G. D., Greene, R. L., & Jeffrey, T. B. (1984). Discriminating between alcoholic and nonalcoholic Blacks and Whites on the MMPI. *Journal of Personality Assessment, 48,* 486–488.

Weiner, I. B. (1995). Psychometric issues in forensic applications of the MMPI-2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI-2* (pp. 48–81). Thousand Oaks, CA: Sage.

* Weisman, C. P., Anglin, M. D., & Fisher, D. G. (1989). The MMPI profiles of narcotics addicts: II. Ethnic and criminal history effects. *International Journal of the Addictions, 24,* 881–896.

Zalewski, C. E., & Gottesman, I. I. (1991). (Hu)Man versus mean revisited: MMPI group data and psychiatric diagnosis. *Journal of Abnormal Psychology, 100,* 562–568.

Received June 15, 1998
Revision received November 16, 1998
Accepted November 24, 1998 ■