Journal of Abnormal Psychology
1995, Vol. 104, No. 3, 464-470

0021-843X/95/$3.00

Ethnicity and Gender in Scales of Psychosis Proneness and Mood Disorders

Phillip M. Chmielewski
University of Illinois at Urbana-Champaign

Leyan O. L. Fernandes
University of Illinois at Urbana-Champaign and
Boston Veterans Affairs Medical Center

Cindy M. Yee
University of California, Los Angeles

Gregory A. Miller
University of Illinois at Urbana-Champaign

Data for Caucasian, African American, Asian American, and Latino college-student samples were compared for several popular self-report scales of psychopathology. Significant group differences were obtained for all scales, with the Caucasian sample consistently having the lowest means. Some gender effects and interactions with ethnic group were also observed. The authors discuss implications of these findings for use of these scales, including implications for use of Caucasian norms with other ethnic samples.

Paper-and-pencil questionnaires are widely used by researchers studying personality or psychopathology in outpatient and nonpatient samples. This method may serve a variety of purposes, including identification of individuals at elevated risk for psychopathology (e.g., Chapman & Chapman, 1985, 1987; Chapman, Chapman, Kwapil, Eckblad, & Zinser, 1994; Edell, 1995), selection of individuals with elevated base rates of personality disorders (e.g., Rosenberger & Miller, 1989), and validation in nonpatient samples of taxonomic distinctions observed in patient samples (e.g., Klein & Miller, 1993). When the goal is the prevention of severe psychopathology, researchers hope that these measures will provide early identification of at-risk individuals and thus facilitate research on etiology and the development of preventive interventions. Whereas the majority of studies have used familial relationships to identify individuals at risk for schizophrenia, questionnaires allow selection on the basis of self-report of symptoms and traits thought to be related to the disorder of interest.

A common problem in the use of such scales, however, is the lack of normative data for individuals who are members of ethnic minority groups. This problem may discourage researchers from including minority-group samples in their investigations, thus limiting the scope of research on at-risk groups. In addition to this lack of ethnic-group norms, there are well-documented gender differences in patterns of symptom expression in schizophrenia (Lewine, 1981), in the prevalence of depression (Goldman & Ravid, 1980; Nolen-Hoeksema, 1987; Weissman & Klerman, 1977; Weissman, Leaf, Holzer, Myers, & Tischler, 1984), and in the tendency to seek help for and report symptoms of depression (Nolen-Hoeksema, 1987, 1990; Weissman & Klerman, 1977). Gender differences in schizophrenia and depression raise the possibility of significant gender differences on scales that measure clinical symptoms and other at-risk indicators. To help overcome these problems, we provide ethnic-group and gender norms for several popular scales that were developed to identify individuals at risk for schizophrenia and affective disorders. The norms were derived from large samples of college students in introductory psychology classes, reflecting the most common use of the scales. Although our goal is to provide normative data, it is also important to note the limitations of our data set. Our data may be applicable to other college-age populations, but these norms are "local" in the sense that they were derived from a single university. Thus, caution should be used when considering their relevance to other locales. Furthermore, because these norms were developed on an exclusively undergraduate population, their applicability to other populations is unclear. On the other hand, the present data set provides a basis for the only available systematic, large-scale evaluation of these scales as a function of ethnicity.

Chapman Scales

The development of a series of scales by Loren and Jean Chapman and their colleagues was guided by Meehl's (1962) model of schizotypy and his clinical descriptions of schizotypic signs (Meehl, 1964). The Chapman scales for physical and social anhedonia (Chapman, Chapman, & Raulin, 1976; Eckblad,

Phillip M. Chmielewski and Gregory A. Miller, Department of Psychology, University of Illinois at Urbana-Champaign; Leyan O. L. Fernandes, Department of Psychology, University of Illinois at Urbana-Champaign and Veterans Affair Medical Center, Boston, Massachusetts; Cindy M. Yee, Department of Psychology, University of California, Los Angeles.

This research was supported in part by National Institute of Mental Health Grants MH09624, MH39628, and MH49296; by the University of Illinois Research Board; and by the University of Illinois Beckman Institute for Advanced Science and Technology.

We gratefully acknowledge the assistance of Jean M. Anhalt, Patricia J. Deldin, Lori Ebert, Marci A. Etienne, and Mindy B. Mechanic in data collection and scoring; Lawrence J. Hubert for statistical consultation; and Tamara L. Brown, Anita A. Davis, Marci A. Etienne, Sung Kim, Daniel N. Klein, Steven R. Lopez, Thom Moore, Jack B. Nitschke, and Stanley Sue for comments on an earlier draft.

Correspondence concerning this article should be addressed to Gregory A. Miller, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820.

Chapman, Chapman, & Mishlove, 1982; Mishlove & Chapman, 1985), perceptual aberration (Chapman, Chapman, & Raulin, 1978), and magical ideation (Eckblad & Chapman, 1983) are measures that have been widely used as indicators of psychosis proneness. The Chapman scales contain true–false, trait-descriptor items. Individuals are typically selected on the basis of scores that are two standard deviations or more above the mean score of their same-gender cohort. The Revised Physical Anhedonia Scale (Chapman & Chapman, 1978; Chapman et al., 1976) consists of 61 items and assesses a self-reported deficit in the ability to experience pleasure from typically enjoyable physical stimuli such as food, sex, and pretty settings (e.g., "Beautiful scenery has been a great delight to me."). The Revised Social Anhedonia Scale (Eckblad et al., 1982) is a 40-item scale that assesses the inability to derive nonphysical pleasures from being with people, such as talking, exchanging expressions of feelings, competing, and loving (e.g., "A car ride is much more enjoyable if someone is with me."). The Perceptual Aberration Scale (35 items; Chapman et al., 1978) assesses several aspects of psychoticlike experiences such as bodily discontinuities and unusual sensory experiences (e.g., "I have felt that something outside my body was a part of my body."). The Magical Ideation Scale (30 items; Eckblad & Chapman, 1983) includes questions concerning beliefs in forms of causation that by conventional standards are not valid but magical (e.g., "I have occasionally had the silly feeling that a TV or radio broadcaster knew I was listening to him.").

Perceptual Aberration and Magical Ideation scores tend to correlate highly and are often used in conjunction in participant selection. Reviews are available describing many studies reporting symptom, social skills, cognitive, and psychophysiological abnormalities in individuals identified with the Chapman scales (Chapman & Chapman, 1985, 1987; Edell, 1995; Fernandes & Miller, 1995; Miller & Yee, 1994; Yee, 1995), and the Chapmans have recently reported on a major longitudinal study of their scales (Chapman et al., 1994).

Because of the limited number of ethnic minority individuals available in their undergraduate population, the Chapmans restricted the development of a normative database for those scales to Caucasians (Chapman & Chapman, 1985, 1987). In the only available independent study of the Chapman scales to use ethnic minority individuals, Kelley and Coursey (1992) reported that a small sample of Asian[1] individuals scored significantly higher than Whites on several measures of schizotypy, including the Chapmans' Magical Ideation Scale. On the basis of their own small non-White groups, Chapman and Chapman (1985) anecdotally noted ethnic-group differences in scores on their scales. They commented that Asian students, as a group, score higher than Whites on the Physical Anhedonia Scale and that Black students do so on the Magical Ideation Scale. In addition to possible ethnic-group differences in Chapman-scale scores, it is also possible that scores may vary somewhat by region. They suggested (Chapman & Chapman, 1985) that investigators develop local norms for the definition of deviant performance on their scales and felt that their own minority-group samples were too small to provide meaningful norms, even for their locale.

General Behavior Inventory

The General Behavior Inventory (GBI; Depue & Klein, 1988; Depue, Krauss, Spoont, & Arbisi, 1989; Depue et al., 1981) was designed to identify individuals at risk for serious affective disorder. Individuals with dysthymia are depressed much of the time, whereas those with cyclothymia vary between depression and hypomania. GBI-identified individuals with dysthymia and cyclothymia suffer significant, diagnosable mood problems that may not reach the acute intensity of major depression or mania but can be serious because they are chronic. It is well documented that such individuals are at elevated risk for major mood disorders (Depue & Klein, 1988), and there is some evidence of psychophysiological abnormalities as well (Miller & Yee, 1994; Yee, 1995).

The GBI is a 73-item, self-report inventory that consists of clinical symptom questions, with a 4-point self-rating scale scored dichotomously (1 or 2 vs. 3 or 4). These questions are divided into Dysthymia, Hypomania, and Biphasic scales; scores for the latter two scales are added. Typically, individuals are selected for dysthymia if they score above the 95th percentile on the Dysthymia scale and below the 85th percentile on the Hypomania + Biphasic combined scale. Individuals are selected for cyclothymia if they score above the 95th percentile on both scales. There have been no known publications of ethnic norms for the GBI. In two large pilot studies involving more than 800 nonclinical individuals, Depue and Klein (1988) found no important gender differences.

Chapman and Chapman (1985) recommended that researchers not use their scales with ethnic groups other than Whites unless they develop their own local norms. For nearly a decade, we have been collecting data to produce such norms by mass testing African American, Asian American, Latino, and Caucasian undergraduates at the University of Illinois. The database of ethnic minority individuals is now sizeable enough to serve as an initial normative data set for use in the selection of non-White individuals with the Chapman scales and the GBI. We hope that the availability of these norms will encourage other researchers to begin including minority-group members in their research and to build their own local normative data sets for sample selection.

Method

Participants completed the questionnaire packet in partial fulfillment of the requirements for an Introduction to Psychology class. They were

[1] The choice and significance of the labels used in describing ethnic-group samples is controversial. Both in the existing literature and in an informal survey of colleagues, we found no consensus, although some views were strongly held. In our discussion of prior literature, we have attempted to respect the ethnic-group category labels used in the publications cited. In referring to our own samples, in most cases we use the actual ethnic-group labels included in the self-report questionnaire packets. Sensibilities about these and alternative terms and category boundaries have evolved considerably in the period over which the data were collected, but it seemed essential to keep the labels consistent in the questionnaire packets across semesters of testing because data would be aggregated. As a result of the categories in our data set, we are not in a position to distinguish, for example, between USA-born Asian Americans and foreign-born, recent immigrants from Asian countries, nor can we differentiate mixed-race individuals. A large, predominantly Caucasian campus that draws heavily from a major city (Chicago) with a predominantly ethnic minority population and from dozens of foreign countries is likely to have some heterogeneity within categories, regardless of the set of categories used. Researchers elsewhere will have to judge the relevance of our norms to their settings.

assigned to our study at random through the departmental subject pool office and were typically tested in groups of approximately 100. Because of the manner in which the subject-pool random-assignment algorithm operated, some testing sessions were same-gender groups, but most were mixed-gender groups. Written instructions were provided and were read to participants by an experimenter. Participants were required to stay no more than 50 min to complete their class requirement but were encouraged to stay longer if necessary to complete the questionnaires. Most of them completed the packet within 50 min.

During the testing period, participants completed computer-readable answer sheets for the GBI and the Physical Anhedonia, Social Anhedonia, Perceptual Aberration, and Magical Ideation scales. They indicated their ethnic group by choosing between "White," "Black," "Hispanic," "Asian," and "Other," and they indicated their gender by answering either "male" or "female." They were excluded if their 13-item Infrequency Scale (Chapman, Chapman, & Miller, 1982) score was greater than 2, if they left more than 5% of the Chapman items or more than 5% of the GBI items blank, or if they left the ethnic-group question blank.

Results

From the fall semester of 1984 through the fall semester of 1992, valid questionnaires were collected from 7,757 students. Of those, .0085% (31 women and 35 men) indicated "Other" as their ethnic group. Analyses are based on the remaining 7,691 participants. Table 1 presents sample sizes by ethnic group and gender. Participants were primarily freshmen, with a median age of 18 years. Although most individuals within every ethnic group were age 18, there was a wider dispersion in age for Blacks and Asians than for Whites and Hispanics, Pearson $\chi^2(9, N = 7{,}691) = 30.17$, $p < .001$. Ethnic group was also related to gender balance, Pearson $\chi^2(3, N = 7{,}691) = 39.92$, $p < .001$. For Whites and Blacks, more women than men participated. The number of Hispanic men and women was equal, but more Asian men than women participated. These slight differences in age and gender among ethnic groups presumably reflect the local undergraduate population, at least those enrolling in the class.

Means and standard deviations were computed for each of the four ethnic groups and for each Ethnic × Gender Group interaction using BMDP2D (Dixon, 1992). These values are presented in Table 1. Table 2 presents values used in typical sample selection with these scales, two standard deviations above the same-gender means for the Chapman scales and cross-gender 85th and 95th percentiles for the GBI scales.[2] Skewness was also computed using BMDP2D. All four ethnic groups produced positively skewed distributions for each scale. Generally, larger groups tended to produce distributions with more skew than smaller groups. For example, the scale that showed the widest range of skewness values was Magical Ideation. On that scale, the distribution for Whites had the most skew (.67), followed by the distributions for Blacks (.43), Asians (.36), and Hispanics (.18). In general, this pattern was present to a much smaller degree for the other scales as well.

A two-way (Ethnic Group × Gender) analysis of variance (ANOVA) was run for each scale using BMDP7D (Dixon, 1992). A highly significant main effect for ethnic group was obtained for every scale, with Whites always significantly lower according to Newman–Keuls post hoc tests. Table 3 presents the results of these analyses. The Physical Anhedonia, Revised Social Anhedonia, and Hypomania + Biphasic scales differed as a function of gender, with men scoring higher than women.

In addition to those main effects, there were significant Ethnic Group × Gender interactions for Physical Anhedonia, $F(3, 7683) = 9.63$, $p < .001$, and Revised Social Anhedonia, $F(3, 7683) = 9.52$, $p < .001$. Simple-effects ANOVAs and Newman–Keuls tests were used to interpret the interactions. For Physical Anhedonia, there was a main effect for ethnic group among women, $F(3, 4039) = 121.83$, $p < .001$, with Whites < Hispanics < Asians < Blacks. The simple main effect among men, $F(3, 3644) = 10.84$, $p < .001$, was Whites < Asian = Hispanic < Blacks. For Revised Social Anhedonia, the simple main effect for ethnic group among women, $F(3, 4039) = 45.94$, $p < .001$, was Whites < Hispanics < Asians < Blacks. For men, $F(3, 3644) = 3.86$, $p < .009$, it was Whites = Asians < Blacks = Hispanics. Thus, for both anhedonia scales, the rank order of the ethnic groups was somewhat different for women and men.

Log-linear analyses were conducted using BMDP4F (Dixon, 1992) to determine whether there were differences in the number of participants from each ethnic group who scored above and below the White-sample-based cutoff scores for the Chapman (mean plus two standard deviations) and Depue (95th percentile) questionnaires. These comparisons are more relevant for the typical use of the scales than are the ANOVAs on group means. The log-linear models used a binary classification for each participant: above or below the cutoff for Whites on a given scale. Table 4 presents the results. Generally, a higher proportion of non-White than White samples fell above the cutoffs. This pattern was reversed for women on the Magical Ideation scale, with fewer non-White participants above the White cutoff.

Discussion

Present results demonstrated clear ethnic-group differences in means for all of the scales, with Whites always scoring lowest. None of the non-White groups was consistently the highest. Blacks were highest on both anhedonia scales, Asians on the other two Chapman scales, and Hispanics on the GBI scales (Depue & Klein, 1988; Depue et al., 1989; Depue et al., 1981). On three of the six scales, gender differences were also obtained. Thus, ethnic group and gender appear to be reliable factors associated with scores on those scales.

Aside from the strong significance levels achieved with the present relatively large sample sizes, the actual size of the observed differences is potentially important. The ANOVA and chi-square effect sizes (Cohen, 1988, 1992) reported in Tables 3 and 4 varied considerably across the scales. For small, medium, and large effect sizes, Cohen suggested values of .10, .25, and .40 in ANOVAs and .10, .30, and .50 in chi-square tests. Effect sizes for four of the ethnic group effects in Table 3 approximated Cohen's criterion for small effects, and the effect size for Physical Anhedonia (.18) was midway to the criterion for medium ANOVA effects. The two anhedonia scales provided substantial gender effect sizes (.31 and .22), but other gender effect sizes were trivial. Only one chi-square in Table 4 reached Cohen's criterion for small effects, but that value (.22), for non-

[2] For completeness, Tables 2 and 4 also present cross-gender values for the Chapman scales and within-gender values for the GBI scales, although the Chapmans have recommended using within-gender norms, and Depue has recommended using cross-gender norms.

Table 1
Ethnic-Group and Gender Means and Standard Deviations for Chapman and GBI Scales

	Chapman scales				GBI scales	
Ethnicity and gender	Phys An	Soc An	Per Ab	Mag Id	Dys	Hyp
White (n = 6,490)						
M	10.82	7.15	6.60	9.20	6.96	5.86
SD	6.44	5.13	6.04	5.80	7.84	4.93
Women (n = 3,378)						
M	8.82	5.97	6.57	9.30	7.27	5.68
SD	5.25	4.34	5.86	5.83	8.02	4.96
Men (n = 3,112)						
M	12.99	8.44	6.64	9.09	6.61	6.06
SD	6.90	5.59	6.23	5.76	7.63	4.90
Black (n = 510)						
M	15.03	8.99	7.31	10.15	9.32	7.24
SD	6.72	5.12	5.64	5.26	8.86	5.41
Women (n = 333)						
M	14.64	8.76	7.16	10.06	9.45	6.90
SD	6.53	4.96	5.58	5.10	8.96	5.25
Men (n = 177)						
M	15.75	9.42	7.59	10.32	9.08	7.89
SD	7.03	5.40	5.75	5.55	8.70	5.65
Asian (n = 491)						
M	12.43	7.84	7.76	10.48	8.59	6.36
SD	6.64	4.85	6.05	5.39	8.59	5.11
Women (n = 232)						
M	10.77	7.47	8.20	10.72	9.15	6.16
SD	6.07	4.60	5.92	5.15	9.04	5.01
Men (n = 259)						
M	13.91	8.17	7.36	10.27	8.09	6.54
SD	6.78	5.06	6.14	5.60	8.15	5.20
Hispanic (n = 200)						
M	12.15	8.18	7.32	10.28	10.42	7.30
SD	6.45	5.52	5.19	5.12	9.26	5.66
Women (n = 100)						
M	10.01	6.57	7.15	10.73	10.50	6.89
SD	5.79	4.94	4.78	4.94	9.52	5.51
Men (n = 100)						
M	14.29	9.80	7.49	9.83	10.35	7.70
SD	6.38	5.62	5.59	5.27	9.03	5.80

Note. The General Behavior Inventory (GBI) includes the subscales Dysthymia (Dys) and Hypomania + Biphasic (Hyp; Depue & Klein, 1988; Depue et al., 1989; Depue et al., 1981). Phys An = Revised Physical Anhedonia Scale (Chapman & Chapman, 1978; Chapman et al., 1976); Soc An = Revised Social Anhedonia Scale (Eckblad et al., 1982); Per Ab = Perceptual Aberration Scale (Chapman et al., 1978); Mag Id = Magical Ideation Scale (Eckblad & Chapman, 1983).

White women falling above the White cutoff on the Physical Anhedonia Scale, was closer to Cohen's threshold for medium chi-square effects. In summary, observed effect sizes ranged from trivial to moderate. Of practical importance, only one of the 18 effect sizes in the cutoff-score analyses (Table 4) was nontrivial (at least "small"), even though 12 of the 18 were statistically reliable.

These findings alone do not resolve whether (a) non-White samples have genuinely elevated rates of psychopathology or risk for psychopathology, (b) the scales assess the same constructs with comparable validity but at different levels of severity in different samples, (c) the scales validly assess qualitatively different constructs in different samples, or (d) the validity of each scale differs across groups. For example, it does not follow from the present results that different norms must be used with different ethnic or gender samples. That choice depends on what one assumes about the basis of the group differences obtained here (see also Sackett & Wilk, 1994). The data clearly indicate that White college norms cannot be assumed to be characteristic of other college samples. That is, whereas White norms might be useful with non-White groups for some purposes, those norms may mean different things in different samples.

Of considerable practical interest is that the group differences were less stark for conventional selection criteria than for group means. Although there were many significant effects for ethnic-group and gender means, several of the scales did not differ in the proportion of participants for various groups scoring above the cutoff score that would normally be used for selecting White participants. In some cases, group differences in standard deviations or in the shapes of the distributions compensated to some degree for the strong group-mean differences.

However, the analyses of cutoff scores must be viewed with some caution. Whereas all of the sample sizes were large enough

Table 2
Mean Cutoff Scores and Percentiles for the Chapman and GBI Scales

Ethnicity and gender	*n*	Chapman scales				GBI scales		
		Phys An	Soc An	Per Ab	Mag Id	Dys (95%)	Hyp 95%	Hyp 85%
White	6,490	23.70	17.41	18.68	20.80	24	15	10
Women	3,378	28.47	19.23	18.59	20.67	25	16	10
Men	3,112	25.71	17.54	19.86	21.26	23	15	10
Black	510	25.05	19.22	17.70	20.52	27	18	12
Women	333	27.70	18.68	18.32	20.26	28	18	11
Men	177	29.81	20.22	19.09	21.42	27	18	13
Asian	491	19.32	14.65	18.29	20.96	25	16	11
Women	232	22.91	16.67	20.04	21.02	27	15	11
Men	259	27.47	18.29	19.64	21.47	23	17	11
Hispanic	200	19.32	14.65	18.29	20.96	29	18	13
Women	100	21.59	16.45	16.71	20.61	31	17	12
Men	100	27.05	21.04	18.67	20.37	28	18	14

Note. Values for the Chapman scales are means + 2 *SD*. Phys An = Revised Physical Anhedonia Scale (Chapman & Chapman, 1978; Chapman et al., 1976); Soc An = Revised Social Anhedonia Scale (Eckblad et al., 1982); Per Ab = Perceptual Aberration Scale (Chapman et al., 1978); Mag Id = Magical Ideation Scale (Eckblad & Chapman, 1983). The General Behavior Inventory (GBI) includes the subscales Dysthymia (Dys) and Hypomania + Biphasic (Hyp; Depue & Klein, 1988; Depue et al., 1989; Depue et al., 1981).

to provide a respectable basis for computing means and standard deviations, the smaller samples for all of the non-White groups provide less confident estimates of the shapes of the tails of the distributions. For example, with only 200 Hispanic individuals contributing to the Dysthymia distribution, the number falling above the 95th percentile would be small enough that sampling error might readily produce a different 95th percentile cutoff in a different sample. Specifically, a few individuals at the high end scoring differently could alter the score falling at that percentile. Our analyses of cutoff scores relied solely on White norms because of this potential instability of non-White percentiles in our data set as well as because the applicability of White norms was a fundamental issue for the study. Although every ethnic group produced positively skewed distributions for each scale, the larger groups did tend to show more skew. This indicates that the somewhat more normal distributions of the smaller groups might actually mitigate the concerns over cutoffs mentioned above. Nonetheless, that the groups differ in at least

Table 3
Main Effects and Newman-Keuls Tests for Ethnic Group and Gender

Main effect	Chapman scales				GBI scales	
	Phys An	Soc An	Per Ab	Mag Id	Dys	Hyp
Ethnic group						
$F(3, 7683)$	76.59	23.70	8.34	13.18	28.25	18.71
$p<$	.001	.001	.001	.001	.001	.001
f	.179	.099	.056	.071	.107	.108
Newman-Keuls	W < H = A < B	W < A < H < B	W < B = H < A	W < B = H = A	W < A < B < H	W < A < B = H
Gender						
$F(1, 7683)$	113.52	53.45	0.00	1.41	2.11	7.00
$p<$	.001	.001	*ns*	*ns*	*ns*	.009
f	.306	.221	.002	.020	.043	.051
Women						
M	9.44	6.30	6.73	9.48	7.64	5.84
SD	5.66	4.50	5.83	5.73	8.24	5.01
Men						
M	13.22	8.51	6.76	9.25	6.94	6.23
SD	6.91	5.56	6.19	5.73	7.81	5.01

Note. Phys An = Revised Physical Anhedonia Scale (Chapman & Chapman, 1978; Chapman et al., 1976); Soc An = Revised Social Anhedonia Scale (Eckblad et al., 1982); Per Ab = Perceptual Aberration Scale (Chapman et al., 1978); Mag Id = Magical Ideation Scale (Eckblad & Chapman, 1983). The General Behavior Inventory (GBI) includes the subscales Dysthymia (Dys) and Hypomania + Biphasic (Hyp; Depue & Klein, 1988; Depue et al., 1989; Depue et al., 1981). W = White; B = Black; A = Asian; H = Hispanic; *ns* = not significant ($p > .10$); f is Cohen's (1988) effect-size index for the analysis of variance.

Table 4
Percentages of Participants Above White-Group Cutoff Scores for Chapman and GBI Scales

Gender and ethnicity	Chapman scales				GBI scales	
	Phys An	Soc An	Per Ab	Mag Id	Dys	Hyp
Women						
White	4.1	5.0	5.2	4.9	5.2	5.1
Black	22.8**	12.3**	5.1	2.1*	6.9	7.5
Asian	9.9**	9.1*	6.0	2.2	7.8	4.7
Hispanic	9.0*	7.0	2.0	2.0	13.0*	9.0
χ^2	192.90	34.00	2.42	10.31	14.66	6.15
$p<$	.001	.001	*ns*	.02	.003	*ns*
w	.218	.092	.024	.050	.060	.039
Men						
White	4.3	4.3	5.1	4.2	5.8	6.6
Black	8.5*	4.5	5.1	4.5	10.7*	14.1**
Asian	5.8	3.5	5.4	4.2	6.9	8.5
Hispanic	6.0	5.0	4.0	1.0	11.0	14.0*
χ^2	7.83	0.56	0.30	2.61	10.89	21.44
$p<$	.05	*ns*	*ns*	*ns*	.02	.001
w	.046	.012	.009	.027	.055	.077
Total						
White	4.9	4.3	5.7	4.6	5.3	6.7
Black	11.4**	6.1	5.5	2.9	8.2*	10.6*
Asian	6.5	3.3	6.1	3.3	7.1	7.3
Hispanic	6.5	8.0*	3.0	1.5	12.5**	13.5**
χ^2	39.87	11.10	2.86	8.70	27.25	22.94
$p<$	.001	.02	*ns*	.04	.001	.001
w	.072	.038	.019	.034	.060	.055

Note. Phys An = Revised Physical Anhedonia Scale (Chapman & Chapman, 1978; Chapman et al., 1976); Soc An = Revised Social Anhedonia Scale (Eckblad et al., 1982); Per Ab = Perceptual Aberration Scale (Chapman et al., 1978); Mag Id = Magical Ideation Scale (Eckblad & Chapman, 1983). The General Behavior Inventory (GBI) includes the subscales Dysthymia (Dys) and Hypomania + Biphasic (Hyp; Depue & Klein, 1988; Depue et al., 1989; Depue et al., 1981); *ns* = not significant ($p > .10$). Chi-square and p values listed are for 4-group main effect. Key for 2-group comparisons (group different from White group): * $p < .05$, ** $p < .001$, reflecting Yates-corrected chi-square tests; w is Cohen's (1988) effect-size index for chi-square. All chi-squares have three degrees of freedom.

some parameters of their distributions is presently clearer than what the differences may be at the extreme ends of the distributions. The significance of group differences in means versus cutoff scores may depend on whether one uses the scales to study sample characteristics as a whole or to select individuals for further study.

The results of the present study replicated Kelley and Coursey's (1992) finding that Asian students scored higher on the Magical Ideation scale than did Whites. Our results also supported the anecdotal reports of Chapman and Chapman (1985) that Asians and Blacks scored higher than Whites on Physical Anhedonia and Magical Ideation, respectively. That we observed more group differences than the Chapmans reported presumably reflects our much larger samples and thus greater statistical power. In any case, our results support their questioning whether White norms are appropriate for non-White subject screening.

Across ethnic groups, women and men scored similarly on Dysthymia, whereas males scored significantly higher on Hypomania + Biphasic. The findings for the Dysthymia Scale are not consistent with frequently reported differences in depression between the men and women. However, given that gender differences in depression do not begin to appear until about 14–15 years of age (Kandel & Davies, 1982; Nolen-Hoeksema, 1990; Rutter, 1986) and that college women have not always shown an elevated rate of depression in other studies (Hammen & Padesky, 1977), present findings can be seen as consistent with the most relevant research to date. The obtained differences may have reached significance in a sample closer to the average age of those on which the majority of epidemiological studies of depression have been conducted. That men scored higher on Hypomania + Biphasic may represent a real tendency in college-age men toward episodic increases in activity level without entailing psychopathology. Another possibility is that it may represent a socially reinforced style of self-presentation or a way of perceiving the self that is prevalent among this group. In either case, the findings with the Hypomania + Biphasic scale raise the possibility of gender-specific scoring for that scale. Accordingly, Tables 1 and 2 provide data on these scales separately for women and men, although Depue and colleagues (Depue et al., 1989; Depue & Klein, 1988; Depue et al., 1981) have not used gender-specific scoring in their use of the GBI.

In summary, present results clearly indicate that investigators should at least consider the use of ethnic-group-specific criteria for the Chapman scales and the GBI. Results also raise the issue of gender-specific cutoffs for the Physical and Social Anhedonia scales and the possibility of gender-specific cutoffs for the GBI Hypomania + Biphasic scale. Were there no such group differences, use of these scales with diverse samples would be straightforward, and samples of convenience would suffice for developing norms. The mere fact of the empirical existence of group differences on the scales does not require within-group norming

when using these scales. Such differences do mean that the investigator must face that choice, realizing that the choice may affect the outcome of the research and the impact on the samples. The choice will depend, at least in part, on what the researcher assumes are the reasons for the group differences, such as genuine group differences in psychopathology validly captured in scale scores or group differences in what is normative in terms of actual symptoms, willingness to report symptoms, and so on. As stated earlier, without external cross-validation, it is not possible to determine the extent to which the group differences found here may generalize to other settings. Nor is there a general answer to whether the relatively small effect sizes would be scientifically significant across varied studies with varied goals. We hope that the present results encourage development of appropriate norms and systematic efforts to address those more conceptual issues, so that choices are based on explicit, well-conceived, and well-supported assumptions.

References

Chapman, L. J., & Chapman, J. P. (1978). *Revised Physical Anhedonia Scale.* (Available from L. J. Chapman, Department of Psychology, 1202 West Johnson Street, University of Wisconsin, Madison, WI 53706)

Chapman, L. J., & Chapman, J. P. (1985). Psychosis proneness. In M. Alpert (Ed.), *Controversies in schizophrenia* (pp. 157–174). New York: Guilford Press.

Chapman, L. J., & Chapman, J. P. (1987). The search for symptoms predictive of schizophrenia. *Schizophrenia Bulletin, 13,* 497–503.

Chapman, L. J., Chapman, J. P., Kwapil, T. R., Eckblad, M., & Zinser, M. C. (1994). Putatively psychosis-prone subjects 10 years later. *Journal of Abnormal Psychology, 103,* 171–183.

Chapman, L. J., Chapman, J. P., & Miller, E. N. (1982). Reliabilities and intercorrelations of eight measures of proneness to psychosis. *Journal of Consulting and Clinical Psychology, 50,* 187–195.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of Abnormal Psychology, 87,* 374–382.

Chapman, L. J., Chapman, J. P., & Raulin, M. L. (1978). Body-image aberration in schizophrenia. *Journal of Abnormal Psychology, 87,* 399–407.

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

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Depue, R. A., & Klein, D. N. (1988). Identification of unipolar and bipolar affective conditions in nonclinical and clinical populations by the General Behavior Inventory. In D. L. Dunner, E. S. Gershon, & J. E. Barrett (Eds.), *Relatives at risk for mental disorders* (pp. 179–202). New York: Raven Press.

Depue, R. A., Krauss, S., Spoont, M. R., & Arbisi, P. (1989). General Behavior Inventory identification of unipolar and bipolar affective conditions in a nonclinical university population. *Journal of Abnormal Psychology, 98,* 117–126.

Depue, R. A., Slater, J. F., Wolfstetter-Kausch, H., Klein, D., Goplerud, E., & Farr, D. (1981). A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: A conceptual framework and five validation studies [Monograph]. *Journal of Abnormal Psychology, 90,* 381–437.

Dixon, W. J. (Ed.). (1992). BMDP statistical software [computer software]. Berkeley: University of California Press.

Eckblad, M., & Chapman, L. J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology, 51,* 215–225.

Eckblad, M. L., Chapman, L. J., Chapman, J. P., & Mishlove, M. (1982). *The revised social anhedonia scale.* (Available from L. J. Chapman, Department of Psychology, 1202 West Johnson Street, University of Wisconsin, Madison, WI 53706)

Edell, W. S. (1995). The psychometric measurement of schizotypy using the Wisconsin scales of psychosis proneness. In G. A. Miller (Ed.), *The behavioral high-risk paradigm in psychopathology* (pp. 3–46). New York: Springer-Verlag.

Fernandes, L. O. L., & Miller, G. A. (1995). Compromised performance and abnormal psychophysiology associated with the Wisconsin psychosis-proneness scales. In G. A. Miller (Ed.), *The behavioral high-risk paradigm in psychopathology* (pp. 47–87). New York: Springer-Verlag.

Goldman, N., & Ravid, R. (1980). Community surveys: Sex differences in mental illness. In M. Guttentag, S. Salasin, & D. Belle (Eds.), *The mental health of women* (pp. 31–55). New York: Academic Press.

Hammen, C. L., & Padesky, C. A. (1977). Sex differences in the expression of depressive responses on the Beck Depression Inventory. *Journal of Abnormal Psychology, 86,* 609–614.

Kandel, D. B., & Davies, M. (1982). Epidemiology of depressive mood in adolescents: An empirical study. *Archives of General Psychiatry, 39,* 1205–1212.

Kelley, M. P., & Coursey, R. D. (1992). Lateral preference and neuropsychological correlates of schizotypy. *Psychiatry Research, 41,* 115–135.

Klein, D. N., & Miller, G. A. (1993). Depressive personality in nonclinical subjects. *American Journal of Psychiatry, 150,* 1718–1724.

Lewine, R. J. (1981). Sex differences in schizophrenia: Timing or subtype? *Psychological Bulletin, 90,* 432–444.

Meehl, P. E. (1962). Schizotaxia, schizotypy, and schizophrenia. *American Psychologist, 17,* 827–838.

Meehl, P. E. (1964). *Manual for use with checklist of schizotypic signs.* Unpublished manuscript.

Miller, G. A., & Yee, C. M. (1994). Risk for severe psychopathology: Psychometric screening and psychophysiological assessment. In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 5, pp. 1–54). London: Jessica Kingsley.

Mishlove, M., & Chapman, L. J. (1985). Social anhedonia in the prediction of psychosis proneness. *Journal of Abnormal Psychology, 94,* 384–396.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 101,* 259–282.

Nolen-Hoeksema, S. (1990). *Sex differences in depression.* Stanford, CA: Stanford University Press.

Rosenberger, P. H., & Miller, G. A. (1989). Comparing borderline definitions. DSM-III borderline and schizotypal personality disorders. *Journal of Abnormal Psychology, 98,* 161–169.

Rutter, M. (1986). The developmental psychopathology of depression: Issues and perspectives. In M. Rutter & C. E. Izard (Eds.), *Depression in young people* (pp. 3–32). New York: Guilford Press.

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist, 49,* 929–954.

Weissman, M. M., & Klerman, G. L. (1977). Gender and depression. *Trends in Neurosciences, 8,* 416–420.

Weissman, M. M., Leaf, P. J., Holzer, C. E., Myers, J. K., & Tischler, G. L. (1984). The epidemiology of depression: An update on sex differences in rates. *Journal of Affective Disorders, 7,* 179–188.

Yee, C. M. (1995). Implications of the resource allocation model for mood disorders. In G. A. Miller (Ed.), *The behavioral high-risk paradigm in psychopathology* (pp. 271–288). New York: Springer-Verlag.

Received December 21, 1993
Revision received October 3, 1994
Accepted January 13, 1995 ■