Racial/Ethnic Variation in the Reliability of DSM-IV Pathological Gambling Disorder

Renee M. Cunningham-Williams, PhD, MPE, Emily L. Ostmann, MPH,**
*Edward L. Spitznagel, PhD,† and Samantha J. Books, MPE**

Abstract: Racial/ethnic disparities in mental disorders, including pathological gambling disorder (PGD), may be either real or artifacts of how they are conceptualized and measured. We aimed to assess racial/ethnic variation in the reliability of self-reported lifetime PGD determined by meeting ≥5 criteria of the Diagnostic and Statistical Manual of Mental Disorders. Using community advertising, we recruited 15–85-year-old Caucasians (n = 225) and African (American/other minorities (n = 87), who had gambled more than 5 times lifetime), for 2 interviews, held 1 week apart, about gambling and associated behaviors. Results indicate substantial to almost-perfect DSM-IV PGD reliability for Caucasians (κ = 0.82) and African Americans/other minorities (κ = 0.68). Reliability for symptoms and for game-specific disorders was fair to almost perfect (κ = 0.37–0.90). After adjusting results for confounding variables and multiple comparisons, racial/ethnic variation in PGD and game-specific reliability failed to persist. Implications exist for increased attention to screening and prevention efforts critical to reducing racial/ethnic disparities in PGD prevalence.

Key Words: Problem/pathological gambling disorder, racial/ethnic disparities, psychiatric diagnoses, psychometrics.

(*J Nerv Ment Dis* 2007;195: 551–559)

There is growing evidence of the significance of race in determining health outcomes (Williams et al., 2003). Reports from the US Surgeon General (US Department of Health and Human Services, 1999, 2001) and the Institute of Medicine (2002) conclude that African Americans and other racial/ethnic minorities receive disparate health and mental health care at various stages along the continuum of care. This continuum includes point-of-symptom recognition (or prevention) through point-of-treatment or service provision and also at several intervening critical junctures. For example, compared with Caucasians, racial/ethnic minorities, particularly African Americans, disproportionately experience differing psychiatric symptom presentation leading to clinical bias and classification errors of over- and under-diagnosis (e.g., for schizophrenia and certain affective disorders), poorer health and mental health care access, inadequate health and mental health treatment, and increased morbidity and mortality (Blake, 1973; Lawson et al., 1994; Mukherjee et al., 1983; Pope and Lipinski, 1978; Strakowski et al., 1993, 1995, 1996, 2003).

Algeria and McGuire (2003) challenge researchers and psychiatric diagnosticians to rethink the "one size fits all" or universalist framework for applying psychiatric criteria and for understanding the psychiatric symptom-disorder relationship. Such rethinking includes the adoption of a "relativist" approach to understanding mental disorder for racially/ethnically diverse populations. Adoption of such a framework would require asking important questions of the existing diagnostic criteria regarding the appropriateness of their application to culturally diverse populations. Such movement would, for example, entail a purposeful analysis of whether there are differing criteria cut-points or symptom thresholds for Caucasians versus African Americans and whether existing mental and substance use disorder criteria apply equally well to various racial/ethnic groups (Canino et al., 1999; Compton et al., 1996; Horton et al., 2000; Hicks, 2004; Robins and Regier, 1991).

Although health disparity research specific to pathological gambling disorder (PGD) is in its infancy, there is emerging epidemiological evidence suggesting that this disorder is not immune from such racial/ethnic variation in prevalence (Cunningham-Williams et al., 1998, 2005; Petry, 2005; Welte et al., 2002). For example, in a national US sample of residents aged 18 and older (n = 2631), Welte et al. (2004) found that being African American, Hispanic, or Asian (versus being Caucasian) was a significant risk factor for gambling disorder, even after controlling for socioeconomic status. Similarly, an analyses of gambling data included in the National Epidemiologic Survey on Alcohol and Related Conditions of 43,093 household and group quarters residents (Petry et al., 2005) found that being African American and living in the Midwest or West, in addition to other

*George Warren Brown School of Social Work, and †Department of Mathematics, Washington University, St. Louis, Missouri.
Supported by the National Institute on Drug Abuse of the National Institutes of Health Grant, GAMCO Project No. K01 DA 00430 (to R.M.C), and GAPP Project No. R01 DA 015032 (to R.M.C.).
This article was presented at the Society for Social Work and Research Conference, San Antonio, TX, January 14–16, 2006.
Send reprint requests to Renee M. Cunningham-Williams, PhD, MPE, LCSW, George Warren Brown School of Social Work, Washington University, One Brookings Drive, St. Louis, MO 63130. E-mail: williamsr@wustl.edu.

ISSN: 0022-3018/07/19507-0551
DOI: 10.1097/NMD.0b013e318093ed13

sociodemographics (e.g., being divorced/separated/widowed or middle-aged), were associated with increased risk for PGD. Concurring with this finding of increased PGD risk for African Americans are recent community-level studies of household residents. For example, Cunningham-Williams and colleagues recently reported that although only 20% of the 913 study participants were African American, they represented nearly half of those with DSM-IV PGD, indicating their vulnerability to this disorder in the context of increased legal gambling opportunities experienced in the St. Louis Metropolitan area (Cunningham-Williams et al., 2005).

In the absence of a "gold standard" for psychiatric diagnoses, what is most widely accepted are the symptoms codified in the Diagnostic and Statistical Manual of Mental Disorders (DSM), first in its third edition (American Psychiatric Association, 1980) and changing with each version through the latest DSM edition (American Psychiatric Association, 1994). It remains to be seen how the publication of DSM-V, expected in 2010 or later, will reflect PGD, but given PGD's historical criteria modification in each DSM edition [as well as a summary of presentations from a recent DSM-V development committee meeting (http://dsm5.org/conference4.cfm accessed 5/25/06)], it is quite possible that PGD criteria and/or the symptoms used to comprise them will likely change again. Thus, psychometric data on DSM-IV PGD may inform future refinements in PGD criteria.

Yet, to our knowledge, there are no data on the reliability of PGD symptoms reported by African American gamblers individually, or in comparison to Caucasian gamblers. Although there are a few investigations focusing specifically on whether DSM criteria capture the disorder in general, most samples were treatment samples (Lesieur, 1988), samples of casino patrons (Fisher, 2000), pilot studies (Beaudoin and Cox, 1999), or preliminary data (Toce-Gerstein, 2005). A notable exception is a study by Stinchfield (2003), which used a general population sample in Minnesota ($n = 803$) and clients sampled from 6 state-supported Minnesota gambling treatment programs ($n = 259$). That study found that DSM-IV PGD had satisfactory internal consistency reliability and construct and convergent validity. Although these findings are promising evidence that DSM-IV PGD, when operationalized into questions, is a psychometrically sound standard for diagnosis, the sample used was 96% Caucasian. Information on racial/ethnic variation in PGD is critical to evidenced-based practice and research aimed at improving our ability to detect disorder early and accurately, and to ultimately reduce racial/ethnic disparities in mental disorder through appropriately targeted and tailored interventions for culturally diverse populations.

This is also the first report of racial/ethnic variation in the reliability of game-specific disorder. In this study, we developed 12 separate computer-derived algorithms to assess DSM-IV PGD overall, and to assess 11 different categories of game-specific disorder (e.g., slot/fruit machine disorder; lottery disorder, etc.). For game-specific disorder assessment, we applied all DSM-IV criteria to each of the 11 different individual categories of gambling activity engaged in more than 5 times lifetime and then determined the presence or absence of racial/ethnic variation within them. Thus, we aimed to test the following hypotheses: (1) Reliability of DSM-IV PGD will not be significantly affected by racial/ethnic variability; (2) Reliability of game-specific disorders will not be significantly affected by racial/ethnic variability; (3) Diagnostic reliability for PGD and game-specific disorders will be substantial to almost perfect in this sample; and (4) Change in PGD diagnostic status between the Test and Retest interviews will not be significantly affected by race/ethnicity.

METHODS

Between November 2003 and November 2004, using community advertising and telephone screening, we recruited 315 individuals meeting age eligibility (i.e., ages 15–85; mean = 46.8; $SD = 17.0$) and minimal gambling involvement inclusion criteria (i.e., >5 times lifetime). Gamblers participated in 2 telephone interviews occurring about 1 week apart (i.e., 1-week test-retest; mean = 7.4 days; $SD = 3.13$; range = 1–29). A different trained, nonclinician interviewer drawn from a pool of trained interviewers who were blinded to the previous interview data conducted each of the 2 interviews. Interviewers were purposefully rotated between administering the test interview session and the retest interview session for each case, to control potential interviewer and social desirability biases and to broaden the experience of the interviewers in what was required of them for conducting each of the 2 administrations. Furthermore, interviewers were instructed to not discuss or debrief any cases with each other, but to do so with the Principal Investigator or the Study Coordinator so as to not bias another interviewer who may be conducting the retest interview for that particular case. Additionally, to further prevent interviewer bias in upcoming interviews, we used a Scheduler who was not a project interviewer to schedule the interviews and to individually notify each interviewer of which cases they were assigned. Thus, given that we used different, blinded interviewers rotating their interview administration across 2 separate time periods, this study is essentially characterized as a "one-week test-retest study with additional controls for interviewer biases" rather than an "inter-rater reliability study."

We excluded from analysis reported herein those who gambled more than 5 times but not on any one activity ($n = 1$) and those who reported not being a gambler at both interviews ($n = 2$), yielding a final data analysis sample of $n = 312$. All interviews were audiotaped and randomly selected for quality assurance review. Informed consent (and parental consent/minor assent) was obtained before the first interview and participants were paid for their time. All procedures were approved by the Washington University Institutional Review Board (IRB), with federal Certificate of Confidentiality assurances.

Operationalization of DSM-IV PGD

This study used the Computerized-Gambling Assessment Module (C-GAM©; Cunningham-Williams et al., 2003). The C-GAM, a family of interviewer- and self-administered computerized and paper/pen assessments, is the only assessment package to date for screening problem gambling and

diagnosing PGD using computer-derived algorithms according to all PGD criteria-inclusive editions of the DSM, as well as the 10th edition of the International Classification of Diseases (ICD-10) (World Health Organization, 1993). The C-GAM was initially developed in paper/pen format, using 10 focus groups and 108 pretests with funding to the first author (R.C.-W.) from the National Center for Responsible Gaming (Cunningham-Williams et al., 2000). Its computerization and psychometric testing was further supported by 2 grants to the first author from the National Institute on Drug Abuse of the National Institutes of Health, namely, the Gambling and Comorbidity Study (GAMCO) and the Gambling and Personality Profile Study (GAPP).

One of the unique features of the C-GAM is its ability to provide separate PGD diagnoses for 11 different gambling activity categories (e.g., card games, lottery, slot machines, market betting, and games of personal skill such as golf or bowling, etc.). It also queries the context of these gambling and betting activities such as the different venues of play (e.g., on the computer/internet, home, street corner, etc.), social networks for gambling initiation and most recent gambling activities, and alcohol- and drug-associated gambling. For DSM-IV, the C-GAM uses 12 items to operationalize the 10 DSM-IV criteria. Additionally, the C-GAM queries the ages of onset and recency for gambling symptoms, behaviors, and associated problems. It provides additional clinically meaningful data not captured by established diagnostic criteria (e.g., gambling-associated "withdrawal-like" and dissociative symptoms), treatment history, gambling wins and losses, demographic data, and interviewer observations of the gambler and of the interviewing condition. The C-GAM also has several stand-alone components with essentially the same core diagnostic items, but reformatted for different administration types (Cunningham-Williams et al., 2005). For this article, we report PGD from the C-GAM as determined by DSM-IV only.

Test-Retest Interview Sessions

After telephone screening, enrollees were sent University-approved informed consent information, a study description, and 2 sealed respondent booklets corresponding to each of the telephone interview sessions. Each booklet contained a label across the seal with the date and time of the scheduled telephone interview and an instruction to not break the envelope seal until the interview begins. The respondent booklet for the Test Interview Session contained cue cards to aid in C-GAM administration (e.g., definition of gambling and betting: "Now I'd like to ask about your experiences with gambling, betting, and playing games for money or for something else of value. This could be at the casino, on the computer, at the track, on the street, at home, or any other place. Please look at Card 3." The envelope also contained other study-related visual aids (e.g., Likert-scale response categories) to accompany the other psychiatric assessments administered. The interview session began with a verbal telephone consent and any updates to locating information. The test/retest attrition rate was 1.05% ($n = 315/n = 301$). We also collected additional gambling information from the South Oaks Gambling Screen (Lesieur and Blume, 1987); the field's most widely used gambling screener. DSM-IV drug and alcohol abuse/dependence was measured with the computerized drug and alcohol companion modules of the C-GAM (C-GAM-DA; Cunningham-Williams and Books, 2003) and depression was assessed using the National Institute of Mental Health's Center for Epidemiological Studies–Depression Scale (CES-D; Radloff, 1977; Radloff, 1991).

For this study, we also used a mixed methods approach for assessing reasons for inconsistent answers and gamblers' reports of their understanding of interview questions using a modification of the Discrepancy Interview Protocol (DIP)/Debriefing procedure (For more detail see Cottler et al., 1994). For psychometric studies where the same instrument is used more than once (e.g. in a test-retest study), this discrepancy and debriefing protocol allows a systematic exploration of research participants' discrepant responses to predetermined interview items across multiple interview administrations. It also allows researchers to evaluate research participants' understanding of the intent of these predetermined items. A separate, different interviewer is used for each interview administration. Procedures are followed to ensure that the subsequent interviewer is blinded to the data collected in the preceding interview(s) until the time of debriefing. At the terminal interview's conclusion, the interviewer checks if discrepancies actually did occur by reviewing the hard-copy DIP Log from each of the interviews containing the responses for those predetermined items. The interviewer then queries respondents on their estimate of the correct response (i.e., resolution) to that item given their discrepant responses across the interviews. Original responses are never changed, however the resolution response is recorded on the DIP log. Then the interviewer records respondents' answers to what they thought the particular item meant using nondirective probes of commonly used responses.

For this study, we modified the DIP/Debriefing procedure by computerizing it, with the aim of making it more efficient and minimizing potential errors in data collection and discrepancy resolution. For example, the computer program screen included columns that allowed a side-by-side review of all original preloaded responses and a blank data entry field for entering the resolved responses in cases of discrepancy. Furthermore, we debriefed all diagnostic stem questions where discrepancies occurred among all PGD diagnostic stem questions included in the C-GAM. We also included a field for entering verbatim the reasons for discrepant items. Only in those cases where the respondent could not think of a reason, the interviewer followed a list of reasons going down consecutively until the respondent endorsed a reason. The order of items on the list of reasons was randomized across 3 separate lists and these lists were rotated every month during data collection. We then asked the respondents, for those particular discrepant items, what they thought was the intent of the question being asked. For the majority of respondents without discrepancies, we obviously did not conduct debriefing, but we did ask the meaning of certain preselected items to assess respondents' understanding of the item's content. For this article, we present reasons for unreliable responses; however detailed information on the respon-

dents' understanding of the intent of items is beyond the scope of the current report.

Additionally, although also not reported herein, half of the sample was simultaneously enrolled in a parallel NIDA-funded GAPP study where they received an additional, yet semi-structured interview, administered by one of several different clinicians, blinded to previous interview data. The clinician interview randomly occurred 1 week before or 1 week after the 2 structured interviews in the GAMCO study. Clinicians also queried 11 different DSM-IV Axis II disorders using the computerized Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) (First et al., 1995).

Statistical Analysis

Data were analyzed using SAS version 9.1 (SAS Institute, 2004). χ^2 analyses described significant bivariate associations among sociodemographic variables, including race/ethnicity and PGD. We also used the simple κ-statistic (Cohen, 1960) to analyze the proportion of agreement corrected for by chance between the test and retest interviews for Caucasian and African American/other minority gamblers. Individuals in the African American/other category includes those who were either biracial, multiracial or of another racial/ethnic minority and is referred to as African Americans/other owing to their low representation ($n = 15$) in the sample. Weighted κ-statistics assessed reliability with values ranging from a low of $\kappa = -1.0$ (indicating poorest agreement) to a high of $\kappa = 1.0$ (indicating highest agreement). Although there are several ways to classify levels of agreement (Cicchetti et al., 2006), we use the earliest guidelines by Landis and Koch (1977) because of their finer gradations in the strength of agreement levels where κ values ≥ 0.81 are considered almost perfect, 0.61–0.80 are substantial, 0.41–0.60 are moderate, 0.21–0.40 are fair, 0–0.20 are slight, and κ values <0 are considered poor (Landis and Koch, 1977). In further understanding racial/ethnic variation in the reliability of DSM-IV PGD, we adjusted for significant bivariate associations with sociodemographics and for the potential instrumentation effects of being dually enrolled in the aforementioned GAPP psychometric study. Additionally, we used Bonferonni post hoc corrections to statistically adjust the p-values for game-specific disorders potentially attributable to Type I error owing to multiple comparisons.

RESULTS

Sample Description

Table 1 summarizes the sociodemographic characteristics of the data analysis sample ($n = 312$) overall and also categorized by race/ethnicity. The sample was 27.9% minority (primarily African Americans at 23.1%), slightly more female (55.8%), predominately middle-aged adults (mean age = 46.8 years; $SD = 17.0$; range 15–85). They were fairly evenly split across 3 of the 4 marital status categories, with 34.7% being never married, with an average of 1.59 children ($SD = 1.64$). Nearly the entire sample had a high school or higher education (94.3%), yet 35% of the sample was currently unemployed. Annual household incomes were fairly evenly distributed among the 4 levels with lowest (<$25,000) being 28.2% and highest (>$75,000) being 21.3%.

There were several significant associations of race/ethnicity and other sociodemographics. Namely, African Americans were significantly younger (mean = 42.6; $SD = 14.4$ vs. mean = 48.35; $SD = 17.6$) and more often never married (47.7% vs. 29.8%). Compared with Caucasians, African Americans in this sample were more often city dwellers (54.0% vs. 29.3%), less likely to live in their own house or apartment (74.7% vs. 88.4%), and more often poor, with nearly half of them reporting annual household incomes less than $25,000 (48.8% vs. 20.4%).

Change in Prevalence in Gambling Symptom Reporting

Of the 12 symptoms operationalizing the 10 DSM-IV PGD criteria, there was a statistically higher proportion of African Americans (versus Caucasians) at the Test interview endorsing: preoccupation ($\chi^2 = 14.57$; $p = .0001$), chasing losses ($\chi^2 = 6.35$; $p = .01$), inability to control or stop gambling ($\chi^2 = 4.88$; $p = .03$), relying on others for a financial bail-out ($\chi^2 = 4.77$; $p = .03$), having gambling that interfered with their responsibilities ($\chi^2 = 4.82$; $p = .03$), and being involved in illegal behaviors because of gambling ($\chi^2 = 8.93$; $p = .003$).

The majority of the sample was consistent in their symptom reporting between the Test and Retest Interviews. Of the 55 gamblers who changed their gambling symptom reports at the Retest Interview 1 week later (i.e., 18.4% discordance), change was primarily in the direction of a reduction in symptoms. The only remaining variation in symptoms at the Retest interview was the significantly higher proportion of African Americans who endorsed chasing losses ($\chi^2 = 7.99$; $p = .005$), relying on others for a financial bail-out ($\chi^2 = 4.66$; $p = .03$), and having gambling that interfered with their responsibilities at home, work or school ($\chi^2 = 7.61$; $p = .006$), with the additional symptom of giving up or reducing activities to gamble or bet ($\chi^2 = 4.07$; $p = .04$).

Table 2 shows that the total gambling symptom prevalence decreased from the Test and Retest interviews (mean = 2.78; $SD = 3.45$; mean = 2.46; $SD = 3.28$, respectively). African Americans had a significantly higher number of gambling symptoms at the Test interview (mean = 3.74; $SD = 3.64$) compared with Caucasians (mean = 2.42; $SD = 3.32$; $t = 2.87$; $df = 135$; $p = .005$), yet their severity level was more comparable with only slight racial/ethnic variation at the Retest interview.

We were also interested not only in racial/ethnic variation in the *net change* in symptom reporting across groups, but more precisely, variation in how *each individual gambler changed* in their endorsement of each symptom between the Test and Retest interviews. This percent change ranged from a low of 3% (for committing illegal acts) to a high of 14% (for gambling to escape problems or a dysphoric mood), all in the direction of a reduction in symptoms by the Retest interview. Compared with Caucasians, African Americans had the largest percent reduction in their individual symptom

TABLE 1. Sociodemographic Characteristics by Race/Ethnicity at Test Interview (n = 312)

Characteristics	*n* (%)* Total Sample (n = 312)	Caucasian (n = 225)	African American (n = 87)	Statistic (df; p)
Gender				χ^2 = 1.30 (1; 0.2547)
Male	138 (44.2)	104 (46.2)	34 (39.1)	
Female	174 (55.8)	121 (53.8)	53 (60.9)	
Age				t = 2.96 (191; 0.0035)
Mean (*SD*)	46.8 (17.0)	48.4 (17.6)	42.6 (14.4)	
Median age (range)	48.0 (15–85)	51.0 (15–85)	44.0 (18–74)	
Highest education attained				χ^2 = 5.58 (2; 0.0615)
Less than high school	18 (5.8)	10 (4.4)	8 (9.2)	
High school/GED	154 (49.4)	106 (47.1)	48 (55.2)	
More than high school/GED	140 (44.9)	109 (48.4)	31 (35.6)	
Current employment				χ^2 = 0.03 (1; 0.8726)
Employed	203 (65.1)	147 (65.4)	56 (64.4)	
Not employed	109 (34.9)	78 (34.7)	31 (28.4)	
Current living arrangements				χ^2 = 9.09 (1; 0.0026)
Own house/apartment	264 (84.6)	199 (88.4)	65 (74.7)	
Someone else's house/apartment/other	48 (15.4)	26 (11.6)	22 (25.3)	
Current residence				χ^2 = 16.56 (1; <0.0001)
City dweller	113 (36.2)	66 (29.3)	47 (54.0)	
Suburban/rural dweller	199 (63.8)	159 (70.7)	40 (46.0)	
Household income (past 12 mo)				χ^2 = 26.71 (3; <0.0001)
Less than $25,000	86 (28.2)	45 (20.4)	41 (48.8)	
$25,000–49,999	93 (30.5)	70 (31.7)	23 (27.4)	
$50,000–$74,999	61 (20.0)	50 (22.6)	11 (13.1)	
$75,000 or higher	65 (21.3)	56 (25.3)	9 (10.7)	
Current marital status				χ^2 = 19.44 (3; 0.0002)
Married	113 (36.3)	98 (43.6)	15 (17.4)	
Widowed	20 (6.4)	12 (5.3)	8 (9.3)	
Never married	108 (34.7)	67 (29.8)	41 (47.7)	
Separated/divorced	70 (22.5)	48 (21.3)	22 (25.6)	
Number of children				t = −1.31 (155; 0.1938)
Mean (*SD*)	1.59 (1.64)	1.51(1.63)	1.78 (1.65)	
Median (range)	1.0 (0–9)	1.0 (0–9)	2.0 (0–5)	
Times married				t = 2.50 (163; 0.0133)
Mean (*SD*)	1.0 (0.9)	1.0 (0.9)	0.76 (0.87)	
Median (range)	1.0 (0–5)	1.0 (0–5 times)	1.0 (0–3 times)	

Three people were excluded because of gambling fewer than 6 times on any one activity and/or being a nonplayer at the test interview or the retest interview; the African American category includes gamblers who self-identify as biracial, multiracial, or other racial/ethnic minority (n = 15; 4.8%); employed = both full part-time workers as well as students; not employed = unemployed due to termination, layoffs, disability, or retirement. Someone else's home/apartment or other category includes those living in a hotel, shelter, or halfway house (n = 2) as well as those living on the street or who consider themselves homeless (n = 8).

*Values indicate n (%) values unless indicated.

reports, specifically for preoccupation (χ^2 = 10.6; p = .001), gambling to escape problems (χ^2 = 5.85; p = .02), inability to control or stop gambling (χ^2 = 6.81; p = .009) and lying/deceit about gambling behavior (χ^2 = 3.77; p = .05).

Given that there were no diagnostically indeterminate cases because of missing data on specific criteria at both interviews, we classified the sample into 3 mutually exclusive groups based on the number of positively endorsed DSM-IV PGD criteria at the Test interview: those meeting criteria (PGD: ≥5 criteria, 27.0%), subsyndromal cases (Sub-Threshold: 1–4 criteria; 33.9%) and those gamblers without problems (Non-Problem; 0 criteria; 39.1%). Table 2 also shows that the reduction in symptoms primarily occurred among gamblers with PGD who were largely represented among the Subthreshold gamblers at the Retest interview.

Reliability in Gambling Symptom Reporting

Table 3 shows the proportion of agreement observed and the findings supporting reliability in reports of DSM-IV gambling symptoms in the substantial to almost-perfect range for the total sample (κ = 0.61–0.88) and moderate to almost-perfect reliability specifically for Caucasians (κ = 0.56–0.90) and African Americans (κ = 0.56–0.84). Statistically significant racial/ethnic variation was also present for 3 of the

TABLE 2. Change in Symptoms and Diagnostic Classification from Test to Retest Interviews by Race/Ethnicity ($n = 299$)

DSM-IV Diagnostic Classification and Symptom Counts	*n* (%)* Total Sample ($n = 312$)	White ($n = 225$)	African American† ($n = 87$)	Statistic (*df*; *p*)
Test interview				$\chi^2 = 11.65$ (2; 0.0029)
PGD	71 (23.7)	46 (21.2)	25 (30.5)	
Subthreshold	112 (37.5)	74 (34.1)	38 (46.3)	
Nonproblem	116 (38.8)	97 (44.7)	19 (23.2)	
Indeterminate	0 (0.0)	0 (0.0)	0 (0.0)	
Retest interview				$\chi^2 = 8.40$ (3; 0.0384)
PGD	63 (21.1)	42 (19.4)	21 (25.6)	
Subthreshold	121 (40.5)	81 (37.3)	40 (48.8)	
Nonproblem	112 (37.5)	91(41.9)	21 (25.6)	
Indeterminate	3 (1.0)	3 (1.4)	0 (0.0)	
Mean symptom count at test interview (*SD*)	2.78 (3.45)	2.42 (3.32)	3.74 (3.64)	$t = -2.87$ (135; 0.0047)
Mean symptom count at retest interview (*SD*)	2.46 (3.28)	2.21 (3.18)	3.10 (3.48)	$t = -2.00$ (135; 0.0474)

*Values indicate *n* (%) values unless indicated.
†The African American category includes gamblers who self-identify as biracial, multiracial, or other racial/ethnic minority ($n = 15$; 4.8%).

TABLE 3. Proportion of Agreement Observed (P_o) and Reliability (κ) of DSM-IV Gambling Symptoms by Race/Ethnicity ($n = 299$)

DSM-IV Gambling Symptoms	Total Sample P_o	Total Sample κ (95% CI)	Among Caucasians P_o	Among Caucasians κ (95% CI)	Among African Americans* P_o	Among African Americans* κ (95% CI)
Financial bail-out	0.97	0.88 (0.82–0.96)	0.98	0.90 (0.82–0.99)	0.94	0.84 (0.70–0.98)
Relationship problems	0.95	0.82 (0.74–0.91)	0.96	0.85 (0.76–0.95)	0.92	0.76 (0.59–0.93)
Lying or hiding gambling	0.92	0.80 (0.73–0.88)	0.94	0.85 (0.77–0.93)	0.87	0.70 (0.53–0.86)
Illegal acts to get gambling money	0.98	0.76 (0.59–0.93)	0.99	0.79 (0.56–1.00)	0.95	0.72 (0.47–0.98)
Preoccupation	0.91	0.75 (0.66–0.84)	0.95	0.83 (0.74–0.93)	0.82	0.60 (0.42–0.77)
Cannot control or quit or cut down	0.90	0.74 (0.66–0.83)	0.93	0.81 (0.72–0.90)	0.82	0.59 (0.41–0.77)
Give up or reduce activities	0.93	0.74 (0.64–0.85)	0.95	0.78 (0.66–0.91)	0.88	0.66 (0.46–0.85)
Gambling to escape	0.87	0.73 (0.65–0.81)	0.91	0.79 (0.71–0.88)	0.78	0.56 (0.38–0.74)
Chasing losses	0.88	0.72 (0.64–0.81)	0.89	0.71 (0.60–0.82)	0.87	0.73 (0.57–0.88)
Interfere with responsibilities at home, work, school	0.93	0.72 (0.61–0.83)	0.93	0.70 (0.56–0.85)	0.90	0.74 (0.56–0.91)
Restless or irritable	0.90	0.70 (0.60–0.80)	0.91	0.71 (0.59–0.83)	0.88	0.68 (0.50–0.86)
Tolerance	0.85	0.61 (0.50–0.72)	0.87	0.56 (0.42–0.71)	0.89	0.70 (0.52–0.88)

*The African American category includes gamblers who self-identify as biracial, multiracial, or other racial/ethnic minority ($n = 15$; 4.8%).

12 gambling symptoms, namely preoccupation, gambling to escape problems, and inability to control or quit gambling, wherein African Americans had moderate reliability on these symptoms ($\kappa = 0.56–0.60$).

Racial/Ethnic Variation in Diagnostic Reliability

When stratified by race/ethnicity, there were significant racial/ethnic differences in DSM-IV PGD at both the test interview ($\chi^2 = 11.7$; $p = .003$) and the retest interview ($\chi^2 = 8.40$; $p = .038$), with substantial to almost-perfect reliability for Caucasians ($\kappa = 0.82$; range = 0.76–0.89; $SE = 0.03$) and African Americans ($\kappa = 0.64$; range = 0.49–0.80; $SE = 0.08$) (Table 4). One of the additional unique features of the C-GAM is its ability to provide an individual diagnosis for each of 11 specific categories of gambling activity. For each category, game-specific disorder reliability was moderate to substantial ranging from $\kappa = 0.51$ (Lottery disorder) to $\kappa = 0.77$ (Card game disorder). The only significant racial/ethnic variation in game-specific disorder reliability was for Slot/Fruit Machine disorder ($\kappa = 0.75$), where reliability was almost perfect for Caucasians ($\kappa = 0.81$) and moderate for African Americans ($\kappa = 0.57$).

We further examined change in reliability where racial/ethnic variation was noted after separately adjusting for potential confounding variables. Specifically, for determining reliability of PGD overall, we adjusted for each separate sociodemographic variable that significantly distinguished Caucasians and African Americans in this sample. There still was a minor race/ethnicity effect after controlling for house-

TABLE 4. Proportion of Agreement Observed (P_o) and Reliability (κ) of DSM-IV PGD and Game-Specific Disorders* by Race/Ethnicity (n = 299)

	Total Sample		Among Caucasians		Among African Americans†	
Gambling Disorder	P_o	κ (95% CI)	P_o	κ (95% CI)	P_o	κ (95% CI)
DSM-IV PGD	0.84	0.79 (0.73–0.84)	0.85	0.82 (0.76–0.88)	0.58	0.68 (0.56–0.81)
Card game (non-VDT) disorder	0.86	0.77 (0.69–0.85)	0.88	0.78 (0.69–0.88)	0.79	0.73 (0.57–0.88)
Slots/fruit machine disorder	0.82	0.75 (0.68–0.82)	0.86	0.81 (0.74–0.89)	0.68	0.57 (0.39–0.74)
Craps/dice game disorder	0.81	0.72 (0.59–0.85)	0.88	0.76 (0.55–0.97)	0.71	0.62 (0.41–0.83)
Parimutuel disorder	0.89	0.72 (0.57–0.88)	0.91	0.76 (0.60–0.92)	0.81	0.58 (0.16–1.0)
Sports disorder	0.85	0.69 (0.54–0.84)	0.85	0.68 (0.48–0.87)	0.86	0.71 (0.46–0.95)
Other casino game disorder	0.85	0.69 (0.54–0.84)	0.85	0.59 (0.36–0.82)	0.85	0.78 (0.57–0.98)
Bingo disorder	0.89	0.62 (0.44–0.81)	0.92	0.70 (0.50–0.91)	0.82	0.44 (0.13–0.75)
Video poker/VDT disorder	0.78	0.62 (0.50–0.75)	0.81	0.65 (0.51–0.79)	0.66	0.50 (0.24–0.76)
Other games/own personal skill game disorder	0.86	0.58 (0.36–0.79)	0.88	0.64 (0.37–0.90)	0.80	0.45 (0.10–0.79)
Lottery disorder	0.79	0.51 (0.39–0.63)	0.85	0.58 (0.42–0.74)	0.63	0.37 (0.17–0.56)

*Reliability statistics were not able to be calculated for Market Betting Disorder due to small cell sizes (n = 11).
†The African American category includes gamblers who self-identify as biracial, multiracial, or other racial/ethnic minority (n = 15; 4.8%).

hold income (χ^2 = 9.49; p = .049) and age (χ^2 = 15.06; p = .005). Specifically across household income and race/ethnicity categories, African Americans with the lowest household annual income (i.e., <$25,000) were the least reliable (κ = 0.59). With respect to age and race/ethnicity, African Americans who were older adults (i.e., ≥60 years) were the least reliable (κ = 0.05), followed by those who were middle-aged (i.e., 45–59 years; κ = 0.45). Yet, after post hoc correction, these results were no longer significant (z-score = 1.84; p = .065).

Slot/Fruit Machine disorder was the only game-specific disorder among the 11 different categories of games showing racial/ethnicity disparity in reliability; with African American gamblers having significantly lower reliability than Caucasian gamblers. In fact, racial/ethnic variation persisted even after controlling for urban or suburban residence (χ^2 = 7.34; df = 2; p = .03), age (χ^2 = 25.2; df = 4; p = .0001), times married (χ^2 = 16.7; df = 3; p = .001), and whether participants were dually enrolled in the parallel GAPP validity study and if enrolled, whether that clinical validation interview came before or after the test-retest reliability interviews (χ^2 = 9.79; df = 3; p = .02). Similar to the analysis of PGD reliability overall, game-specific reliability varied by race/ethnicity. Specifically, across race/ethnicity and age categories, African American gamblers who were middle-aged (45–59 years) and older (≥60 years) were equally reliable (κ = 0.18 for both age categories). In terms of multiple marriages, African Americans who were married ≥2 times were the least reliable (κ = 0.14). Finally, across race/ethnicity and study enrollment status, African Americans who were enrolled in the parallel GAPP study and had their reliability interviews randomly occurring before the clinical validation interview were the least reliable (κ = 0.04). Similar to reliability in the overall PGD diagnosis, after Bonferroni post hoc correction, racial/ethnic variation in the reliability of Slot/Fruit Machine disorder failed to persist (z-score = 2.54; p = .011 × 10 = .110).

Reasons for Unreliability

We used a modified DIP /Debriefing procedure to tease out self-reported reasons why answers may be different when asked at a subsequent interview. Only gamblers with discrepant answers on at least one of the PGD diagnostic stem questions (n = 136 discrepancies) were asked about discrepant answers from the interviews. The most prevalent reason given for discrepancies in the DSM-IV PGD symptoms was having misunderstood the question asked (n = 92), followed by not paying attention to the question when asked (n = 17), and attributing discrepancies to errors in interviewer coding or to the interviewer's own misunderstanding of responses given (n = 7). The remaining gamblers either did not know why they gave an inconsistent answer (n = 6) or offered another reason for the discrepancy (n = 14). Self-reported reasons for discrepancies did not significantly differ by race/ethnicity, other sociodemographics, or by PGD diagnostic classification.

DISCUSSION

DSM-IV PGD, as operationalized by the C-GAM has substantial test-retest reliability overall. Furthermore, the diagnosis applies equally well for Caucasians and African Americans, with reliability falling in the moderate to almost-perfect range, with moderate to almost-perfect reliability at the symptom level as well for both groups of gamblers. When inconsistent responses to diagnostic stem questions were evident between time 1 and time 2, it was infrequent (<20%), with inconsistencies not attributable to racial/ethnic variation.

All but one hypothesis was supported in that reliability for game-specific disorders for African Americans, while fair to moderate, was less than substantial for 6 of 10 game-specific disorders assessed. However, our results showed that this difference in reliability was not owing to racial/ethnic variation. In fact, when modest racial/ethnic variation appeared to persist for slot/fruit machine disorder, as well as for

the overall PGD disorder, after we controlled for significant confounding sociodemographic variables and made post hoc corrections, these differences were no longer apparent. Moreover, gamblers primarily attributed their inconsistent responses to misunderstanding items, and this attribution also failed to be distinguished by race/ethnicity.

There are several important limitations to consider. First, this is a convenience sample of gamblers responding to community-advertising for those gambling more than 5 times at any point in their lifetimes. Thus, these findings may not be generalizable to the general population of gamblers, treatment-seeking gamblers, and to those currently experiencing gambling problems. Also, the African Americans in this sample had low educational attainment and were predominately very poor. Therefore, our African American sample may not be representative of African American gamblers or gamblers of other races/ethnicities. As this is the first investigation examining racial/ethnic disparity in PGD diagnostic reporting, these findings need to be replicated using a more representative sample of gamblers with racial/ethnic diversity.

It is important to coordinate these data with additional psychometric data collected in the GAMCO and GAPP protocols and replicate these findings through the future work of others. Moreover, there are likely additional confounding variables that may help shed light on the possibility of racial/ethnic variation in PGD reliability. Furthermore, future research may need to examine PGD as a multidimensional (rather than categorical) construct assessing whether there are different cut-points or profiles of gamblers by race/ethnicity.

CONCLUSIONS

Despite these limitations, DSM-IV PGD as operationalized by the C-GAM may be useful as it provides important information on the symptoms that comprise current criteria, does so for specific gambling categories, and generally performs equally well for Caucasians and African Americans. Although a game-specific diagnosis is not yet officially recognized in the current DSM nosology, the C-GAM data may be informative in making inroads into this uncharted area in mental health research. DSM-IV PGD as operationalized by the C-GAM may also be a useful tool in the clinician's armamentarium aimed at reducing racial/ethnic disparities in the prevalence of PGD.

ACKNOWLEDGMENTS

The authors acknowledge the GAMCO and GAPP staff and participants; Julie M. Kapp, M.P.H., Ph.D. for conference abstract data analysis; the Waterman Research Group, St. Louis, Missouri for technical consultation; Center for Mental Health Services Research, George Warren Brown School of Social Work (5P30 MH068579) for technical and administrative assistance, Karen L. Dodson at the Academic Publishing Service, Washington University for editing services, Catharine Mennes, MSW, Ph.D. for helpful comments on an earlier draft; and NIDA Career Development Award preceptors through various stages of the GAMCO project: Linda Cottler, Ph.D, Edward L. Spitznagel, Ph.D., Wilson M. Compton, M.D., M.P.E.

REFERENCES

Algeria M, McGuire T (2003) Rethinking a universal framework in the psychiatric symptom-disorder relationship. *J Health Soc Behav*. 44:257–274.

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th ed). Washington, DC: American Psychiatric Association.

Beaudoin CM, Cox BJ (1999) Characteristics of problem gambling in a Canadian context: A preliminary study using a DSM-IV-based questionnaire. *Can J Psychiatry*. 44:483–487.

Blake W (1973) The influence of race on diagnosis. *Smith Coll Stud Soc Work*. 43:184–192.

Canino G, Bravo M, Ramirez R, Febo VE, Rubio-Stipec M, Fernandez RL, Hasin D (1999) The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): Reliability and concordance with clinical diagnoses in a Hispanic population. *J Stud Alcohol*. 60:790–799.

Cicchetti D, Bronen R, Spencer S, Haut S, Berg A, Oliver P, Tyrer P (2006) Rating scales, scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. *J Nerv Ment Dis*. 194:557–564.

Cohen JA (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 20:37–46.

Compton WM, Cottler LB, Dorsey KB, Spitznagel EL, Mager DE (1996) Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. *Drug Alcohol Depend*. 41:179–187.

Cottler LB, Compton WM, Brown L, Shell A, Keating S, Shillington A, et al. (1994) Discrepancy Interview Protocol: A method for evaluating and interpreting discordant survey responses. *Int J Methods Psychiatr Res*. 4:173–182.

Cunningham-Williams RM, Books SJ (2003) *Computerized-Gambling Assessment Module-Drug and Alcohol (C-GAM-DA©)*. St. Louis (MO): Washington University.

Cunningham-Williams RM, Cottler LB, Compton WM (1998) Taking chances: Problem gamblers and mental health disorders—Results from the St. Louis Epidemiologic Catchment Area (ECA) Study. *Am J Public Health*. 88:1093–1096.

Cunningham-Williams RM, Cottler LB, Compton WM, Books SB (2003) *Computerized Gambling Assessment Module (C-GAM©)*. St. Louis (MO): Washington University.

Cunningham-Williams RM, Cottler LB, Mangiaforte S (2000) The development of the Gambling Assessment Module (GAM-IV), in *11th International Conference on Gambling and Risk-Taking*, Las Vegas, NV, June 12–16, 2000.

Cunningham-Williams RM, Grucza R, Cottler LB, Womack SB, Books SJ, Przybeck T, et al. (2005) Prevalence of pathological gambling among St. Louis area household residents. *J Psychiatr Res*. 39:377–390.

First MB, Spitzer RL, Gibbon M, Williams JBW, Davies M, Borus J, et al. (1995) The structured clinical interview for DSM-III-R Personality Disorders (SCID-II). Part II: Multi-site test-retest reliability study. *J Personal Disord*. 9:92–104.

Fisher S (2000) Measuring the prevalence of sector-specific problem gambling: A study of casino patrons. *J Gambl Stud*. 16:25–51.

Hicks JW (2004) Ethnicity, race and forensic psychiatry: Are we color-blind? *J Am Acad Psychiatry Law*. 32:21–33.

Horton J, Compton W, Cottler L (2000) Reliability of substance use disorder diagnoses among African-Americans and Caucasians. *Drug Alcohol Depend*. 57:203–209.

Institute of Medicine (2002) *Unequal Treatment: Confronting Racial Ethnic Disparities in Health Care*. Washington, DC: National Academy Press.

Keck PE Jr, Arnold LM, Collins J, Wilson RM, Fleck DE, Corey KB, et al. (2003) Ethnicity and diagnosis in patients with affective disorders. *J Clin Psychiatry*. 39:157–162.

Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics*. 33:159–174.

Lawson WB, Helper N, Holladay J, Cuffel B (1994) Race as a factor in inpatient and outpatient admissions and diagnosis. *Hosp Community Psychiatry*. 45:72–74.

Lesieur H (1988) Altering the DSM-III criteria for pathological gambling. *J Gambl Stud*. 4:38–47.

Lesieur HR, Blume SB (1987) The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. *Am J Psychiatry*. 144:1184–1188.

Mukherjee S, Shukla S, Woodle J, Rosen AM, Olarte S (1983) Misdiagnosis of schizophrenia in bipolar patients: A multi-ethnic comparison. *Am J Psychiatry*. 140:1571–1574.

Petry NM (2005) *Pathological Gambling: Etiology, Comorbidity and Treatment*. Washington, DC: American Psychological Association.

Petry NM, Stintson FS, Grant BF (2005) Comorbidity of DSM-IV pathological gambling and psychiatric disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry*. 66:564–574.

Pope HG, Lipinski JF (1978) Diagnosis in schizophrenia and manic-depressive illness. *Arch Gen Psychiatry*. 35:811–828.

Radloff LS (1977) The CES-D Scale: A self-report depression scale for research in the general population. *Appl Psychological Measurement*. 1:385–401.

Radloff LS (1991) The use of the Center for Epidemiologic Studies Depression Scale in adolescents and young adults. *J Youth Adolesc*. 20:149–166.

Robins LN, Regier D (1991) Psychiatric Disorders in America: The Epidemiologic Catchment Area Study. New York: The Free Press.

SAS Institute (2004) *Statistical Analyses System (SAS®) 9. 1 User's Guide*. Cary, NC: SAS Institute.

Stinchfield R (2003) Reliability, validity and classification accuracy of a measure of DSM-IV diagnostic criteria for pathological gambling. *Am J Psychiatry*. 160:180–182.

Strakowski SM, Keck PE Jr, Arnold LM, Collins J, Wilson RM, Fleck DE, et al. (2003) Ethnicity and diagnosis in patients with affective disorders. *J Clin Psychiatry*. 64:747–754.

Strakowski SN, Lonczak HS, Sax KW, West SA, Crist A, Mehta R, et al. (1995) The effects of race on diagnosis and disposition from a psychiatric emergency service. *J Clin Psychiatry*. 56:101–107.

Strakowski SN, McElroy SL, Keck PE, West SA (1996) Racial influence on diagnosis in psychotic mania. *J Affect Disord*. 39:157–162.

Strakowski SN, Shelton RC, Kolbrener ML (1993) The effects of race and comorbidity on clinical diagnosis in patients with psychosis. *J Clin Psychiatry*. 35:1368–1377.

Toce-Gerstein M (2005) Using statistics to explore the DSM-IV criteria for pathological gambling, in *Proceedings of the 19th annual conference on prevention research and treatment of problem gambling.* Session IV: Towards the DSM-V. New Orleans, LA: National Council on Problem Gambling, Washington DC, June 23–25, 2005.

US Department of Health and Human Services (1999) *Mental Health: A Report of the Surgeon General*. Rockville (MD): Public Health Service, Office of the Surgeon General.

U.S. Department of Health and Human Services (2001) *Mental health: Culture, race and ethnicity: A supplement to mental health: A report of the surgeon general*. Rockville, MD: Public Health Service, Office of the Surgeon General.

Welte JW, Barnes GM, Wieczorek WF, Tidwell MC (2004) Gambling participation and pathology in the United States: A socio-demographic analysis using classification trees. *Addict Behav*. 29:983–989.

Welte JW, Barnes GM, Wieczorek WF, Tidwell MC, Parker JC (2002) Gambling participation in the US—Results from a national survey. *J Gambl Stud*. 18:313–337.

Williams DR, Neighbors HW, Jackson JS (2003) Racial/ethnic discrimination and health: Findings from community studies. *Am J Public Health*. 93:200–208.

World Health Organization (1993) *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva, Switzerland: World Health Organization.