BRIEF REPORT

The Convergence of Personality Disorder Diagnoses Across Different Methods Among Monolingual (Spanish-Speaking Only) Hispanic Patients in Substance Use Treatment

Douglas B. Samuel
Purdue University

Luis M. Añez, Manuel Paris, and Carlos M. Grilo
Yale School of Medicine

Methods for diagnosing personality disorders (PDs) within clinical settings typically diverge from those used in treatment research. Treatment groups in research studies are routinely diagnosed using semistructured interviews or self-report questionnaires, yet these methods show poor agreement with clinical diagnoses recorded in medical charts or assigned by treating clinicians, reducing the potential for evidence-based practice. Furthermore, existing research has been limited by focusing on primarily White and English-speaking participants. Our study extended prior research by comparing 4 independent methods of PD diagnosis, including self-report questionnaire, semistructured interview, chart diagnoses, and ratings by treating clinicians, within a clinical series of 130 monolingual (Spanish only) Hispanic persons (69% male; *M* age 37.4), in treatment for substance use. The authors examined the convergence of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*) PD diagnoses across these methods. PD diagnoses appeared infrequently within medical charts but were diagnosed at higher levels by independent treating clinicians, self-report questionnaires, and semistructured interviews. Nonetheless, diagnostic concordance between clinical diagnoses and the other methods were poor ($\kappa <$ .20). Convergence of PD diagnoses across diagnostic methods for Spanish-speaking Hispanic persons are comparable to other groups allaying concerns about cross-cultural application of PD diagnoses. Additionally, the results of this study echo previous research in suggesting that clinicians' PD diagnoses overlap little with self-report questionnaires or semistructured diagnostic interviews and suggest that PDs are underdiagnosed using standard diagnostic approaches. Implications for the clinical application of empirically supported research are discussed.

Keywords: personality disorder, Spanish, Hispanic, agreement, substance use

Although personality disorders (PDs) have great clinical import because of their public health burden and negative impact on psychosocial functioning, their diagnosis has historically been fraught with difficulties (Perry, 1992). Within clinical practice, PD diagnoses are primarily assigned by treating therapists based on their clinical interactions and unstructured interviews (Zimmerman, 1994). These diagnoses are often unstructured in the sense that they are not derived from any systematic assessment of the PD criteria. Within research settings, however, PD diagnoses are typically generated based on self-report questionnaires, semistructured diagnostic interviews, or both that comprehensively assess

each criterion for all PDs (Widiger & Samuel, 2005). Unfortunately, the sparse research performed to date comparing different diagnostic methods for assessing PD has generally found weak agreement (Samuel et al., 2013).

Relatively little research has compared different diagnostic methods for PD, and most existing studies addressed narrow questions, such as the concordance between two methods. These studies have, for example, reported poor agreement between self-report questionnaires and clinician-generated categorical PD diagnoses (e.g., *Mdn K* = .08; Hyler, Rieder, Williams, & Spitzer, 1989). Agreement was only moderately better when considered dimensionally with correlations for individual PD diagnoses ranging from .05 (Chick, Sheaffer, Goggin, & Sison, 1993) to .36 (Klein et al., 1993). Further, routine discordance between clinician-generated PD diagnoses and those from semistructured diagnostic interviews (Samuel et al., 2013) raises important questions about whether practicing clinicians can have confidence that the diagnoses they generate will inform treatment selection and planning. For instance, although research has identified effective treatments for specific PDs, such as borderline (BPD), the patient groups in those studies are diagnosed using semistructured diag-

---

nostic interviews. Thus, their applicability to clinical practice might be questioned. Additional research that compares the convergence across these sources within diverse populations is crucial to further address these issues.

Three studies have compared clinician-assigned diagnoses to both semistructured interview and self-reported diagnoses (Samuel et al., 2013; Samuel & Widiger, 2010; Tenney, Schotte, Denys, van Megen, & Westenberg, 2003). Further complicating these comparisons is that even the method of collecting clinicians' diagnoses has varied across studies. Whereas, Tenney et al. (2003) considered the categorical PD diagnoses that were recorded in the clients' clinical charts, Samuel et al. (2013) used the Personality Assessment Form (PAF; Shea et al., 1990) which collects dimensional ratings of each PD on a 6-point Likert-type scale. No previous study has even compared clinicians' chart diagnoses with clinicians' unstructured ratings of PDs, let alone compared these two methods with diagnoses assigned by semistructured interviews and self-report questionnaires.

A second important limitation of the existing literature on diagnostic convergence across methods is that the clinician and client participants have been overwhelmingly White and largely English speaking. Most of the studies have been conducted in the United States, but even the others have been confined to predominantly White, northern European countries. Thus, it is important to investigate patterns of diagnosing PDs in different cultures, as research suggests that ethnicity–race factors influence diagnostic (Chavira et al., 2003) and treatment (Bender et al., 2007) practices for PDs.

Hispanics are a particularly relevant and understudied group in the field of PDs. Hispanic persons represent the largest (16% of the U.S. population) and fastest growing minority group in the United States (Ennis, Rios-Vargas, & Albert, 2011). Investigating PD diagnostic practices within Hispanic groups is particularly relevant given well-documented differences from White groups in values and interpersonal factors that can potentially impact on the assessment of personality-related constructs (Añez, Paris, Bedregal, Davidson, & Grilo, 2005). For example, Hispanic–Latin groups are more collectivistic than non-Hispanic Anglo groups and have a heightened need for harmonious relationships that might drive differences in dependent and antisocial PDs (Cuellar & Paniagua, 2000). Further, Hispanic individuals tend to place a greater emphasis on the present, rather than the future, which may produce differences on diagnoses such as obsessive–compulsive PD as well as those that have components of impulsivity (e.g., BPD). Nonetheless, we are unaware of any research that has previously examined the rates of cross-method diagnostic convergence within Hispanic populations.

A unique research need for this large minority group, is for research conducted with Spanish with Spanish-language instrumentation (Grilo, Añez, & McGlashan, 2003). In the United States, many Hispanic individuals do not speak English, including estimates that roughly 50% of Hispanic persons do not speak English well (U.S. Bureau of the Census, 1995). Even among those who do speak English, Spanish remains the preferred language for sensitive topics—which are required for appropriate assessment and diagnosis (Altarriba & Santiago-Rivera, 1994).

This report addresses these limitations by examining the diagnostic convergence of PD diagnoses assigned by four different perspectives: self-report questionnaire, semistructured interview,

clinical chart diagnoses, and ratings provided by the treating clinicians. These data are significant because they are from a sample of monolingual (Spanish-speaking only) Hispanics, but also because they will provide the most comprehensive evaluation to date of the convergence between clinicians' PD diagnoses and those assigned by other research-based methods. Furthermore, the data are from within a clinical population, substance users, who are known to have high rates of PDs (Trull, Jahng, Tomko, Wood, & Sher, 2010).

Method

Participants

Participants were drawn from a community-based outpatient program located in an urban area in the northeastern United States. This program, based within a larger community mental health center, serves monolingual (Spanish-speaking only) Hispanic adults. The specialty focus of this program is on intensive aftercare services for individuals with a substance use history and required abstinence during treatment. This patient group has been described previously (Ansell et al., 2010) and is composed of a series of 130 monolingual Hispanic adult patients assigned to a particular treatment team over one year. Patients were assigned to this team based on caseflow within the larger clinic and not on the basis of any clinical or demographic variables. All 130 participants had lifetime diagnoses of alcohol use disorder, and a majority also had additional lifetime substance use disorder diagnoses, assigned by clinical interview. Medical records indicated that all participants were abstinent for at least 60 days prior to their participation. Of the 130 participants, 90 (69%) were men, and 40 (31%) were women; the mean age was 37.4 years (SD = 10.5), and 76 (58%) were married. The majority (66%) were born in Puerto Rico and had been in the United States for 12.2 years (SD = 9.3) on average.

Procedures

All procedures were approved by the institutional review board and were described, and written informed consent was obtained, in Spanish, from all participants. Participants were compensated $50 for their time for completing the research assessments. This study included four types of assessment methods for PD diagnoses gathered completely independently from each other using: (a) a semistructured diagnostic interview administered by research clinicians; (b) a self-report questionnaire completed by clients; (c) a structured rating form completed by the treating clinicians; and (d) medical-record chart diagnoses, which were determined by an intake clinician and then confirmed–modified by the team psychiatrist following a rounds discussion.

Participants first completed the self-report Spanish version of the Personality Diagnostic Questionnaire—4 (PDQ–4; Hyler, 1994). Immediately following the completion of the S–PDQ–4, the Spanish-Language Version of the Diagnostic Interview for *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) Personality Disorders (DIPD–IV; Zanarini, Frankenburg, Sickel, & Yong, 1996) was administered by experienced, bilingual Hispanic, doctoral-level research clinicians who were blind to the S–PDQ–4 results. The research clinicians received intensive training in PD

assessment and the specific use of the DIPD–IV by the senior author. These interviewers were completely independent of, and had no overlap with, the bilingual treating clinicians who provided diagnostic ratings of the participant via the Personality Assessment Form (PAF; Shea et al., 1990). The PAF was completed after the DIPD–IV had been administered but without access to the DIPD–IV findings. Finally, chart diagnoses were obtained from medical records by research personnel.

Measures

Spanish Diagnostic Interview for Personality Disorders (S–DIPD–IV; Grilo, Añez, & McGlashan, 2003). Like the original version (Zanarini et al., 1996), the S–DIPD–IV is a semistructured diagnostic interview that assesses for all *DSM–IV* personality disorders and criteria. The S–DIPD–IV required that criteria must be present and pervasive for at least 2 years and be characteristic of the person during adulthood. A score of 2 is given if the criterion is present and clinically significant; 1 if present, but of uncertain clinical significance; and 0 if not present. The Spanish-language version of the DIPD–IV was developed through a rigorous established process of translation and back-translation (Arnold & Matus, 2000) by doctoral-level research psychologists who were bilingual in English and Spanish and trained by the DIPD–IV's developer and who received further training by the senior author. This process and initial analysis of the S–DIPD–IV are described elsewhere (Grilo et al., 2003). Interrater reliability was evaluated using pairs of independent ratings for 27 randomly selected taped assessments from this sample. Kappa coefficients for the personality disorders averaged .83 (SD = .16) and were acceptable for all PDs (range 0.70–1.0) except for paranoid PD (.38). Cronbach's alpha values ranged from .72 (schizotypal) to .93 (avoidant), with a median of .85.

S–PDQ–4. Like its English-language version the S–PDQ–4 contains 98 statements that are rated as true or false. The PDQ–4 is a very widely used measure and contains one item correspond-ing to each of the diagnostic criteria for the *DSM–IV* PDs. The Spanish-language version of the PDQ–4 was developed by the same fully bilingual research psychologists as the S–DIPD–IV using the same rigorous process of translation and back-translation. Its psychometric properties have not previously been examined, but alpha values in the current sample ranged from .48 (BPD) to .82 (avoidant), with a median of .72. Histrionic also obtained a relatively low value at .52, but all others were .66 or higher.

PAF. The PAF provides a standardized method to quantify clinicians' PD diagnoses and thus closely mirrors the type of PD ratings and diagnoses made in clinical practice. The PAF contains 3–4 sentence prototype descriptions for the 10 *DSM–IV* PDs, which the clinician rates on a 1–6 scale, where 1 indicated *not at all* and 6 indicated that the patient matched the prototype *to an extreme degree*. Consistent with previous research (Shea et al., 1990), a score ≥4 indicated a categorical diagnosis. The PAF was completed in English by the bilingual clinicians.

Results

Table 1 shows the number of participants diagnosed with each specific PD by each of the four methods. The bottom row of each column presents the number and percentage who met criteria for any of the 10 *DSM–IV* PDs. Noteworthy here is that borderline PD (n = 12) and antisocial PD (ASPD; n = 11) were the only PD diagnoses that were recorded within the medical charts. In addition, another three participants were classified as PDNOS, and a large number (n = 51) had Axis II diagnoses deferred. In contrast, all 10 PDs were diagnosed in multiple participants by each of the other three methods and yielded a much larger percentage of participants as having one or more PDs.

Our primary aim was to index agreement across methods. Because only two PDs were recorded in the medical charts, agreement could not be calculated for the other eight PDs. The categorical agreement (kappa coefficients) for the BPD chart

Table 1

Frequency of Diagnostic and Statistical Manual of Mental Disorders *(4th ed.; DSM–IV) PDs According to Four Different Methods*

Variable	Clinical chart (n = 128)		Clinician ratings (n = 126)		Semistructured interview (n = 130)		Self-report questionnaire (n = 115)	
	n	%	n	%	n	%	n	%
Paranoid	0	0.0	20	15.9	16	12.3	63	55.8
Schizoid	0	0.0	14	11.1	2	1.5	55	47.8
Schizotypal	0	0.0	11	8.7	6	4.6	52	46.8
Antisocial	11	8.6	33	26.2	17	13.1	34	29.6
Borderline	12	9.4	23	18.3	39	30.0	52	48.1
Histrionic	0	0.0	13	10.4	4	3.1	12	10.5
Narcissistic	0	0.0	9	7.1	4	3.1	25	22.9
Avoidant	0	0.0	15	11.9	34	26.2	51	45.5
Dependent	0	0.0	22	17.7	19	14.6	33	28.7
OCPD	0	0.0	11	8.7	34	26.2	49	44.1
PDNOS	3	2.3						
Deferred	51	39.8						
Any PD	22	17.2	69	54.8	74	56.9	90	78.3

Note. OCPD = obsessive-compulsive personality disorder; PDNOS = personality disorder not otherwise specified; PD = personality disorder.

Table 2

Diagnostic Agreement Among Therapist, Self-Report, and Semistructured Interview Personality Disorder Diagnoses

Variable	Clinician to interview		Clinician to self-report		Interview to self-report	
	κ	r	κ	r	κ	r
Paranoid	.29	.23*	.12	.17	.15	.57**
Schizoid	−.03	.03	.02	.06	.02	.46**
Schizotypal	.08	.24**	.09	.13	.06	.51**
Antisocial	.17	.35**	.23	.33**	.58	.63**
Borderline	.17	.38**	.20	.24	.48	.49**
Histrionic	−.05	.09	−.19	.04	.23	.34**
Narcissistic	.12	.01	.09	−.03	.12	.37**
Avoidant	.10	.05	.02	.11	.42	.74**
Dependent	.21	.20*	.00	.17	.24	.69**
OCPD	.05	.00	.06	.04	.39	.67**
M[a]	.11	.16	.06	.13	.27	.56

Note. Pairwise *n* ranged from 102–112 because of missing data. OCPD = obsessive-compulsive personality disorder.
[a] Correlations were transformed to *z* scores using Fisher's method and averaged before returning to correlations.
* $p < .05$. ** $p < .01$ (two-tailed).

diagnoses were .00, .25, and .09 with the PAF, S–DIPD–IV, and S–PDQ–4, respectively. The agreement of the ASPD chart diagnoses were .12 with the PAF, .36 with the S–DIPD–IV, and .26 with the S–PDQ–4.

Table 2 presents the categorical and dimensional agreement among the other methods. For instance, the first columns indicate that the kappas between the PAF and the S–DIPD–IV ranged from a low of −.05 (histrionic) to a high of .29 (paranoid), with an overall mean of 0.11. The agreement increased slightly when calculated dimensionally, with correlations ranging from .00 (OCPD) to .38 (BPD), with a mean of 0.16. Agreement between the PAF and the self-report S–PDQ–4 were even lower, with only two PDs (BPD and ASPD) obtaining values greater than .20. Kappas ranged from −.19 (histrionic) to .23 (ASPD), with a mean of 0.06. The dimensional agreement was similarly low, ranging from −.03 (narcissistic) to .33 (ASPD), with a mean of 0.13.

Higher agreement values were obtained between the semistructured interview and the self-report questionnaire, particularly when they were scored dimensionally. The kappa values ranged from .06 (schizotypal) to .58 (ASPD), and all but four values were larger than .20, with an overall mean of 0.27. The dimensional agreement between the S–PDQ–4 and the S–DIPD–IV was even more robust, ranging from .34 (histrionic) to a high of .74 (avoidant), with an overall mean of 0.56.

Table 2 results were further summarized by averaging the convergent values for each individual method with all others to arrive at a global indicator of its contribution, as was done by Samuel and Widiger (2010). For example, when one averages the correlation of the therapists' PAF ratings with the self-report measure (*r* = .12) and the interview (*r* = .16), it indicates that on average agreement for the therapists with other methods was .14. In contrast, comparable values for the self-report (mean *r* = .34) and semistructured interview (mean *r* = .36) were notably higher.

Discussion

In this study, we found that among a series of monolingual (Spanish-speaking only) Hispanic patients at a community mental health center, the rates of PDs diagnosed by four independent methods show only moderate-to-low levels of agreement. Our research within this important and understudied group extends previous work that has considered other combinations of these methods within predominantly White and English-speaking samples (Samuel et al., 2013). Prior research has generally indicated meager agreement among PD diagnoses assigned by alternate methods (Klonsky, Oltmanns, & Turkheimer, 2002), and the agreement in this study appears equally low. When considered categorically, there were only a handful of comparisons between any two diagnostic methods that exceeded .20.

A novel contribution of this study was the inclusion of both unstructured (i.e., chart diagnoses) and structured (i.e., PAF ratings) diagnoses provided independently by separate clinicians involved in the treatment of clients. These methods represent the standard approaches to PD diagnosis in clinical practice yet have not been compared with one another previously. Strikingly, the agreement between chart diagnoses and independent ratings by treating clinicians showed extremely low agreement for the two diagnoses that were recorded in charts (kappas of .00 for BPD and .12 for ASPD). Furthermore, the low interrater reliability might help to explain why clinicians' diagnoses obtained the lowest consensus agreement with the other methods both in this study and in prior research (Samuel & Widiger, 2010). Another strength of our method was the administration of a self-report questionnaire prior to the administration of the semistructured interview. This mirrors how self-report questionnaires are used in routine clinical settings and obviates any potential confound of practice effects from having completed an interview first. This yields confidence that our findings reflect the level of agreement that is obtained between self-report and clinical diagnoses in practice settings.

There did appear to be patterns of greater cross-method agreement for certain PD diagnoses. ASPD had among the highest convergence across diagnostic methods. Schizoid and histrionic, on the other hand, routinely obtained among the lowest agreement, even negative kappa values in several instances. It is unclear whether these findings are attributable to the frequencies of these PDs within the given sample or properties of the PDs. For example, the criteria for ASPD are the most behaviorally specific of any PD, and this fact may increase the agreement across sources by reducing opportunities for the introduction of subjectivity. Nonetheless, ASPD was among the most frequent diagnoses according to each method, as would be expected given the sample's history of substance use (Trull et al., 2010). Replication of these findings within additional samples, with different rates PDs would be helpful in clarifying this issue.

Although the agreement of PD diagnoses across methods and particularly for the diagnoses provided by treating clinicians within routine clinical practice was low by objective standards, it is important to note that this is not likely unique to PDs. The present dataset does not have information regarding the rates of other, non-PD diagnoses, but meta-analyses suggest that agreement across methods for other mental disorders are quite comparable (Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009).

Thus, our findings likely reflect diagnostic practices more broadly across the *DSM*.

It is important to note, however, that the convergence does not itself suggest which diagnostic method should be considered the most valid. Thus, it would be inappropriate to conclude from our findings that clinicians' PD diagnoses are necessarily less valid than other methods. Nonetheless, our findings are consistent with prior research that has examined predictive validity. Samuel et al. (2013) compared the validity of clinicians' PD diagnoses to those assigned by a self-report questionnaire and a semistructured interview for predicting prospective psychosocial functioning 5 years longitudinally. They noted that both self-report and interview diagnoses regularly incremented the prediction of therapists' ratings, yet the reverse was only rarely true, even when accounting for clinicians' familiarity with the individual. Thus, converging evidence suggests that the PD diagnoses assigned by clinicians within routine clinical practice or when using a brief measure such as the PAF are less valid than more systematic approaches.

This study also provides data indicating PDs are likely to be underdiagnosed within typical clinical practice. The overall rate of *DSM–IV* PD diagnoses within medical charts was notably lower than the rates diagnosed by semistructured interview, self-report questionnaire, and even independent treating clinicians using a structured rating system (i.e., the PAF). One possible explanation for the underdiagnosis within medical charts is that clinicians do not routinely obtain enough information to definitively assess the presence of a PD. The fact that Axis II PD diagnoses were deferred for 51 participants provides support for this hypothesis. In light of this finding, we recommend the incorporation of more systematic diagnostic tools within general practice to facilitate the identification of PDs.

Limitations

This study extends research on the clinical diagnosis of PDs in important ways, but it is not without limitations. It provided the first evaluation of monolingual Hispanic individuals but was confined to this ethnic group. Generalizability to bilingual Hispanic persons and to those of other ethnic–racial groups is uncertain. Our findings also may not generalize to other clinicians, different settings, alternative diagnostic procedures, or different measures. Nonetheless, our findings do converge with the existing literature that has examined these issues in a piece-meal fashion (e.g., self-report compared to semistructured interview). Moreover, our findings dovetail with those reported by Samuel et al. (2013) regarding the agreement of three methods within a large multisite study suggesting clinic-level idiosyncrasies or demographic variables have not solely shaped our results.

Additionally, although the size was moderate ($n = 130$), larger clinical groups might have allowed more precise estimates of convergence, particularly for less frequently occurring PDs. Notably, our sample consisted of individuals who were in treatment of substance abuse. This surely impacted the frequency of specific PD diagnoses (Trull et al., 2010). The highest cross-method agreement for individual PD diagnoses was repeatedly found for the diagnoses of ASPD and BPD and is perhaps attributable to the somewhat higher rates of these two PDs within our patient group. Nonetheless, even higher frequency of these PDs did not increase the convergent coefficients appreciably as the agreement between

the clinicians' PAF ratings and the semistructured interview diagnoses remained $\kappa < .20$ for these PDs, alleviating this concern.

A more specific limitation is that the S–DIPD–IV paranoid PD scale obtained inadequate interrater reliability ($K = .38$) within this sample. Although internal consistency of the paranoid PD in this sample was .80, the poor interrater reliability raises concerns about the findings from this scale. In this regard, though, we noted that the dimensional and categorical agreement for the S–DIPD–IV paranoid PD scale was quite comparable with other S–DIPD–IV scales. Finally, it should be noted that the S–PDQ–4 was translated for this study and does not have previously reported psychometric properties. Although these scales generally performed as might be expected (given their relationships with the interview measure), the alpha values for the BPD and histrionic scales were somewhat lower than ideal.

Conclusion

A novel contribution of this study is that convergence of PD diagnoses across diagnostic methods for Spanish-speaking Hispanic persons are comparable to other groups. This provides additional evidence allaying concerns about cross-cultural (and cross-language) application of PD constructs and perhaps the diagnostic manual as a whole. More globally, the results provided a unique perspective on the convergence of clinicians' PD diagnoses, both as recorded within medical charts and as assigned by systematic measures, with other methods of diagnosis. Important to note is that our results echo previous research in suggesting that clinicians' PD diagnoses share very little overlap with those assigned by self-report questionnaires or semistructured diagnostic interviews and extend them to suggest that PDs are underdiagnosed using standard diagnostic approaches.

References

Altarriba, J., & Santiago-Rivera, A. L. (1994). Current perspectives on using linguistic and cultural factors in counseling the Hispanic client. *Professional Psychology: Research and Practice, 25,* 388–397. doi: 10.1037/0735-7028.25.4.388

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Añez, L. M., Paris, M., Jr., Bedregal, L. E., Davidson, L., & Grilo, C. M. (2005). Application of cultural constructs in the care of first generation Latino clients in a community mental health setting. [Case Reports]. *Journal of Psychiatric Practice, 11,* 221–230. doi:10.1097/00131746-200507000-00002

Ansell, E. B., Pinto, A., Crosby, R. D., Becker, D. F., Añez, L. M., Paris, M., & Grilo, C. M. (2010). The prevalence and structure of obsessive-compulsive personality disorder in Hispanic psychiatric outpatients. *Journal of Behavior Therapy and Experimental Psychiatry, 41,* 275–281. doi:10.1016/j.jbtep.2010.02.005

Arnold, B. R., & Matus, Y. E. (2000). Test translation and cultural equivalence methodologies for use with diverse populations. In I. Cuellar & F. A. Paniagua (Eds.), *Handbook of multicultural mental health* (pp. 121–136). San Diego, CA: Academic Press. doi:10.1016/B978-012199370-2/50008-0

Bender, D. S., Skodol, A. E., Dyck, I. R., Markowitz, J. C., Shea, M. T., Yen, S., . . . Grilo, C. M. (2007). Ethnicity and mental health treatment utilization by patients with personality disorders. *Journal of Consulting and Clinical Psychology, 75,* 992–999. doi:10.1037/0022-006X.75.6.992

Chavira, D. A., Grilo, C. M., Shea, M. T., Yen, S., Gunderson, J. G., Morey, L. C., . . . Mcglashan, T. H. (2003). Ethnicity and four personality disorders. *Comprehensive Psychiatry, 44,* 483–491. doi:10.1016/S0010-440X(03)00104-4

Chick, D., Sheaffer, C. I., Goggin, W. C., & Sison, G. F. (1993). The relationship between MCMI personality scales and clinician-generated *DSM–III–R* personality disorder diagnoses. *Journal of Personality Assessment, 61,* 264–276. doi:10.1207/s15327752jpa6102_8

Cuellar, A., & Paniagua, F. A. (Eds.). (2000). *Handbook of multicultural mental health*. San Diego, CA: Academic Press.

Ennis, S. R., Rios-Vargas, M., & Albert, N. G. (2011). The Hispanic Population: 2010. *2010 Census Briefs*.

Grilo, C. M., Añez, L. M., & McGlashan, T. H. (2003). The Spanish-language version of the diagnostic interview for DSM–IV personality disorders: Development and initial psychometric evaluation of diagnoses and criteria. *Comprehensive Psychiatry, 44,* 154–161. doi:10.1053/comp.2003.50006

Hyler, S. E. (1994). *Personality Diagnostic Questionnaire—4 (PDQ–4)*. New York, NY: State Psychiatric Institute.

Hyler, S. E., Rieder, R. O., Williams, J. B., & Spitzer, R. L. (1989). A comparison of clinical and self-report diagnoses of *DSM–III* personality disorders in 552 patients. *Comprehensive Psychiatry, 30,* 170–178. doi:10.1016/0010-440X(89)90070-9

Klein, M. H., Benjamin, L. S., Rosenfeld, R., Treece, C., Husted, J., & Greist, J. H. (1993). The Wisconsin Personality Disorders Inventory: Development, reliability, and validity. *Journal of Personality Disorders, 7,* 285–303. doi:10.1521/pedi.1993.7.4.285

Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice, 9,* 300–311. doi:10.1093/clipsy.9.3.300

Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *The American Journal of Psychiatry, 149,* 1645–1653.

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18,* 169–184. doi:10.1002/mpr.289

Samuel, D. B., Sanislow, C. A., Hopwood, C. J., Shea, M. T., Skodol, A. E., Morey, L. C., . . . Grilo, C. M. (2013). Convergent and incremental predictive validity of clinician, self-report, and structured interview diagnoses for personality disorders over 5 years. *Journal of Consulting and Clinical Psychology, 81,* 650–659.

Samuel, D. B., & Widiger, T. A. (2010). Comparing personality disorder models: Cross-method assessment of the FFM and *DSM–IV–TR*. *Journal of Personality Disorders, 24,* 721–745. doi:10.1521/pedi.2010.24.6.721

Shea, M. T., Pilkonis, P. A., Beckham, E., Collins, J. F., Elkin, I., Sotsky, S. M., & Docherty, J. P. (1990). Personality disorders and treatment outcome in the NIMH treatment of depression collaborative research program. *The American Journal of Psychiatry, 147,* 711–718.

Tenney, N. H., Schotte, C. K. W., Denys, D. A. J. P., van Megen, H. J. G. M., & Westenberg, H. G. M. (2003). Assessment of *DSM–IV* personality disorders in obsessive-compulsive disorder: Comparison of clinical diagnosis, self-report questionnaire, and semi-structured interview. *Journal of Personality Disorders, 17,* 550–561. doi:10.1521/pedi.17.6.550.25352

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised Nesarc personality disorder diagnoses: Gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders, 24,* 412–426. doi:10.1521/pedi.2010.24.4.412

U.S. Bureau of the Census. (1995). *The Hispanic population in the United States: March 1993* (P20, No. 475). Washington, DC: U.S. Government Printing Office.

Widiger, T. A., & Samuel, D. B. (2005). Evidence-based assessment of personality disorders. *Psychological Assessment, 17,* 278–287. doi:10.1037/1040-3590.17.3.278

Zanarini, M. C., Frankenburg, F. R., Sickel, A. E., & Yong, L. (1996). *The Diagnostic Interview for DSM–IV Personality Disorders (DIPD-IV)*. Belmont, MA: McLean Hospital.

Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry, 51,* 225–245. doi:10.1001/archpsyc.1994.03950030061006