Published in final edited form as:

Personal Disord. 2014 April ; 5(2): 172–177. doi:10.1037/per0000033.

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,

Yale School of Medicine

Manuel Paris, and

Yale School of Medicine

Carlos M. Grilo

Yale School of Medicine

Abstract

Methods for diagnosing personality disorders (PDs) within clinical settings typically diverge from those employed in treatment research. Treatment groups in research studies are routinely diagnosed using semi-structured 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 *four* independent methods of PD diagnosis, including self-report questionnaire, semi-structured interview, chart diagnoses, and ratings by treating clinicians, within a clinical series of 130 monolingual (Spanish-only) Hispanics (69% male; mean age 37.4), in treatment for substance use. We examined the convergence of the *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 semi-structured 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 Hispanics are comparable to other groups allaying concerns about cross cultural application of PD diagnoses. Additionally, our results echo previous research in suggesting that clinicians' PD diagnoses overlap little with self-report questionnaires or semi-structured diagnostic interviews and suggest that PDs are under-diagnosed using standard diagnostic approaches. Implications for the clinical application of empirically supported research are discussed.

Correspondence concerning this article should be addressed to: Douglas B. Samuel, Department of Psychological Sciences, Purdue University, 703 Third St, West Lafayette, IN 47907. dbsamuel@purdue.edu.
Douglas B. Samuel, Department of Psychological Sciences, Purdue University; Luis M. Añez Manuel Paris, and Carlos M. Grilo, Department of Psychiatry, Yale School of Medicine.

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 and/or semi-structured diagnostic interviews 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., in press).

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., $\kappa = .08$, 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 semi-structured diagnostic interviews (Samuel et al., in press), 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 semi-structured diagnostic 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 semi-structured interview and self-reported diagnoses (Samuel et al., in press; 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 and colleagues (2003) considered the categorical PD diagnoses that were recorded in the clients' clinical charts, Samuel and colleagues (in press) utilized 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* to clinicians' *unstructured ratings* of PDs, let alone compared these two methods with diagnoses assigned by semi-structured 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 (Anez, 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 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, Anez, & McGlashan, 2003). In the U.S., many Hispanics 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).

The current report addresses these limitations by examining the diagnostic convergence of PD diagnoses assigned by *four* different perspectives: self-report questionnaire, semi-structured interview, clinical chart diagnoses, and ratings provided by the treating clinicians. These data are significant because they are from a sample of monolingual (Spanish-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, that is 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 comprised of a series of 130 monolingual

Hispanic adult patients assigned to a particular treatment team over a one-year. Patients were assigned to this team based on caseload 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; 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: (1) a semi-structured diagnostic interview administered by research clinicians, (2) a self-report questionnaire completed by clients, (3) a structured rating form completed by the treating clinicians, and (4) 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 DSM-IV 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 semi-structured 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). Inter-rater 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 .70 to 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 corresponding 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 ten *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 ten 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 there 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 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 .11. The agreement increased slightly when calculated dimensionally, with correlations ranging from .00 (OCPD) to .38 (BPD), with a mean of .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 $.06$. The dimensional agreement was similarly low, ranging from $-.03$ (narcissistic) to $.33$ (ASPD), with a mean of $.13$.

Higher agreement values were obtained between the semi-structured 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 $.27$. 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 $.56$.

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 semi-structured interview (mean $r = .36$) were notably higher.

Discussion

The study found that among a series of monolingual (Spanish-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., in press). Prior research has generally indicated meager agreement among PD diagnoses assigned by alternate methods (Klonsky, Oltmanns, & Turkheimer, 2002) and the agreement in the current 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 the present 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 to 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 inter-rater 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 semi-structured 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 if 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 the present 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, the present findings are consistent with prior research that has examined predictive validity. Samuel and colleagues (in press) compared the validity of clinicians' PD diagnoses to those assigned by a self-report questionnaire and a semi-structured interview for predicting prospective psychosocial functioning five 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.

The present study also provides data indicating PDs are likely to be under-diagnosed within typical clinical practice. The overall rate of *DSM-IV* PD diagnoses within medical charts was notably lower than the rates diagnosed by semi-structured interview, self-report questionnaire, and even independent treating clinicians using a structured rating system (i.e., the PAF). One possible explanation for the under-diagnosis 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

The present study extends research on the clinical diagnosis of PDs in important ways, but is not without limitations. The study provided the first evaluation of monolingual Hispanics, but was confined to this ethnic group. Generalizability to bi-lingual Hispanics 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 semi-structured interview). Moreover, our findings dovetail with those reported by Samuel et al. (in press) regarding the agreement of three methods within a large multi-site 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 semi-structured 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 inter-rater reliability raises concerns about the findings from this scale. In this regard, though, we note that the dimensional and categorical agreement for the S-DIPD-IV paranoid PD scale was quite comparable to 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.

Conclusions

A novel contribution of the current study is that convergence of PD diagnoses across diagnostic methods for Spanish-speaking Hispanics 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. Importantly, our results echo previous research in suggesting that clinicians' PD diagnoses share very little overlap with those assigned by self-report questionnaires or semi-structured diagnostic interviews and extend them to suggest that PDs are under-diagnosed using standard diagnostic approaches.

Acknowledgments

Research was supported by grants from the NIH (MH50850).

References

- Altarriba J, Santiago-Rivera AL. Current perspectives on using linguistic and cultural factors in counseling the Hispanic client. *Professional Psychology: Research and Practice*. 1994; 25(4):388–397.10.1037//0735-7028.25.4.388
- Anez LM, Paris M Jr, Bedregal LE, Davidson L, Grilo CM. Application of cultural constructs in the care of first generation Latino clients in a community mental health setting. [Case Reports]. *Journal of Psychiatric Practice*. 2005; 11(4):221–230. [PubMed: 16041232]
- Ansell EB, Pinto A, Crosby RD, Becker DF, Anez LM, Paris M, Grilo CM. The prevalence and structure of obsessive-compulsive personality disorder in Hispanic psychiatric outpatients. *Journal of Behavior Therapy and Experimental Psychiatry*. 2010; 41(3):275–281.10.1016/j.jbtep.2010.02.005 [PubMed: 20227063]
- Arnold, BR.; Matus, YE. Test translation and cultural equivalence methodologies for use with diverse populations. In: Cuellar, I.; Paniagua, FA., editors. *Handbook of Multicultural Mental Health*. San Diego, CA: Academic Press; 2000. p. 121-136.
- Bender DS, Skodol AE, Dyck IR, Markowitz JC, Shea MT, Yen S, Grilo CM. Ethnicity and mental health treatment utilization by patients with personality disorders. *Journal of Consulting and Clinical Psychology*. 2007; 75(6):992–999.10.1037/0022-006x.75.6.992 [PubMed: 18085915]
- Chavira DA, Grilo CM, Shea MT, Yen S, Gunderson JG, Morey LC, Mcglashan TH. Ethnicity and four personality disorders. *Comprehensive Psychiatry*. 2003; 44(6):483–491.10.1016/S0010-440x(03)00104-4 [PubMed: 14610727]
- Chick D, Sheaffer CI, Goggin WC, Sison GF. The relationship between MCMI personality scales and clinician-generated DSM-III—R personality disorder diagnoses. *Journal of Personality Assessment*. 1993; 61(2):264–276.10.1207/s15327752jpa6102_8 [PubMed: 8229636]
- Cuellar, A.; Paniagua, FA., editors. *Handbook of Multicultural Mental Health*. San Diego, CA: Academic Press; 2000.
- Ennis SR, Rios-Vargas M, Albert NG. *The Hispanic Population: 2010*. 2010 Census Briefs. 2011 Retrieved 11/20/2012, 2012.
- Grilo CM, Anez LM, McGlashan TH. The Spanish-language version of the diagnostic interview for DSM-IV personality disorders: Development and initial psychometric evaluation of diagnoses and criteria. *Comprehensive Psychiatry*. 2003; 44(2):154–161.10.1053/comp.2003.50006 [PubMed: 12658625]
- Hyler, SE. *Personality Diagnostic Questionnaire-4 (PDQ-4)*. New York State Psychiatric Institute; New York: 1994.
- Hyler SE, Rieder RO, Williams JB, Spitzer RL. A comparison of clinical and self-report diagnoses of DSM-III personality disorders in 552 patients. *Comprehensive Psychiatry*. 1989; 30(2):170–178.10.1016/0010-440x(89)90070-9 [PubMed: 2920552]
- Klein MH, Benjamin LS, Rosenfeld R, Treece C, Husted J, Greist JH. The Wisconsin Personality Disorders Inventory: Development, reliability, and validity. *Journal of Personality Disorders*. 1993; 7(4):285–303.
- Klonsky ED, Oltmanns TF, Turkheimer E. Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology-Science and Practice*. 2002; 9(3):300–311.
- Perry JC. Problems and Considerations in the Valid Assessment of Personality Disorders. *American Journal of Psychiatry*. 1992; 149(12):1645–1653. [PubMed: 1443240]
- Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research*. 2009; 18(3):169–184.10.1002/mp.289 [PubMed: 19701924]

- Samuel DB, Sanislow CA, Hopwood CJ, Shea MT, Skodol AE, Morey LC, Grilo CM. Convergent and incremental predictive validity of clinician, self-report, and diagnostic interview assessment methods for personality disorders over five years. *Journal of Consulting and Clinical Psychology*. (in press).
- Samuel DB, Widiger TA. Comparing personality disorder models: Cross-method assessment of the FFM and DSM-IV-TR. *Journal of Personality Disorders*. 2010; 24(6):721–745.10.1521/pedi.2010.24.6.721 [PubMed: 21158596]
- Shea MT, Pilkonis PA, Beckham E, Collins JF, Elkin I, Sotsky SM, Docherty JP. Personality disorders and treatment outcome in the NIMH treatment of depression collaborative research program. *American Journal of Psychiatry*. 1990; 147:711–718. [PubMed: 2343912]
- Tenney NH, Schotte CKW, Denys DAJP, van Megen HJGM, Westenberg HGM. 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*. 2003; 17(6): 550–561.10.1521/pedi.17.6.550.25352 [PubMed: 14744080]
- Trull TJ, Jahng S, Tomko RL, Wood PK, Sher KJ. Revised Nesarc Personality Disorder Diagnoses: Gender, Prevalence, and Comorbidity with Substance Dependence Disorders. *Journal of Personality Disorders*. 2010; 24(4):412–426. [PubMed: 20695803]
- U.S. Bureau of the Census. *The Hispanic population in the United States*: March 1993. Washington, DC: US Government Printing Office; 1995. (P20, No. 475)
- Widiger TA, Samuel DB. Evidence-based assessment of personality disorders. *Psychological Assessment*. 2005; 17(3):278–287.10.1037/1040-3590.17.3.278 [PubMed: 16262454]
- Zanarini, MC.; Frankenburg, FR.; Sickel, AE.; Yong, L. *The Diagnostic Interview for DSM-IV Personality Disorders (DIPD-IV)*. Belmont, MA: McLean Hospital; 1996.
- Zimmerman M. Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry*. 1994; 51(3):225–245.10.1001/archpsyc.1994.03950030061006 [PubMed: 8122959]

Table 1

Frequency of DSM-IV PDs According to Four Different Methods

	Clinical Chart (n = 128)		Clinician Ratings (n = 126)		Semi-structured 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

Notes: PDNOS = Personality Disorder Not Otherwise Specified. PDNOS and Deferred diagnoses available for chart data only. Any PD refers to the number of percent of individuals who met criteria for any of the 10 DSM-IV PD diagnoses by each method.

Diagnostic Agreement among Therapist, Self-Report, and Semi-structured Interview Personality Disorder Diagnoses

Table 2

	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**
mean ^a	.11	.16	.06	.13	.27	.56

Notes: Pairwise n ranged from 102–112 because of missing data. OCPD = Obsessive Compulsive Personality Disorder.

* $p < .05$,
** $p < .01$ (two tailed).

^aCorrelations were transformed to z scores using Fisher's method and averaged before returning to correlations.