

THE SPANISH VERSION OF THE WISCONSIN PERSONALITY DISORDERS INVENTORY-IV (WISPI-IV): TESTS OF VALIDITY AND RELIABILITY

Tracey L. Smith, PhD, Marjorie H. Klein, PhD, Carmen Alonso, PhD, José Salazar-Fraile, MD, PhD, Elena Felipe-Castaño, PhD, Clara Lopez Moreno, MD, Silvia Raquel Acosta, PhD, Lisbeth Iglesias Rios, PhD, and Vicent Martí-Sanjuán, PhD

Personality disorders (PD) are a prevalent class of mental disorders that interfere with functioning and cause subjective distress while increasing the intensity and duration of Axis I clinical syndromes, and therefore assessing PD is important even when PDs are not the focus of treatment. The purpose of these studies was to develop and test a new Spanish version of a self-report measure of PD, the Wisconsin Personality Inventory-IV (WISPI-IV) that would be psychometrically equivalent to the English version while also maintaining the same interpersonal content, which is based on Benjamin's analysis of the PD criteria using her Structural Analysis of Social Behavior (SASB) model (1974). Study 1 participants completed the WISPI-IV twice over a two-week interval. For Study 2, participants from two sites in Spain and one site in Argentina completed Spanish versions of the WISPI-IV and other personality measures. SASB-analysis of the translated items showed high correspondence between the interpersonal content of the English version and the Spanish version demonstrating theoretical validation in relation to other PD measures. The Spanish WISPI-IV showed satisfactory reliability based on test-retest correlations and alphas for internal consistency. Study 2 showed the Spanish WISPI-IV had good convergent validity with the Spanish versions of the IIP and SCID-II and performed similarly to the English versions of these measures. Our goal in translating the WISPI-IV was to extend this measure to Spanish-speakers in language that would be understood by different Hispanic sub-

From University of Wisconsin (T. L. S., M. H. K., C. A.); Centro de Salud Mental de Paterna. Valencia CIBER Salud Mental. Instituto Carlos III. (J. S.-F.); University of Extremadura, Spain (E. F.-C.); CAECE University, Buenos Aires, Argentina (C. L. M., S. R. A.); University of New Mexico (L. I. R.); and Centro de Salud Mental de Paterna. Valencia (V. M.-S.).

We would like to thank: Cristina Miralles Cardona (Wisconsin) and Jennifer Vicuna (New Mexico) for their valuable assistance with data collection and recruiting; and to Dr. Scott Tonigen for serving as the faculty sponsor at the New Mexico site. We are also grateful to Ken Critchfield and Aidan Wright for their many helpful comments on a draft of this article.

Address correspondence to Tracey L. Smith, Ph.D., University of Wisconsin, Dept. of Psychiatry, 6001 Research Park Blvd., Madison, WI 53719; E-mail: Tracey.Smith2@va.gov

groups, however research team members and subjects had a variety of suggestions for changes in item wording. This reflects the difficulty with creating a “neutral” Spanish version of any assessment given regional differences.

The WISPI-IV (Klein & Benjamin, 1996) is a self-report measure of DSM-IV personality disorders (PD) derived from an interpersonal perspective. The aim in constructing the WISPI-IV was to create items that were consistent with the DSM-IV descriptions of the personality disorders (American Psychiatric Association, 2000) and also with Benjamin’s (1993, 1996) interpersonal analysis of the DSM PD criteria using Benjamin’s (1974) Structural Analysis of Social Behavior (SASB) model. This model characterizes interpersonal behavior according to three distinctions: (1) focus (on other with self, self with other, and within self) and the orthogonal circumplex dimensions of (2) affiliation versus hostility and (3) interdependence versus autonomy (see Figure 1). The model suggests links between early object relationships, social patterns, and the development of the self-concept (e.g., Benjamin, 1987). Its construct validity, reliability, internal consistency, and circumplex structure have been demonstrated empirically (Benjamin, 1974, 1984).

As applied to the personality disorders (PD), Benjamin (1993, 1996) first developed formulations and descriptions of each of the PD categories, con-

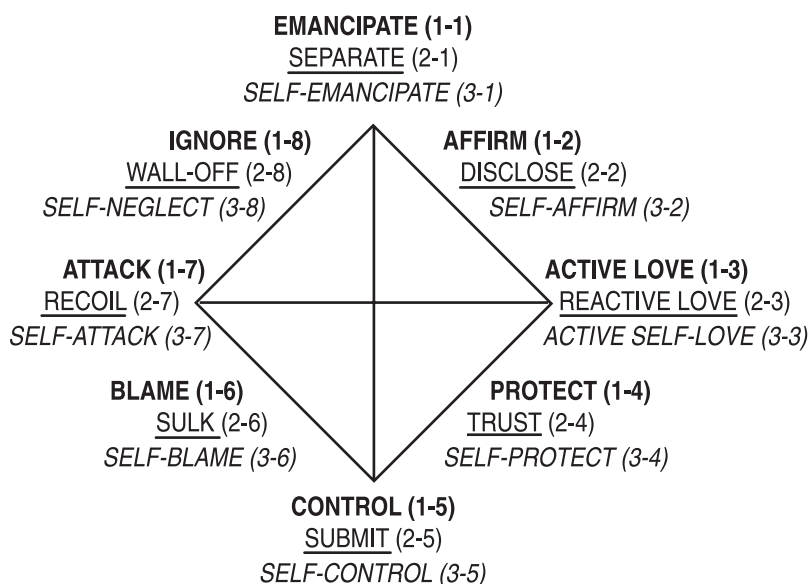


FIGURE 1. The simplified SASB cluster model (adapted from Benjamin, 1996).

sidering the interpersonal context in which the behavior patterns associated with each PD are hypothesized to have developed and suggest interpersonal issues that are likely to continue in adulthood. For example, for the Avoidant (AVD) PD, early experiences of nurturance and bonding gave him or her sense of attachment and social bonding; but as the developing AVD was also subject to strong parental controls to maintain a pleasing social image, any flaws were cause for humiliation and embarrassment. Thus the adult AVD has learned to perform adequately but is very concerned about making social mistakes and being the subject of ridicule and rejection. Benjamin coded these developmental formulations, using the SASB model (Figure 1), to reflect the interpersonal feature of each PD (1993, 1996). For AVD, for example, the early loving nurturance was coded as receiving Active Love and Protection from others. At the same time, the parental Control and Blame lead to the development of AVD's vulnerability to Self-Blame and needs for Self-Control. This formulation was then employed as the basis for SASB-coding of the Axis II criteria and for the construction of the corresponding WISPI-IV items. For example DSM-IV (American Psychiatric Association, 2000) criterion #1 "avoids occupational activities that involve significant interpersonal contact, because of fears of criticism, disapproval, or rejection is reflected in items such as "Because I get embarrassed so easily, I avoid jobs and social situations which would force me to be with people more." This item was coded 1-6 for fear of Blame from others and 2-1 for the self Separating from others. Another example is the item for criterion #6 "I am very shy in social situations because I know I'm awkward and unappealing," which was coded 2-8 for Walling-Off and 3-6 for Self-Blame. The three most frequent SASB codes for the 16 Avoidant PD items were 2-8 for Walling-Off, 1-6 for fear of Blame from others, and 3-5 for Self-Control. Therefore, each WISPI item conforms to DSM-IV criteria and Benjamin's SASB-coded interpersonal conception of these criteria and attempts to capture the phenomenology of a person with a given PD.

As illustrated above, the content of the WISPI-IV PD items reflect Benjamin's conceptions of the interpersonal patterns and self-directed styles associated with each of the PDs. For example, in the case of Paranoid PD, the primary focus is on fears of being Attacked (1-7), Blamed (1-6), or Ignored (1-8) by others. Schizoid and Schizotypal PDs share patterns of Walling-Off from others (2-8), but for different reasons; while in Schizoid PD there is no need or interest in relationships, the Schizotypal is concerned with protection from perceived Blame (1-6) or Attacks (1-7) by others. Histrionic and Narcissistic PDs share needs for Affirmation (1-2), Love (1-3), or Protection (1-4) from others, but in Histrionic PD this is obtained through Control (1-5), while the Narcissistic PD is focused on Self-Love (3-3) and Blaming others (1-6) for their shortcomings. In the case of Antisocial PD, there are patterns of Attacking (1-7) or Blaming (1-6) others, accompanied by the need for Autonomy (2-1). Fears of being Ignored (i.e.,

abandoned) (1-8) are prominent in Borderline PD, but Self-Neglect (3-8) is also present. Avoidant and Dependent PDs share fears of being Blamed (1-6) or Attacked (1-7) by others, but the Avoidant is most likely to Wall-Off (2-8) from others in response while the Dependent internalizes the Blame and Attack and Self-Blames (3-6). The primary concern of the Obsessive-Compulsive PD is to Control (1-5) others or to Submit (2-5) in service of needs for perfection, but the cost is Self-Neglect (3-8) in achieving goals and maintaining relationships.

The current English version (WISPI-IV; Klein & Benjamin, 1996) is the third revision of this personality disorder measure that coincided with changes in the DSM. It consists of 214 items for 11 PD scales (including Passive-Aggressive) and the Marlowe-Crowne Social Desirability Scale (Greenwald & Satow, 1970). Subjects are asked to think of their "usual selves during the past five years or more" and rate each item on a 1-10 scale, with 1 "never or not at all true of you" to 10 "always to extremely true of you." Scores for each scale are the mean of ratings of all items. Analyses can also use ipsatized or normative WISPI-IV scores. For the two studies reported here we chose to focus on the ten primary PDs in the DSM-IV, thus excluding Passive-Aggressive PD.¹ The WISPI-IV version was tested for concordance with the SCID-II (First, Gibbon, Spitzer, Williams, & Benjamin, 1996) using a psychiatric inpatient population selected for potential PD status (Smith, Klein, & Benjamin, 2003). Although the correspondence between categorical WISPI and SCID-II diagnoses was poor, the effect sizes for the difference in WISPI-IV means for groups with and without SCID-II diagnoses were large ($>.80$; Cohen, 1988). When SCID-II and WISPI-IV dimensional scores were considered the average r between profiles was .61 (median = .58) and the correlation between corresponding PD scales averaged .48; compared with $r = .18$ for off-diagonal correlations. Alphas for the 11 PD scales ranged from .74 to .91. When an earlier English DSM-III-R version of the WISPI was studied in a mixed sample of student volunteers and psychiatric outpatients, the scales demonstrated strong internal consistency. Alphas for the 11 PD scales of the WISPI-III-R ranged from .81 to .95 (Klein et al., 1993). Tests of validity of the WISPI-III-R included: a test of content validity in which two clinicians sorted WISPI items into the DSM PD categories, yielding kappas of .81 and .90; comparison of patients versus nonpatients on the WISPI (effect sizes ranged from .06 for Histrionic to .79 for Avoidant, $M = .35$; Median = 32; Klein et al., 1993) comparisons of patients with PD to patients without PD to test discriminant validity (effect sizes ranged from .14 for Antisocial to .84 for Schizotypal, $M = .56$; Median = .59; Barber & Morse, 1994; Klein et al., 1993); and tests of concurrent validity with other measures of PD (Barber & Morse, 1994; Klein et al., 1993).

1. This diagnosis is listed in DSM-IV's Appendix B ("Criteria Sets and Axes Provided for Further Study") because of controversy and the need for further research on how to also categorize the behaviors.

SPANISH TRANSLATIONS OF THE WISPI-IV

The WISPI-IV was first translated into Spanish in a Mexican mental health setting (Robles-García, Nabel, & Agraz, 2003). When administered to 227 subjects, its psychometric properties were similar to those reported for the original DSM-III-R version of the WISPI (Klein et al., 1993), however the correspondence to the underlying SASB theoretical structure was not tested. Furthermore, when the Robles translation was analyzed by a Spanish-speaking clinical psychologist (CA) who applied the SASB codes to the Robles' WISPI-IV items, we discovered 23 items that had major errors as well as 55 items with minor problems either in the translation and/or in their correspondence to the SASB codes (representing 36% of the WISPI items). Errors in translation were often due to misinterpretation of English idiomatic expressions. For example, an item with major problems was the translation for the English WISPI item 75 for Paranoid PD, "I've been harassed by people, so when I detect the slightest hint of trouble, I strike back or get away instantly." The translation for the Spanish version of this item was "I've been harassed by people, so when I detect the root of the problem, I retire immediately." This translation changes the meaning of the item and fails to capture both the vigilance of the person with Paranoid PD but also their tendency to attack in retaliation. An example of an item with minor problems was English WISPI item 162 for Histrionic PD, "I enjoy flirting and usually succeed in arousing people's interest." The translation for the Mexican version of this item was "I enjoy flirting and sometimes that is attractive to people." This translation changes the interpersonal focus of the item from the HPD's interpersonal perception of others to the reaction of others to that person. Therefore, the purpose of our new translation was to develop a Spanish language version of the WISPI that maintained the same underlying interpersonal conception as the English version, while also conforming to the DSM-IV PD criteria.

After describing the translation and coding procedures we will report two studies of the psychometric properties and tests of validity in Spanish-speaking participants in the United States (Study 1) and in Spain and Argentina (Study 2). In study 1 the validity criterion is the underlying interpersonal content of the WISPI-IV; in study 2 the validity criteria is the DSM-IV, as reflected in the SCID-II and the interpersonal dimensions tapped by the Inventory of Interpersonal Problems (IIP-64 also called the IIP-C; Alden, Wiggins, & Pincus, 1990; Horowitz, Alden, Wiggins, & Pincus, 2000) which is based on an interpersonal circumplex model that shares many features of the Benjamin model (Horowitz, Rosenberg, & Bartholomew, 1993). The IIP measures the kinds of interpersonal problems people have and the distress those problems cause them. The IIP-64 yields 8 octant scores organized around two orthogonal axes of dominance and nurturance. The SCID-II and the IIP-64 have been translated into Spanish (SCID-II: First, Gibbon, Spitzer, Williams, & Benjamin, 1999; IIP-64: Salazar, Martí, Soriano, Beltran, & Adam, 2010).

TRANSLATION PROCEDURE

Our goal was to create a Spanish language translation of the WISPI-IV that was both psychometrically equivalent to the English version and which could be readily understood by Spanish speakers of varying backgrounds through the use of neutral Spanish that was free of cultural idiosyncrasies (e.g., Mexican idioms can be quite different from Cuban idioms). We followed translation guidelines that have been developed by professional translators and other developers of psychosocial assessments who have translated various mental health measures (e.g., Matías-Carrelo et al., 2003). To begin the translation process, we first had a bilingual Spanish-national clinical psychologist (CA), who was well-versed in the theoretical model, which underlies the WISPI-IV; make the first translation from English into Spanish. This translation began after reviewing the SASB codes associated with each WISPI-IV item with a trained and reliable SASB coder so that the translation would capture the same interpersonal content structure as the original measure. This first translation was then given to a bilingual colleague in Spain who has a B.A. in Spanish Philology (the humanistic study of language and literature). This colleague edited the measure with the goal of making sure the translation captured the subtleties of the original measure and used neutral Spanish.

VERIFICATION OF THE SASB CODE CONCORDANCE

In order to confirm that the translated WISPI items reflected Benjamin's (1993, 1996) interpersonal conception of the PDs we first had a professional Spanish-English translator create a back-translation (from the edited Spanish version back to English). Then the translation team reviewed the items in the back-translation and compared them to the original English items. After making revisions to the Spanish version and back-translating the changed items, the SASB coder (TLS) coded the English back-translation to make sure that the items conformed to the underlying interpersonal content that is the basis of the measure. The SASB codes from the back-translation were compared to the original SASB codes created by Dr. Benjamin during the creation of the WISPI-IV with revisions made as needed. The kappas for agreement between the SASB codes for the PD scales of the English WISPI-IV and the codes for the back-translation range from .66 (Dependent) to .85 (Borderline), median .81. It is notable that these values are equivalent and in some cases higher than those found for the agreement among SASB coders in other studies (Critchfield, 2002; Harder & Henry, 1998; Humes & Humphrey, 1994).

STUDY 1

Study 1 was designed to provide psychometric information including test-retest and alpha reliabilities.

METHOD

The study design included two arms, a bilingual arm for participants who met criteria for fluency in English and Spanish at screening and a monolingual arm for primary Spanish speakers. Those in the bilingual group took the WISPI-IV once in English and once in Spanish (order randomly determined) two weeks apart; monolingual Spanish-speakers took the Spanish WISPI-IV twice two weeks apart. Human Subjects IRBs at the Wisconsin and New Mexico research and recruitment sites approved the study.

Participants

Participants were recruited by means of flyers posted in Madison, Wisconsin and Albuquerque, New Mexico and from the bilingual recruiters' social networks. In addition to the university campuses, local medical and mental health clinics that served Spanish-speaking people were targeted. Potential subjects were invited to contact the research staff by telephone to receive an initial description of the study and, if interested, given the choice of coming to the research site for further discussion and consent or to have the study explained over the phone. During this initial contact subjects were screened for their fluency in English and/or Spanish to establish their eligibility for the bilingual or monolingual Spanish arm of the study; 36 (43%) were considered monolingual Spanish, 48 (57%) bilingual Spanish/English. A bilingual research assistant assessed fluency using a standardized interview (procedure adapted from Mitchell, 1987). Subjects were asked: (1) if they could understand the TV or radio news in both English and Spanish; (2) to explain their occupation or field of study in a 2–3 minute monologue in both English and Spanish; (3) to describe the settings (home, work, reading, and watching TV) in which they use English and Spanish (needed to use both languages in at least 2 settings); and (4) to rate their ability to read and understand both English and Spanish [need a rating of at least 6 on a 1 (not at all) to 10 (proficient) scale]. To ensure that participants had an adequate reading ability for this study, we required participants to have completed high school or obtained a high school equivalency certificate.

Test Procedure

Those who expressed interest over the phone were sent the consent form and the initial WISPI-IV test and demographic forms to complete and return by mail; those who met in person with the research staff and signed the consent form and were given the initial test materials to complete in private at the research site. Two weeks later all participants were sent the second WISPI-IV to complete and return by mail. The 48 (57%) participants who were considered bilingual were given one WISPI-IV in English and other in Spanish (order randomly assigned); the remaining 36 (43%)

completed both WISPI-IVs in Spanish. Participants were paid \$20 for participating in the study.

Of the 100 participants who signed consent forms, 91 completed the first test and 84 (35 in Wisconsin and 49 in New Mexico) also completed the retest. Ninety-two percent of the retests were filled out within 9 to 23 days of the first test. The time between tests for the remaining 5 participants ranged from 28 to 85 days ($M = 49$ days). Among the completers, 55 were female and 29 male. Their average age was 34 ($SD = 11.84$, median = 32, range from 18 to 60); 1% had not completed high school, 13% were high school graduates, 21% had some college or technical school, 33% were college or technical school graduates, and 29% had advanced degrees. When asked if they had ever received mental health services, 23 (32%) reported past treatment, 7 reported current treatment, and 3 reported both past and current treatment. The majority of participants were of Mexican heritage (55%) but others were of South American (30%), Central American (7%), or European (4%) heritage.

RESULTS

Mean scores for the WISPI-IV in Spanish ranged from 1.66 (Antisocial) to 3.52 (Obsessive-Compulsive) among bilingual participants (grand $M = 2.39$, median = 2.36) and from 1.42 (Antisocial) to 3.67 (Obsessive-Compulsive) among monolingual participants (grand $M = 2.42$, median = 2.50). There were no significant differences in mean PD scores by test language, by research site, gender or by patient/nonpatient status. The test-retest correlations speak most directly to the aims of this study. Overall, for the 10 PD scales, Pearson r 's ranged from .69 (Schizoid) to .89 (Histrionic and Dependent), median .83. When these correlations for bilingual tests were compared with monolingual tests, there was no significant difference between the correlations for the monolingual and bilingual group [paired $t(9) = 1.91$; $p = .08$].

Alphas for the PD scales provide another indicator of reliability. For the PD scales overall, they ranged from .81 to (Schizoid) to .94 (Antisocial, Avoidant) with a median of .91. When test language was considered, median alphas for the Spanish and English versions were .91 and .92, respectively. A final measure of reliability was the within-subject correlations of the profiles of WISPI-IV means for the first and second tests. For 77 subjects with scores for all of the 10 PD categories, the average Pearson r was .85 ($SD = .24$), median r was .86, and these ranged from $-.55$ to 1.00. Two significant group differences were found. The profile correlations were higher in females ($M = .81$; $SD = 0.16$) than males [$M = .68$; $SD = 0.26$; $t(75) = -2.70$, $p < .01$]. Among those who took the two tests in English and Spanish, profile correlations were higher among those who took the Spanish version first, $t(45) = 2.25$, $p < .05$ (Spanish first $M = .853$; $SD = 0.13$; English first $M = .75$; $SD = 0.17$). A between-site comparison was not significant. Finally, profile r 's were not correlated with the number of days between the two tests or with participant educational levels.

DISCUSSION STUDY 1

In general the translation of the WISPI-IV in this study shows very satisfactory levels of test-retest reliability. Alphas for internal consistency are comparable to what has been reported for translations of other personality measures, including the Temperament and Character Inventory (Gutiérrez et al., 2001), the Personality Assessment Inventory (Rogers, Flores, Ustad, & Sewell, 1995) and for several measures of the Five Factor Model (Aluja et al., 2006; Aluja, García, & García, 2003; Gomà-i-Freixanet, Valero, Puntí, & Zuckerman, 2004; Gutiérrez-Zotes et al., 2008; Hendriks, Hofstee, & DeRaad, 1999; Rodríguez-Fornells, Lorenzo-Seva, & Andrés-Pueyo, 2001; Silva et al., 1994). Even more important is that we demonstrated the theoretical validity of the translation, in that the SASB codes of the back-translation agreed substantially with the SASB codes of the original version.

A limitation of this study was that it did not include any other measures of personality disorder, such as the Structured Clinical Interview for Axis II, or related personality traits, such as the Inventory of Interpersonal Problems, or measures of temperament, such as the NEO-PI. The collaboration with sites in Spain and Argentina reported in study 2 includes some of these measures and will allow us to get further feedback on their experience with the WISPI items in these two countries.

STUDY 2

This study was designed to validate the Spanish translation of the Wisconsin Personality Inventory (WISPI-IV) in relation to two other widely-used personality assessment methods, the self-report Inventory of Interpersonal Problems (IIP-64; Alden et al., 1990; Horowitz et al., 2000) and the Structured Clinical Interview for Axis II interview for personality disorder diagnosis (SCID-II; First et al., 1996). Two sites in Spain (Valencia and Extremadura) and one site in Argentina (Buenos Aires) collected data. The protocol was approved by Investigational Review Boards at each site and at Wisconsin where the data analysis was done.

METHOD

Participants

At each site 50 subjects were recruited, either from the community by posted announcement or from clinics by provider referral or posted announcements. We insured that at least 25 subjects who have received mental health services in the past month were recruited from each site. A total of 149 subjects provided complete data, 50 males and 99 females. The Extremadura site in Spain contributed 49 subjects; Valencia, Spain, 50 and (Buenos Aires), Argentina, 50. With respect to patient status, 77 were nonpatients and 72 were currently in mental health treatment. Their

average age was 34 ($SD = 12.51$), median = 32, range from 18 to 65; 20% had not graduated from high school, 24% were high school graduates, 24% had attended college or technical schools, 22% were college graduates and 12% had advanced degrees.

Procedures

A research staff member explained the study to potential subjects and, if interested, and gave them time to read and sign the consent form. Subjects needed to be 18 years of age or older and a high school graduate (or the equivalent). Subjects were excluded if they had: current active psychosis; presence of an organic brain disorder or severe cognitive impairment; or had electroconvulsive therapy (ECT) within the last three months. Controls were clinic staff (Valencia and Argentina), students or university staff recruited by flyers, e-mail, or class announcements (Extremadura). The clinical sample was recruited by referrals from Mental Health Centers. At the Spanish sites, instruments were given in the following order: IIP, WISPI, SCID-II Questionnaire (except for Valencia), followed by the SCID-II interview. In Argentina, the order was IIP, SCID-II Questionnaire, the SCID-II interview and then the WISPI. The Valencia site reported good inter-rater reliability ($kappa > .80$) for independent ratings of videotaped interviews. Extremadura had only one interviewer who had previous research experience using the Spanish SCID-II however inter-rater reliability was not determined. In Argentina, 4 SCID-II interviewers were trained on the administration of the SCID-II Interview, recorded a practice interview, and then rated the recorded interviews of the other interviewers. The inter-rater reliability was Cronbach = .74. In Valencia and Cáceres, Spain, 93 completed the study and no subjects withdrew. A total of 149 participants provided complete data.

Measures

As described above, the WISPI-IV consists of 214 items (204 items for 11 PDs and a 10-item social desirability scale). Subjects are instructed to "think of your usual self during the past five years or more" and to rate themselves on a scale from 1 (never or not at all true of you) to 10 (always or extremely true of you). Scores are the mean of all items in each PD scale or ipsatized scores. Other measures included a brief demographic questionnaire and a Spanish version of the paper-and-pencil SCID-II Personality Questionnaire to screen for the presence or absence of 118 PD criteria (First et al., 1999). After screening, subjects were interviewed using the Spanish version of the Structured Clinical Interview for DSM-IV Personality Disorders (First et al., 1999). In the protocol for the SCID-II interview, interviewers generally only asked about items that had been endorsed in the Personality Questionnaire, but queried unendorsed items if the participant was one criterion away from meeting a diagnosis or when a par-

ticipant disclosed new information for a previously unendorsed item. In Valencia where the SCID-II Screening Questionnaire was not administered, every item of the SCID-II was assessed. Three PD scores were obtained from the interview, (1) a total score for each PD by summing criteria which had values of 1 for absent or false, 2 for subthreshold, and 3 for threshold or true and (2) a percent score obtained by dividing the total by the total number possible for each scale, and (3) a yes/no score for each diagnostic category met.

The second personality-related measure administered was the Spanish version of the Inventory of Interpersonal Problems (IIP-64; Salazar et al., 2010). The IIP-64 is a circumplex version (Alden et al., 1990; Horowitz et al., 1993) of the original 127-item IIP (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). Items in the IIP assess interpersonal behavior that is hard for (one) to do or one does too much. Each item is rated on a five-point Likert scale ranging from 0 (not at all) to 4 (extremely). Eight trait scores were obtained corresponding to eight points on the circumplex model that has been validated (IIP-64; Alden et al., 1990; Horowitz et al., 1993). The major dimensions of the model are: Domineering versus Non-assertive on the vertical dimension of the model; and Overly Nurturant versus Cold on the horizontal. Points in between are Intrusive, Exploitable, Avoidant, and Vindictive. These octant scores were used to create circular and structural summary statistics (Wright, Pincus, Conroy, & Hilsenroth, 2009).

RESULTS

Comparisons by gender and patient/control status on the WISPI-IV PD scales revealed that, similar to other studies (e.g., Golumb, Fava, Abraham, & Rosenbaum, 1995), males ($N = 50$, $M = 3.10$) scored significantly higher than females ($N = 99$, $M = 2.63$) on Narcissistic, $F(1, 138) = 6.82$, $p < .01$ and Antisocial scales, $F(1, 138) = 11.88$, $p < .001$. Means for males ranged from a low of 1.67 for Antisocial to 3.51 for Paranoid. Means for females ranged from a low of 1.34 for Antisocial to 3.77 for Obsessive-Compulsive. Patients ($N = 72$, $M = 3.01$) scored higher than controls ($N = 77$, $M = 2.58$) on Narcissistic, $F(1, 138) = 6.43$, $p < .01$, Antisocial, $F(1, 138) = 8.42$, $p < .01$, Avoidant, $F(1, 138) = 16.05$, $p < .001$, Dependent, $F(1, 138) = 8.32$, $p < .01$, and Schizoid, $F(1, 138) = 4.82$, $p < .05$ scales. Means for controls ranged from a low of 1.33 for Antisocial to 3.64 for Obsessive-Compulsive. Means for patients ranged from a low of 1.59 for Antisocial to 3.70 for Obsessive-Compulsive. There were no significant differences between sites in Spain and Argentina.

Table 1 shows the standardized Alphas for each of the 10 WISPI-IV scales, the number of SCID-II diagnosed cases, and the Pearson correlations between the WISPI mean scores and the SCID-II percent scores. Alphas are uniformly high, ranging from .82 (Narcissistic and Antisocial) to .93 (Avoidant and Dependent), average .86. Examination of the number of

TABLE 1. Standardized Alpha Coefficients WISPI-II Scales, Number of Diagnosed Cases on SCID-II, and Correlations with SCID-II Percent Scores

PD Scale	WISPI α	SCID-II Cases	r (N = 150) All Subjects	r (N = 78) Controls	r (N = 72) Patients
Paranoid	.84	2	<u>.33**</u>	<u>.33**</u>	<u>.31**</u>
Schizoid	.84	2	<u>.32**</u>	<u>.23*</u>	<u>.38**</u>
Schizotypal	.84	2	<u>.29**</u>	<u>.31**</u>	<u>.26*</u>
Histrionic	.83	2	<u>.33**</u>	<u>.44**</u>	<u>.26*</u>
Narcissistic	.82	4	<u>.33**</u>	.05	<u>.42**</u>
Antisocial	.82	—	.09	.06	.12
Borderline	.88	14	<u>.59**</u>	<u>.52**</u>	<u>.62**</u>
Avoidant	.93	12	<u>.55**</u>	<u>.53**</u>	<u>.53**</u>
Dependent	.93	8	<u>.53**</u>	<u>.46**</u>	<u>.56**</u>
Compulsive	.93	22	<u>.37**</u>	<u>.39**</u>	<u>.35**</u>
Mean	.86		.38	.34	.40

Note. Underline indicates the highest correlation was with the corresponding PD on the other measure.

diagnosed cases of PD on the SCID-II interview shows that the prevalence rates of PDs were relatively low in this sample. The Pearson correlations between the WISPI mean scores and the SCID-II percent scores showed that for controls and patients combined, the r s ranged from .09 (Antisocial) to .59 (Borderline), average .38. For controls the r s ranged from .05 (Narcissistic) to .53 (Avoidant); for patients the range was from .12 (Antisocial) to .62 (Borderline). These correlations are generally lower than those reported in Smith et al. (2003) using the same measures (in English) with an inpatient sample with higher levels of personality pathology in contrast to the relatively lower levels of personality pathology in this sample. As Table 1 also shows, 9 of the 10 of the SCID dimensional scores in patients and controls combined had their highest correlation (as indicated by the underlined r s) with the corresponding WISPI-IV score. In the case of controls, 8 of the 10 r s were the highest; for patients, 9 of the 10 r s were the highest.

Table 2 shows the Cohen effect sizes (d) when the WISPI-IV mean scale scores are compared between the participants who met criteria for a par-

TABLE 2. Cohen Effect Sizes and Measures of Overlap and Nonoverlap Based on WISPI-IV, Mean Scale Scores for PDs with More than Five Diagnosed Cases on the SCID-II

PD Scale	M Dx SCID-II	WISPI z	M w/o Dx SCID-II	d	% Nonoverlap	% of PD above M of Non PD
Borderline	4.66	1.65	2.28	1.38	38.46	92.31
Avoidant	6.38	2.14	2.86	1.99	16.67	100.00
Dependent	4.96	1.90	2.43	1.46	25.00	100.00
Compulsive	4.36	0.89	3.56	0.7	0.00	78.95

Note. M Dx on SCID—II = Mean WISPI—IV scale score for group diagnosed on the SCID—II. WISPI—z = Mean WISPI—IV z score for group diagnosed on the SCID—II. d = Cohen's Effect Size. %Nonoverlap = percentage of nonoverlap between the SCID—II diagnosed PD and not diagnosed PD groups. % of PD Above M of Non PD = percentage of patients with a particular SCID—II diagnosed PD scoring above the mean of those patients not diagnosed on the SCID—II with that particular PD.

ticular PD on the SCID-II and those who did not. This statistic describes the distance between the means of the disordered and nondisordered groups in pooled standard deviation units (Cohen, 1988; Hsu, 2002). This table also presents two measures of the percentage of overlap and non-overlap between the two distributions of WISPI-IV mean scores. Here, the WISPI-IV distinguished between those who were diagnosed with a particular PD on the SCID-II and those who were not, as demonstrated by the large effect sizes ($d \geq .80$; Cohen, 1988) for three of the four PDs that had more than five diagnosed cases on the SCID-II. These compare favorably to the effect sizes reported earlier on the English versions of the WISPI-IV (Smith et al., 2003).

Another way of examining convergence between the two PD measures is to examine the correspondence between the profiles of the 10 PD scales on each measure for each participant. Figure 2 shows the profiles of mean percentage of endorsement on the Spanish versions of the WISPI-IV and SCID-II PD scales. This within-subject procedure considers all of the PD dimensions at once and derives an index of overall congruence. To make the WISPI-IV scores metrically commensurate with the SCID percentage scores, we calculated WISPI percentages from the sum of all item endorsements within a scale, divided by the total score possible for that scale. After reversing the data matrix (participants become the columns and the 10 PDs the rows), each participant's profile of percentage dimensional scores on the SCID-II was correlated with his or her corresponding WISPI-IV profile of percentage scores. The resulting statistic (a within-subjects Pearson product moment correlation) represents the congruence of the two measures across all PDs per subject. The average r (after r to z conver-

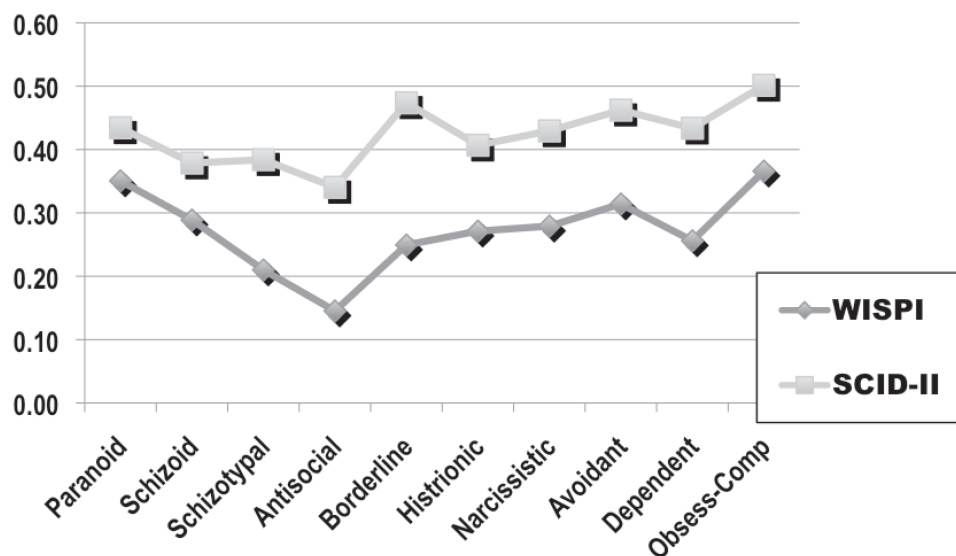


FIGURE 2. Profiles of mean percentage of endorsement on the Spanish versions of the WISPI-IV and SCID-II PD scales.

sion) between the congruence scores across all participants was .57 with a range of $-.51$ to $.95$ (median = $.56$). This is consistent with earlier studies of the English versions of the WISPI and SCID-II, which found mean profile correlation of $.53$ (Barber & Morse, 1994) and $.61$ (Smith et al., 2003).

Table 3 shows the IIP-64 octant scores for all the patients and controls as well as for those who met criteria for a PD on the WISPI-IV (as determined by a z -score ≥ 1.96). We only show here those PDs that had 5 or more instances of a WISPI z -score ≥ 1.96 . We found in our earlier work (Smith et al., 2003) that using these z -scores was a reasonable way to establish diagnosis on the WISPI (i.e., the mean correct classification in the 2003 study was $.71$ in comparison to SCID-II diagnoses). We used circular statistics and structural summary scores as was recently suggested by Wright and colleagues (2009) to take advantage of the richness of circumplex description of the IIP-64 (for a detailed description of these statistics please see Wright et al., 2009). The structure of the predicted sinusoidal (i.e., cosine curve) pattern of responses to the IIP-64 can be summarized using the following statistics: elevation, amplitude, and angular displacement (see Figures 3 and 4). When examining the IIP-64 in this manner, elevation (or the mean response level across scales) can be interpreted as general level of interpersonal distress. Amplitude, is equivalent to vector length, and is indicative of the degree of differentiation among the interpersonal aspects in the group profile. Angular displacement, on the other hand, is indicative of the core interpersonal theme of the profile. R^2 is a goodness-of-fit statistic that summarizes the degree to which the profile can be captured by the three structural parameters just mentioned. Wright and colleagues describe this as a measure of "interpersonal prototypicality" although others have described it as a measure of profile complexity (Gurtman & Balakrishnan, 1998). Values of R^2 below $.70$ in this context suggest that the interpersonal problem profile of responses cannot be adequately summarized using circular statistics and therefore the summary angle should be interpreted with caution.

TABLE 3. Interpersonal Comparison Using Controls, Patients, and Those PDs with 5 or More Cases of a WISPI-IV z -score > 1.96

	Controls	Patients	PAR	SZD	SZT	HST	BOR	AVD	DEP	OCPD
Circular Statistics										
Mean	79.45	177.80	134.47	135.01	170.51	63.05	161.05	203.87	214.75	29.14
Variance	72.94	73.19	64.08	70.68	63.02	82.01	62.36	57.97	82.51	72.93
95% CI High	95.64	194.71	166.90	166.80	225.75	123.81	201.79	236.67	259.61	68.78
95% CI Low	63.27	160.89	102.04	103.23	115.28	2.30	120.31	171.07	169.90	-10.51
Structural Summary										
Displacement	89.99	179.04	147.56	169.16	164.55	86.36	163.42	195.19	201.50	62.39
Amplitude	0.24	0.45	0.49	0.32	0.46	0.31	0.49	0.67	0.22	0.26
Elevation	0.38	0.52	0.20	0.24	0.13	0.35	0.56	0.48	0.69	0.16
R^2	<u>0.71</u>	<u>0.80</u>	<u>0.93</u>	<u>0.71</u>	0.61	0.39	<u>0.76</u>	<u>0.94</u>	0.46	<u>0.89</u>
N	78	72	15	19	5	7	9	12	13	13

Note. PAR—Paranoid; SZD—Schizoid; SZT—Schizotypal; HST—Histrionic; BOR—Borderline; AVD—Avoidant; DEP—Dependent; OCPD—Obsessive-Compulsive.

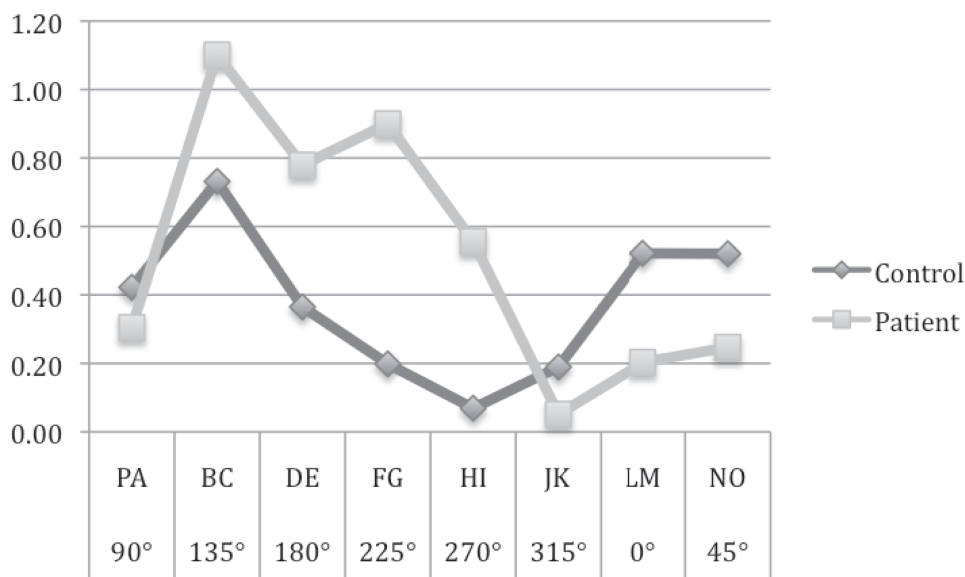


FIGURE 3. Interpersonal comparison of patients and controls using IIP octant z-scores. PA—Domineering; BC—Vindictive; DE—Cold; FG—Avoidant; HI—Nonassertive; JK—Exploitable; LM—Overly-Nurturant; NO—Intrusive.

When calculating these circumplex statistics, the standard procedure is to use normative reference sample as a basis for calculating the group IIP profiles (e.g., Wright et al., 2009). We used two different reference samples: a nonpatient sample to examine the IIP profiles of all the control subjects in our study in comparison to all patients; and a patient reference sample to examine the IIP profiles for those who met PD criteria on the WISPI-IV. Examining patients and controls in relation to a normal reference group allowed us examine overall differences between these two groups. Later we used a clinical reference group to examine within group differences among those individual who met our criteria for a PD on the WISPI-IV. PDs are generally differentiated from normal functioning by higher levels of coldness/hostility on the IIP therefore the clinical reference group accounts for those common characteristics allowing for a more targeted, finer grained analysis of the specific interpersonal profiles among the patients meeting various PD criteria.

Since the IIP-64 was just recently translated into Spanish we used the mean IIP-64 octant scores of a nonpatient group collected for another study, which examined the validity of the Spanish version of IIP-64 and its use in screening personality disorders in clinical practice (Salazar et al., 2010). When we examined our groups of patients and controls across the three sites, the nonpatient reference sample for this analysis consisted of 66 subjects from a nonpsychiatric population recruited from professional staff with neither psychiatric diagnoses nor personality disorders. This is not a representative sample of the general population so they cannot be

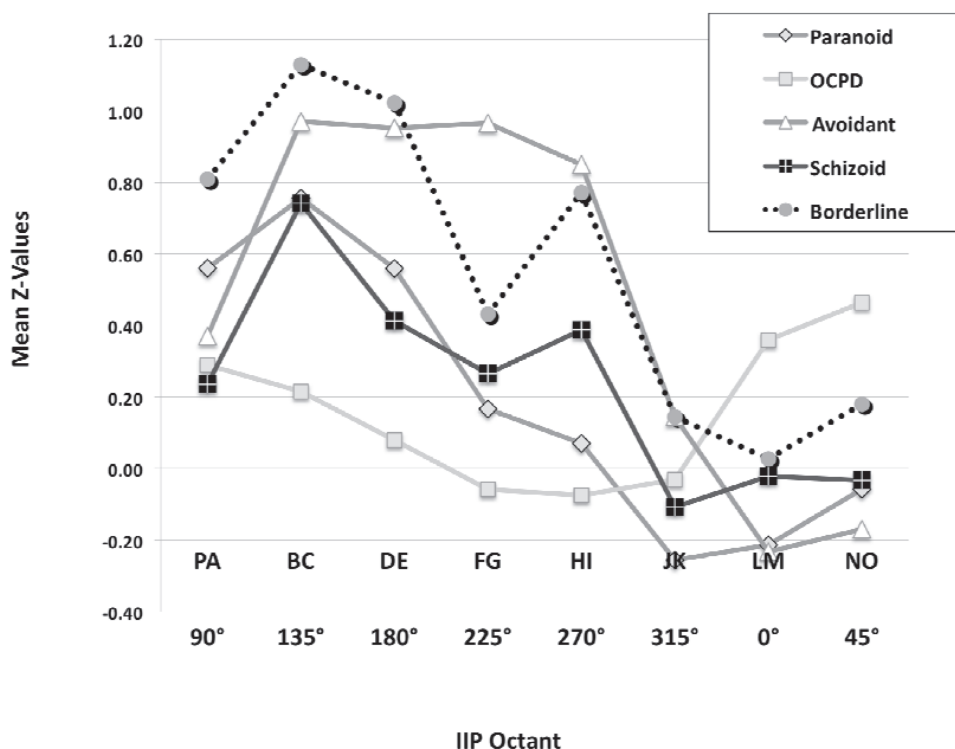


FIGURE 4. IIP z-scores for personality disorders with more than 5 cases of a WISPI-IV z-score >1.96 and a R^2 value $>.70$. PA—Domineering; BC—Vindictive; DE—Cold; FG—Avoidant; HI—Nonassertive; JK—Exploitable; LM—Overly-Nurturant; NO—Intrusive.

considered truly normative scores. Subjects were 70% female with a mean age of 36 years living in Valencia, Spain. Given that $\frac{2}{3}$ of our study sample were Spanish; 66% were female, and the mean age was 34 years, this control group seemed to be a reasonable if not an ideal choice for the normal reference sample we needed to calculate these statistics. Nevertheless, this limitation should be kept in mind in reviewing our results. The octant means and standard deviations for this normative control sample were: LM = 4.7(5.0); NO = 7.9(4.2); PA = 5.9(3.8); BC = 4.9(3.7); DE = 4.8(3.8); FG = 8.4(5.5); HI = 11.9(5.3); JK = 12.8(5.4). Similarly, when we examined the PD groups of patients, the normative clinical sample we used came from the same study (Salazar et al., 2010) which consisted of 190 patients (60 males, 130 females) attended consecutively at public health services network and private practice in Valencia, Spain. The octant means and standard deviations for this normative clinical sample were: LM = 7.5(6.1); NO = 10.4(5.8); PA = 7.4(5.0); BC = 7.9(5.6); DE = 7.3(5.9); FG = 13.3(7.8); HI = 15.3(7.1); JK = 14.8(6.0).

Table 3 and Figure 3 shows that patients and controls can be differentiated based on their interpersonal problem profiles and that these two

groups have reasonable goodness of fit (based on their R^2 scores). The angular displacement for patients was in the DE (Cold) octant while controls was in the PA (Domineering) octant. Controls had less interpersonal distress and fewer interpersonal problems based on the lower elevation and amplitude of their profile.

Table 3 also shows the interpersonal comparison of the profiles of those groups of individuals whose z -scores on the WISPI-IV PD scales were ≥ 1.96 . Figure 4 shows a subset of 5 out of 7 PD, since we graphed only the PD profiles with a R^2 value $>.70$. We set the criterion that a curve with an R^2 value $<.70$ was not adequately summarized by its structural components and therefore a sinusoidal curve was not a strong representation of the underlying data (see Wright et al., 2009 for a detailed discussion of this issue). The angular displacement for Paranoid PD ($N = 15$) was in the BC (Vindictive) octant; Schizoid ($N = 19$) and Borderline ($N = 9$) PD were in the DE (Cold) octant; Avoidant PD ($N = 12$) was in the FG (Socially Avoidant) octant; while OCPD ($N = 13$) was in the NO (Intrusive) octant. Borderline PD shows an interesting split in their profile in that they are reporting being both Vindictive and Nonassertive which is consistent with the interpersonal instability characteristic of this PD. Avoidant PD showed more interpersonal distress than Paranoid and Obsessive-Compulsive PDs based on the higher elevation of their profiles. Borderline PD shows the greatest level of interpersonal distress compared to the other PDs shown in Figure 4. These PD were well differentiated based on their IIP octant patterns. In addition the patterns for Avoidant and Paranoid PD were quite similar to those reported using a Norwegian version of the IIP-64 (referred to as the IIP-C) and the SCID-II (Monsen, Hagtvet, Havik, & Eilertsen, 2006).

DISCUSSION STUDY 2

This study demonstrates that the Spanish version of the WISPI-IV has high internal consistency and good convergent validity with the Spanish versions of the SCID-II interview and the IIP-64, and performed similarly to their English language counterparts. Examination of the correlations between SCID-II dimensional scores and the WISPI-IV scales demonstrated good convergence (mean r between corresponding PD scales = .38). Profile analysis provided the most compelling evidence for convergence (mean $r = .57$) demonstrating that individuals show a similar pattern of responses across the 10 PD scales on the 2 measures. The results of this profile analysis were similar to those previously reported between these two measures ($r = .61$ in Smith et al., 2003 and .53 in Barber & Morse, 1994). The Spanish versions of the WISPI and the IIP showed a similar pattern of correlations to those reported for Norwegian translations of the SCID-II and the IIP, again demonstrating similar properties to other translations of these instruments (Leising, Rehbein, & Eckardt, 2009). The pat-

tern of IIP-64 octant scores for some PDs obtained with this Spanish version of the WISPI was quite similar to those obtained with a Norwegian sample using the SCID-II to diagnose PD (Monsen et al., 2006).

There are a number of limitations to our study. While the training of interviewers was adequate, the inter-rater reliability for the SCID-II interview was not uniformly measured across sites. Ideally we should have had several videotaped SCID-II interviews conducted in Spanish that all sites and interviewers could have rated to establish that reliability across sites using the same procedures and statistics. However, given that there were no site differences on any of the measures, this lack of procedural consistency did not seem to have a large effect on our results. The most serious limitation was the relatively low rates of PD pathology in this sample. This may have attenuated some of our analyses although the SCID-II to WISPI-IV profile concordance suggests that both measures were finding similar levels of pathology. In addition the octant patterns for various PDs on the IIP-64 were similar to that reported in another study (Monsen et al., 2006).

GENERAL DISCUSSION

In summary, Study 1 demonstrated that this Spanish version of the WISPI-IV demonstrated good internal consistency, test-retest reliability, and theoretical validity in that the SASB codes of the translation agreed substantially with the SASB codes of the original version. Study 2 showed that this Spanish version WISPI-IV had good convergent validity with the SCID-II interview and the IIP-64. Although more research is needed, both studies suggest evidence of measurement equivalence between the Spanish and English version of the WISPI-IV.

Our effort to translate the WISPI-IV is consistent with recommendations for the need to develop culturally-sensitive assessments for psychological disorders in ethnic minorities (Bernal, 2006; Hall, 2001; Miranda, Nakamura, & Bernal, 2003), making it possible to gather and analyze comparable data to characterize issues that are specific to different ethnic groups (López & Guranaccia, 2000). It is essential to develop an understanding that illnesses in the diagnostic nomenclature are embedded in the cultural origins and economic status of different ethnic groups and subgroups. Thus development of assessment procedures such as the WISPI-IV by which various ethnic groups can be compared and contrasted is needed. With respect to Hispanic speakers, we recognize the geographic and cultural variety of their origins as well the stresses associated with immigration and acculturation processes.

Our goal in translating the WISPI-IV was to extend this measure to the 417 million Spanish-speakers in the world (Lewis, 2009), and in language that would be understood by different Hispanic subgroups. However, comments from some participants, particularly in New Mexico, indicated difficulties understanding particular items, or while understanding the intent of some items, the opinion was that the translation could have been

improved. This suggests that our process of review and back-translation should be extended to representatives of the larger Hispanic groups, including Mexican, Puerto Rican, Cuban, Central and Latin American, and European so that translators from these areas could review, discuss, and possibly resolve language issues. In some cases, regional Spanish words or phrases might be inserted in parentheses. Each site had different suggestions for changes in item wording. This reflects the difficulty with creating a "neutral" Spanish version of any assessment given regional differences.

Our hope is that this Spanish translation of the WISPI-IV, which is consistent with the interpersonal theoretical roots of the original English version, will begin to assist others in addressing the gaps in personality disorder research, clinical assessment, and treatment planning that accounts for personality pathology among Spanish-speaking populations so that their mental health needs can be met with effective and culturally sensitive service delivery systems.

REFERENCES

- Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the inventory of interpersonal problems. *Journal of Personality Assessment*, 55, 521-536.
- Aluja, A., García, Ó., & García, L. F. (2003). Psychometric properties of the Zukerman-Kuhlman personality questionnaire (ZKPQ-III-R): A study of a shortened form. *Personality and Individual Differences*, 34, 1083-1097.
- Aluja, A., Rossier, J., García, L. F., Angleitner, A., Kuhlman, M., & Zuckerman, M. (2006). A cross-cultural shortened form of the ZKPQ (ZKPQ-50-cc) adapted to English, French, German, and Spanish languages. *Personality and Individual Differences*, 41, 619-628.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- Barber, J. P., & Morse, J. Q. (1994). Validation of the Wisconsin personality disorders inventory with the SCID-II and PDE. *Journal of Personality Disorders*, 8, 307-319.
- Benjamin, L. S. (1974). Structural analysis of social behavior. *Psychological Review*, 81, 392-425.
- Benjamin, L. S. (1984). Principles of prediction using structural analysis of social behavior (SASB). In R. A. Zucker, J. Aronoff, & A. J. Rabin (Eds.), *Personality and prediction of behavior* (pp. 121-174). New York: Academic Press.
- Benjamin, L. S. (1987). Use of the SASB dimensional model to develop treatment plans for personality disorders: I. Narcissism. *Journal of Personality Disorders*, 1, 43-70.
- Benjamin, L. S. (1993). *Interpersonal diagnosis and treatment of personality disorders*. New York: Guilford Press.
- Benjamin, L. S. (1996). *Interpersonal diagnosis and treatment of personality disorders* (2nd ed.). New York: Guilford Press.
- Bernal, G. (2006). Intervention development and cultural adaptation research with diverse families. *Family Process*, 45, 143-151.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Critchfield, K. L. (2002). An empirical analysis and initial taxonomy of patient reported relational patterns based on Interpersonal Case Formulation. Dissertation. University of Utah.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J.B.W., & Benjamin, L. S. (1996). *Structured clinical interview for the DSM-IV axis II personality disorder*

- ders (SCID-II). New York: New York State Psychiatric Institute.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J.B.W., & Benjamin, L. S. (1999). Guía del usuario para la entrevista clínica estructurada para los trastornos de la personalidad del eje II del DSM-IV®: SCID-II. 1ª ed. Barcelona, Spain: Masson SA.
- Golumb, M., Fava, M., Abraham, M., & Rosenbaum, J. F. (1995). Gender differences in personality disorders, *American Journal of Psychiatry*, 152, 579–582.
- Gomà-i-Freixanet, M., Valero, S., Puntí, J., & Zuckerman, M. (2004). Psychometric properties of the Zuckerman—Kuhlman personality questionnaire in a Spanish sample. *European Journal of Psychological Assessment*, 20, 134–146.
- Greenwald, H. J., & Satow, Y. (1970). A short social desirability scale. *Psychological Reports*, 27, 760–733.
- Gurtman, M. B., & Balakrishnan, J. D. (1998). Circular measurement redux: The analysis and interpretation of interpersonal circle profiles. *Clinical Psychology: Science and Practice*, 5, 344–360.
- Gutiérrez, F., Torrens, M., Boget, T., Martín-Santos, R., Sangorrín, J., Perez, G., & Salamero, M. (2001). Psychometric properties of the temperament and character inventory (TCI) questionnaire in a Spanish psychiatric population. *Acta Psychiatrica Scandinavica*, 103, 143–147.
- Gutiérrez-Zotes, J. A., Gutiérrez, F., Valero, J., Gallego, E., Baillés, E., Torres, X., Labad, A., & Livesley, W. J. (2008). Structure of personality pathology in normal and clinical samples: Spanish validation of the DAPP-BQ. *Journal of Personality Disorders*, 22, 389–404.
- Hall, G.C.N. (2001). Psychotherapy research with ethnic minorities?: Empirical, ethical, and conceptual issues. *Journal of Consulting and Clinical Psychology*, 69, 501–510.
- Harder, S., & Henry, W. P. (1998). Reliability of interpersonal case formulation in a psychotic population. Unpublished manuscript, University of Copenhagen.
- Hendriks, A.A.J., Hofstee, W.K.B., & De Raad, B. (1999). The five-factor personality inventory (FPI). *Personality and Individual Differences*, 27, 307–325.
- Horowitz, L. M., Alden, L. E., Wiggins, J. S., & Pincus, A. L. (2000). *Inventory of interpersonal problems manual*. Odessa, FL: The Psychological Corporation.
- Horowitz, L. M., Rosenberg, S. E., & Bartholomew, K. (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. *Journal of Consulting and Clinical Psychology*, 61, 549–560.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885–892.
- Hsu, L. M. (2002). Diagnostic validity statistics and the MCMI-III. *Psychological Assessment*, 14, 410–422.
- Humes, D. L., & Humphrey, L. L. (1994). A multimethod analysis of families with a polydrug-dependent or normal adolescent daughter. *Journal of Abnormal Psychology*, 103, 676–685.
- Klein, M. H., & Benjamin, L. S. (1996). *The Wisconsin personality disorders inventory-IV*. (Available from Dr. M. H. Klein, Dept. of Psychiatry, Wisconsin Psychiatric Institute and Clinics, 6001 Research Park Blvd., Madison, WI 53719-1179.)
- Klein, M. H., Benjamin, L. S., Rosenfeld, R., Treece, C., Husted, J., & Greist, J. H. (1993). *Journal of Personality Disorders*, 7, 285–303.
- Leising, D., Rehbein, D., & Eckardt, J. (2009). The inventory of interpersonal problems (IIP-64) as a screening measure for avoidant personality disorder. *European Journal of Psychological Assessment*, 25, 16–22.
- Lewis, M. P. (Ed.). (2009). *Ethnologue: Languages of the world* (16th ed.). Dallas, Tex.: SIL International.
- López, S. R., & Guarnaccia, P. J. (2000). Cultural psychopathology: Uncovering the social world of mental illness. *Annual Review of Psychology*, 51, 571–598.
- Matías-Carrelo, L. E., Chávez, L. M., Negrón, G., Canino, G., Aguilar-Gaxiola, S., & Shoppe, S. (2003). The Spanish translation and cultural adaptation of five mental health outcome measures. *Culture, Medicine and Psychiatry*, 27, 291–313.
- Monsen, J. T., Hagtvet, K. A., Havik, O. E., & Eilertsen, D. E. (2006). Circumplex structure and personality disorder correlates of the interpersonal problems model (IIP-C): Construct validity

- and clinical implications. *Psychological Assessment*, 18, 165–173.
- Miranda, J., Nakamura, R., & Bernal, G. (2003). Including ethnic minorities in mental health intervention research: A practical approach to a long-standing problem. *Culture, Medicine and Psychiatry*, 27, 467–486.
- Mitchell, R. (1987). Assessing the language skills of bilingual primary pupils in SCRE spotlight 6. Edinburgh: Scottish Council for Research in Education. (www.scre.ac.uk/spotlight/spotlight6.html) Retrieved April 14, 2010.
- Robles-Garcia, L. C., Nabel, T., & Agraz, F. P. (2003). Study of translation and reliability of the Wisconsin personality disorders inventory (WISPI-IV). *Actas españolas de psiquiatría*, 31, 133–137.
- Rodriguez-Fornells, A., Lorenzo-Seva, U., & Andres-Pueyo, A. (2001). Psychometric properties of the Spanish adaptation of the Five Factor Personality Inventory. *European Journal of Psychological Assessment*, 17, 145–153.
- Rogers, R., Flores, J., Ustad, K., & Sewell, K. W. (1995). Initial validation of the personality assessment inventory—Spanish version with clients from Mexican American communities. (1995). *Journal of Personality Assessment*, 64, 340–348.
- Salazar, J., Martí, V., Soriano, S., Beltran, M. & Adam, A. (2010). Validity of the Spanish version of the inventory of interpersonal problems and its use for screening personality disorders in clinical practice. *Journal of Personality Disorders*, 24, 499–515.
- Silva, F., Avia, D., Sanz, J., Martínez-Arias, R., Graña, J. L., & Sánchez-Bernados, L. (1994). The five-factor model—I. Contributions to the structure of the NEO-PI. *Personality and Individual Differences*, 17, 741–753.
- Smith, T. L., Klein, M. H., & Benjamin, L. S. (2003). Validation of the Wisconsin personality disorders inventory-IV with the SCID-II. *Journal of Personality Disorders*, 17, 173–187.
- Wright, A. G., Pincus, A. L., Conroy, D. E., & Hilsenroth, M. J. (2009). Integrating methods to optimize circumplex description and comparison of groups. *Journal of Personality Assessment*, 91, 311–322.