



Theses and Dissertations

2019-07-01

Personality Inventory DSM-5: A Spanish Translation for Hispanics in the United States

Jessica Abigail Carmona
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

BYU ScholarsArchive Citation

Carmona, Jessica Abigail, "Personality Inventory DSM-5: A Spanish Translation for Hispanics in the United States" (2019). *Theses and Dissertations*. 8567.
<https://scholarsarchive.byu.edu/etd/8567>

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Personality Inventory DSM-5: A Spanish Translation for Hispanics
in the United States

Jessica Abigail Carmona

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Patrick R. Steffen, Chair
Gary Michael Burlingame
Scott Ronald Braithwaite
Bruce Noble Carpenter
Mikle Don South

Department of Psychology
Brigham Young University

Copyright © 2019 Jessica Abigail Carmona

All Rights Reserved

ABSTRACT

Personality Inventory DSM-5: A Spanish Translation for Hispanics in the United States

Jessica Abigail Carmona
Department of Psychology, BYU
Doctor of Philosophy

The Personality Inventory DSM-5 (PID-5) was created to measure personality pathology and help in the development of a dimensional conceptualization of personality disorders (Krueger, Derringer, Markon, Watson, & Skodol, 2012). It measures five maladaptive personality traits: Negative Affect, Detachment, Antagonism, Disinhibition and Psychoticism. The PID-5 has also garnered significant support for its hierarchical structure, five-factor structure across samples and translations, and its ability to predict variance in internalizing and externalizing disorders (Krueger & Markon, 2014). The current study builds on this literature by translating the PID-5 into Spanish spoken in Latin America and testing the replicability of the five-factor structure, reliability, and validity of the PID-5-Sp facets in a Hispanic sample. Using Mechanical Turk, 305 participants completed the PID-5-Spanish, Patient Health Questionnaire (PHQ-9), Generalized Anxiety Disorder (GAD7), Aggression Questionnaire-Revised (AQ-R) and the Big Five Inventory (BFI). EFA suggested a three-factor structure that resulted in two small factors that were conceptually similar to Antagonism and Detachment and one large global general distress factor. CFA results indicated that a five-factor solution had a poor fit for the current sample. Reliability was acceptable for most facets ($\alpha = .60-.95$, $M = .85$). In general, PID-5-Sp domains showed moderate to strong correlations with theoretically congruent normative traits, with exception of Psychoticism, which was not significantly correlated with Openness to Experience ($r = -.08$, $p = > .05$). As expected, Detachment and Negative Affect predicted GAD-7 and PHQ-9 scores. Aggression scores were predicted by Negative Affect, Antagonism and Disinhibition. Overall, the PID-5-Sp partially replicated previous validity and reliability findings. However, future research is needed to further test the five-factor structure and its replicability in non-Western samples.

Keywords: PID-5, Spanish, Hispanics, Five-Factor Model, personality assessment

ACKNOWLEDGMENTS

This dissertation would not have been possible without the constant support of Bruce N. Carpenter, PhD, who first gave me the opportunity to join him as a graduate mentee in January 2013. I am very grateful that he continued to offer his unwavering support even after retirement. Thank you for always being willing to go the extra mile. The support kept me hopeful and focused on the end goal. Dr. Patrick Steffen, thank you for stepping in and helping me through the finish line. My committee, of course, also deserves bouts of gratitude for the learning opportunity offered throughout the development and completion of this project.

I also want to thank the clinicians who were willing to help me refine the translation of the PID-5-Sp; without your efforts the translation would not have been made possible. Liliana Carmona Garces deserves my deep gratitude for her role in the translation and editing of the final version of the PID-5-Sp. The Spanish translation would make a lot less sense without the many hours that she invested in its grammatical composition.

Además, quiero agradecerle a mi familia por su apoyo y amor incondicional. Sin ustedes no tendría la confianza de ser la persona que soy. Gracias, mamá y papá, fue su validación de la importancia de la educación que me permitió convertirme en la primera integrante de la familia en obtener un doctorado. Este logro es tanto mío como de ustedes. Claudia y Lili, gracias por el apoyo emocional y su inquebrantable fe en mis habilidades, espero algún día poder hacerle justicia a sus expectativas.

Por último, le quiero dedicar este logro a mi abuelita, Gabina Hernandez Perez, quien se nos adelantó en el camino de la vida. Gracias por interceder ante Dios por mí, es gracias a sus plegarias que Dios nunca me ha abandonado.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	vi
LIST OF FIGURES	vii
Introduction.....	1
Classification.....	3
Assessment of Personality.....	5
FFM and Psychopathology.....	7
Assessment of Personality Psychopathology	9
The PID-5.....	11
Model and Structure	11
The Measure.....	12
PID-5 and other FFM Measures.....	14
Five-Factor Structure.....	15
Psychopathology Correlates.....	18
The PID-5 as a Diagnostic Tool.....	19
The Big Five and Culture.....	22
PID-5 in Spanish	24
Purpose.....	26
Method	27
Translation.....	27
Data Collection.....	28
Figure 1	32
Participants	33
Measures.....	34
Personality Inventory for DSM-5, Spanish version (PID-5-Sp).....	34
Spanish Big Five Inventory (BFI).....	34
Aggression Questionnaire-Refined (AG-R).....	35
Patient Health Questionnaire-9 (PHQ-9).....	36

Generalized Anxiety Disorder-7 (GAD-7).....	36
Statistical Analysis	36
Exploratory factor analysis.....	37
Confirmatory factor analysis.....	38
Results.....	39
Descriptives and Distributions	39
Table 1.....	40
Reliability.....	41
Factor Structure.....	41
Table 2.....	43
Construct Validity	44
Table 3.....	45
Discussion.....	45
Table 4.....	46
Limitations	55
Future Research and Conclusions	57
References.....	58

LIST OF TABLES

Table 1 Descriptive Statistics for PID-5-Sp Facets, Domains and Outcome Measures	40
Table 2 PID-5-Sp Exploratory Factor Analysis.....	43
Table 3 Two-Tailed Pearson’s Correlations of PID-5-SP Domains and BFI Traits.....	45
Table 4 Regressions between PID-5 Domains and Outcomes.....	46

LIST OF FIGURES

Figure 1 Flow of Participants Included in Final Analyses..... 32

Personality Inventory DSM-5: A Spanish Translation for Hispanics in the United States

Introduction

The assessment of personality psychopathology has a long history in academic and applied research, but at times it has suffered from stagnation, not so much because of a lack of new ideas and approaches, but because of entrenched inertia among users and comfort with the well know, if at times limited in scope, categorical models. Also, developing appropriate normative data for instruments measuring the constructs of interest is a massive undertaking and establishing a strong research basis can takes years. Thus, when a new approach with considerable promise emerges, it is worthy of our effort to aid in the development of norms, building a research base, removing barriers to accessibility, and exposing practitioners to its benefits. The Personality Inventory for DSM-5 (PID-5) presents such an opportunity and is the focus of this research study. The PID-5 is a hierarchically organized set of item clusters (each representing a facet of dysfunctional personality) closely tied to the “PSY-5,” a version of the Five Factor Model (FFM) for dysfunctional personality attributes (American Psychiatric Association, 2013). When contrasting it to the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Tellegen, & Kreammer, 1989), a widely accepted measure of psychopathology, the PID-5 is model-driven, is less than half the length of the MMPI, provides information about both specific and broad attributes (facets and dimensions), has no item overlap across scales, is relatively straight-forward in interpretation, parallels data needed for DSM-5 diagnoses under the new dimensional DSM-5 Personality model, will be familiar to practitioners who use the NEO-PI-3 (the most widely used measure of normative personality; Costa & McCrae, 1992), and is free to use. Also, despite its recent publication in 2013, it has already

garnered support from hundreds of research articles, which is a testament to the curiosity of the field regarding its properties and future utility (Krueger & Markon, 2014).

However, even good assessment instruments can flounder in the presence of barriers that have traditionally existed, such as cost in time, cost in money, along with associated copyright restrictions that complicate independent research, norms not representative of the population of interest, complex interpretation and language barriers(most measures are in English only, and developed in Western countries). The PID-5 removes some of these barriers by being shorter than other personality inventories and being free to use (APA, 2013). The current project seeks to remove a language barrier by (a) translating the measure to Spanish as mostly spoken by native Latin America Spanish speakers, (b) providing initial data towards the development of norms for Hispanics living in the United States, and (c) providing validity data assessing the PID-5's Latin America Spanish translation ability to function in parallel to the original English version and its consistency with a FFM of personality.

This last theoretical component is particularly relevant because a significant percentage of the importance of the PID-5 comes from its foundational role in the discipline's journey towards the development of dimensional conceptualizations of psychopathology. The PID-5 is part of a larger model that is supported by many renowned researchers and the APA (2013); as such it has the potential to become an integral part of how psychologists talk and understand personality. Therefore, we'll first focus on providing context about the maladaptive model of personality proposed in the DSM-5, then move on to how the PID-5 fits within this model, culminating with the relevance of having a Latin American Spanish translation that is specific for Hispanics in the American continent.

Classification

For centuries classification systems have been integral to science and the meaningful integration of knowledge. In biology, the classification of living organisms into a taxonomy has provided scientists with a robust hierarchical structure that organizes organisms with increasing levels of specificity. This classification system has allowed biologists and zoologists to build a coherent body of literature regarding the origins of life, evolution, and the species that inhabit Earth. As science has evolved, technological advances have made information about genomes available and although the current taxonomy has not fully incorporated this new information, proposals exist towards its integration (Godfray, 2002). The desire to keep this taxonomy alive in the natural sciences is not only the product of nostalgia, it is an attempt to maintain findings that have accumulated over the years together in a coherent narrative that can be understood by scientists across the world.

Given the utility and cohesiveness of taxonomies, psychologists have tried for years to replicate similar models in their study of human behavior. However, these have been met with mixed results. Classification systems in the natural sciences tend to be reliable because most of their subjects of study have physical characteristics that can be objectively observed and described. In contrast, the subjects of study in psychology are more abstract and definitions depend heavily on theoretical conceptualizations. It is not uncommon for the same event to be described differently depending on theoretical orientations (Fournier, Di Domenico, Weststrate, Quitasol, & Dong, 2015). This diversity in theories maintains the field actively engaged in research, but it also puts restraints in its ability to test whole theories efficiently, which means that advancement is slow and sometimes cluttered.

One area in which this has been a prevalent issue is in the field of personality. Since the Hellenistic period, the desire to understand behavior and drives has consumed humans (Robinson, 1995), and with the contributions of philosophers, biologists, physicists, and later social scientist this field of inquiry has grown to become what is now known as the study of personality. Many theories have been postulated, but a consensus has remained elusive. Some theorists focus on interpersonal interactions and how these aid in the development of the self (Sullivan, 1953). Others, postulate that personality has a hierarchy of stable dimensional traits that can be identified through the study of language and behavioral patterns (Digman, 1990). The later theory has received empirical support sporadically throughout the years, mainly due to fluctuating levels of interest and lack of communication between researchers in the field. Although support also exists for more situationally based theories of personality a full review of that literature would be beyond the scope of the current research study (Fournier et al., 2015).

Currently, the most widely accepted trait theory of personality is the “Big Five,” or the FFM of personality. This model divides personality into five stable traits that are assumed to be present in all humans and across cultures to varying degrees. The five traits are commonly named: conscientiousness, neuroticism, extraversion, agreeableness, and openness to experience (Costa & McCrae, 1992). The “Big Five” are often attributed to the work of Costa and McCrae (1992), but in reality, this is a model that had been in the making since the early 1920’s. Digman (1990) attributes the roots of the model to the theoretical work of William McDougall. McDougall (1926) argued that binary models such as those proposed by Jung that divided individuals into introverts and extroverts were insufficient to account for personality and human idiosyncrasies. He believed a more inclusive models of personality were needed. McDougall (1926) proposed that a comprehensive personality model would require the integration of at least

five factors—not to be confused with our current understanding of factors, rather he meant five different components—with varying degrees of presentation and relevance depending of the individual. The proposed factors or classes were intellect, disposition, temper, temperament, and character. Despite his proposal, research remained stagnant and he later lamented about the lack of specificity and cohesiveness in the personality literature in an editorial introducing the *Journal of Personality* (McDougall, 1932).

Assessment of Personality

The foundational research that lead to the development of the “Big Five,” began in linguistics, with research conducted based on a “lexical hypothesis” (Digman, 1990). The rationale assumed humans could encode socially meaningful characteristics into their natural language. It was expected that themes that were salient and differentiating in the human experience would emerge as single word characteristics, descriptive of human personality. Therefore, research efforts focused on discovering personality characteristics in the structure of language (John, Angleitner, & Ostendorf, 1988). Initially this research was conducted in Germany, but by the mid 1930’s researchers from the United States started taking note. Perhaps one of the most influential studies published in the topic was the work of Allport and Odbert (1936). They created a list of 17,953 words that they considered personality relevant terms which were then divided into four categories: personal traits, temporary states, social evaluations, and miscellaneous (words with metaphorical and/or doubtful meanings). It was that first category, personality traits, which ultimately became fundamental for the definition of stable traits in our current conceptualization of personality trait models (John et al., 1988). In fact, Cattell used this list as a starting point for the development of his personality model and its accompanying inventory, the Sixteen Personality Factors Questionnaire (16PF). Cattell’s model

included 16 factors that were derived from statistical and intuitive methods. Unfortunately, other laboratories were unable to replicate his 16-factor model, finding five-factors instead (Digman, 2002; Goldberg, 1981; John et al., 1988). Cattell and Mead (2008) argued that the 16PF contains 16 factors because it allows for “real world” intercorrelations between traits and reported that when data is “forced [into] orthogonal definitions” (p. 139) it is possible to identify five, second-order, global factors that are similar to those defined by the FFM. Although the 16PF and its model is less used today, it continues to be relevant and its items have provided the groundwork for the development of other trait-based models. For example, using items from the 16 PF, Tupes and Christal (1961) found five factors that they identified as: surgency, agreeableness, conscientiousness, emotional stability, and culture. Although the factors found in earlier studies were labeled differently than the five factors defined under the Costa and McCrae (1992) model, the underlying structure remains very similar.

Norman (1967) also created a list of terms descriptive of personality traits based on the work of Allport and Odbert (1936). He focused on compiling terms that could create a structured taxonomy that would facilitate scientific communication and personality assessment. Norman’s updated list contained 2,800 terms once he excluded terms that were ambiguous, unusual, evaluative, or descriptive of physical characteristics. Goldberg (1981) continued refining Norman’s work and used his final product to collect self-ratings and peer-ratings that when factor analyzed yielded five factors. Costa and McCrae (1976, 1992) continued this line of research, focusing on refining personality trait measurement. Their research began with 16 PF data that was collected from veterans. Analyses suggested a three-cluster structure. These clusters were identified as neuroticism, extraversion, and openness to experience (Costa & McCrae, 1976). As their research progressed, they concluded that the PF16 was a good measure

for the first two clusters, but inadequate to measure the openness to experience cluster that they had unearthed. To meet this measurement need, they developed the Neuroticism, Extraversion, Openness to Experience Personality Inventory (NEO-PI; Costa & McCrae, 1992) which consisted of three major domains that were further divided into 18 facets, six for each domain. These facets were created based on a review of existing personality literature. Further development of the NEO-PI led to testing the correspondence between its three-factor structure and the five factors that had emerged from the assessment of the English language in linguistics research (McCrae & Costa, 1985). Results showed high ($r > 0.50$) correlations between the three NEO-PI domains and their corresponding Norman (1963) factors. Ultimately, the data indicated that a five-factor solution was more appropriate. The consistency of this five-factor structure in personality data eventually led Costa and McCrae (1992) to revise their personality model and personality inventory. Using the same method of development of the first version of the NEO-PI they added two more domains, agreeableness, and conscientiousness (Costa, McCrae, & David, 1991). The final product included 240 items divided into 36 facets that were further nested within five broad personality domains. In more recent years the NEO-PI-R-3 and Costa and McCrae's (1992) model has become synonymous with the FFM and the "Big Five," even if the later term was originally attributed to Goldberg (Digman, 1990).

FFM and Psychopathology

Given its popularity and robust research base the FFM is constantly used to study personality in various areas of research, such as academic performance (Cupani, Garrido, & Tavella, 2013), morality (Pohling, Bzdok, Eigenstetter, Stumpf, & Strobel, 2016), physical inactivity (Sutin et al., 2016), harassment in the workplace (Nielson, Glaso, & Einarsen, 2017) and psychopathology (Sleep, Hyatt, Lamkin, Maples-Keller, & Miller, 2017). This last area has

been especially relevant to the study of personality psychopathology, which is the focus of the current study. The study of personality disorders became a prolific area of research after the introduction of Axis II in the third edition of the Diagnostic Statistical Manual of Mental Disorders (DSM III; APA, 1980). The new system introduced categorical diagnoses, had a stronger base of research than previous DSM editions, and conceptualized personality disorders as discrete entities. Ultimately, the categorical system was instrumental to the development of the abnormal personality literature. It allowed scientists to communicate and provided clinicians with a common language to describe individuals with chronic impairments. Unfortunately, the system did not withstand the test of time and empirical inquiry. Categorical diagnoses exhibited high comorbidity rates among personality disorders and mental health disorders with temporal presentations, such as depressive and anxiety disorders (Widiger & Samuel, 2005). Boundaries between normal and abnormal personality traits were ambiguous and although the categories were useful means of communication among clinicians and researchers, the utility of the system was limited (Widiger & Simonsen, 2005). The different permutations that could lead to a personality disorder diagnosis were too general to provide reliable information about treatment and the overlap between categories made consistency in diagnoses difficult (First et al., 2002).

Given the vast support amassed from research conducted on the dimensional structure of normative personality using adjectives and personality inventories, it was only natural that a transition towards a dimensional model of personality pathology would occur (Widiger & Samuel, 2005). The FFM was slightly revised to accommodate for the psychopathology and dysfunction found in personality disorders. Under this modified FFM personality disorders are interpreted as maladaptive or extreme variants of normative personality traits (Widiger, 2011). Research using personality inventories have shown that personality disorders can be

conceptualized using the FFM, without compromising reliability or validity (Verheul, 2005). Neuroticism, extraversion, and agreeableness have been especially useful in this line of inquiry, showing strong correlations with DSM personality disorder categories. In fact, neuroticism, extraversion, and agreeableness can partially account for the comorbidity between borderline personality disorder and narcissistic personality disorder (Widiger & Costa, 1994).

Conscientiousness and openness to experience are more problematic, but there is evidence indicating that the lack of strong associations in these two traits is likely a matter of measurement breadth (Haigler & Widiger, 2001). This is particularly relevant when research is conducted with personality inventories that were created to measure normative personality traits, such as the NEO-PI and IPIP. Haigler and Widiger (2001) provided support for this hypothesis. They found that 90% of the items in the NEO-PI are positively keyed towards adaptability. In fact, when the wording of the items is changed to be more indicative of maladaptive traits the strength of the correlations between the conscientiousness facets and OCPD increase significantly.

Assessment of Personality Psychopathology

Historically, there has been some disconnect between assessment of normal and pathological attributes (Digman, 1990). Consequently, the most widely used approaches, such as the Millon Clinical Multiaxial Inventory (MCMI), Personality Assessment Inventory (PAI), MMPI, and Rorschach, have little or no connection to normal personality models, and often lack a model of any kind, but are rather a potpourri of pathological traits found in clinical populations. Measures of normal personality, when used in clinical settings, have been criticized for missing important elements of pathology, being insensitive to variation at the extremes where pathology typically resides, and being poorly normed for such applications (Widiger & Costa, 1994). Similarly, using measures of pathology to assess normative personality traits—even though

normed on normal populations and even though pathology-based models of personality were common in the mid-20th Century—have been criticized for framing normal functioning in pathological terms, and for ignoring modern conceptions of personality (Graham, 2012).

Harkness, McNulty, and Ben-Porath (1995) made the first attempt to bridge existing measures of psychopathology to modern personality theory with their “marker” scales for the MMPI-2, organizing a subset of items into new scales that captured what they called the Psychopathology Five (PSY-5). These dimensions were developed directly from DSM-III-R personality disorders clinical criteria. Concerned with the likelihood of false positives in a dimensional measure they decided to include both normative and maladaptive characteristics. They hoped that a wide coverage would increment the scale’s sensitivity to underlying pathological factors (Harkness et al., 1995). The final version overlapped with four of the normative Big Five. Particularly interesting is the addition of Psychoticism, which does not directly overlap with Openness to Experience. This scale intended to account for reality testing differences among personality disorders, which could not be achieved with the Openness to Experience scale. Although useful, the measure only described very broad attributes, as there were simply not enough items to reliably assess a broad spectrum of more specific attributes. Later, attempts were made to adapt the PSY-5 scales to modern demands by adding a hierarchy (Arnau, Handel, Archer, Bisconer, and Gross, 2004). Unfortunately, the approach was, once again, restricted by the availability of existing MMPI-2 items and the PSY-5 subscales were found to be inadequate, with dubious reliability and validity (Quilty & Bagby, 2007).

Seeing these criticisms, some have ignored formal assessment methods in favor of interview data only (after all, “are they not both self-report?”). In fact, in almost all cases, DSM diagnoses are assumed to be drawn exclusively from patient and family interviews, perhaps with

observation thrown in. Indeed, formal assessment is almost exclusively the domain of psychologists, who see the value of formally assessing validity and reporting biases, attaching normative data to reports, and enhancing variability to make finer discriminations between etiologies in dysfunction. However, it is possible that the introduction of theoretically and psychometrical sound measurement tools that are easy to interpret can make a difference in the willingness of practitioners to use assessment tools in the provision of mental health services. The PID-5 with its grounded theoretical foundation and focus on measuring the idiosyncrasies of personality pathology has the potential to fill this gap.

The PID-5

Model and Structure

The PID-5 was created to measure pathological personality traits with the purpose of improving the conceptualization, diagnosing, and treatment of personality disorders. As such, the PID-5 is the psychometric half of a hybrid dimensional/categorical model included in section III of the DSM-5 that seeks to improve categorical diagnosis. In this model, there are seven possible categorical personality disorder diagnoses that can be broken down to the presence of certain personality traits and impairments in interpersonal functioning (DSM-5). The hybrid maintains six of the categorical diagnoses presented in the Personality Disorders section of the DSM-5. The personality disorders that were proposed for retention include: antisocial, avoidant, borderline, narcissistic, obsessive-compulsive, and schizotypal. After completing an extensive review of the personality disorders literature, the Work Group in charge of developing the alternative DSM-5 personality model concluded that there was not enough scientific evidence to support the delineation of a specific categorical diagnosis for paranoid, schizoid, histrionic or dependent personality disorder (Skodol et al., 2011). Whatever coverage is missed from not

including these four PDs is expected to be captured by a seventh diagnosis which is trait specified and is used when personality pathology exists that is not accounted for by the other six personality disorder categories. The hybrid model includes both specific dysfunctional personality traits, and information about the magnitude of impairments the individual is experiencing in their perception of self and interpersonal relationships. Samuel and colleagues (Samuel, Hopwood, Krueger, Thomas, & Ruggero, 2013) tested the ability of the PID-5 to aid in the measurement of specific dysfunctional personality traits and found that the PID-5 accounted for a significant amount of variance in categorical disorders as defined by DSM-IV criteria when using a cut off *T* score of ≤ 65 as the measure of clinical significance. The amount of variance accounted for was similar to that accounted for by the Personality Diagnostic Questionnaire – 4 (PDQ-4) which is a widely used measure in the diagnosis of categorical personality disorders in personality research. Although the PDQ-4 has acceptable psychometric properties as a diagnostic tool for personality disorders it is based on a categorical model and only asks yes or no questions to DSM-IV personality disorder criteria, limiting its ability to capture the nuances that the DSM-5 dimensional personality model seeks to provide to clinicians and researchers.

The Measure

The PID is a personality inventory that consists of 220 questions. It measures five personality domains that are further divided into 25 facets for specificity. The five domains are: Negative Affect, Detachment, Antagonism, Disinhibition, and Psychoticism. The inventory is meant to be used in clinical populations aged 18 years and older. Each question is answered in a 4-point scale that ranges from 0 (*very false or often false*) to 3 (*very true or often true*). Domain and facet scores are obtained by averaging the client's responses within the scales. Guidelines provided by Krueger and colleagues (Krueger, Derringer, Markon, Watson, & Skodol, 2012)

recommend that facets with more than 25% of missing data should not be used. In cases where 25% or less of the data is missing the score should be prorated by multiplying the summation of the answered items times the number of items in the facet and then dividing the product by the number of items that were answered in the facet. The reliability of the inventory has been shown to be adequate. In the original article, reported Cronbach's alphas ranged from .72 to .96, with a median of .86 (Krueger et al, 2012; Samuel et al., 2013). Subsequent studies have shown similar reliability coefficients, with clinical populations showing greater stability. Quilty and colleagues (Quilty, Ayearst, Chmielewski, Pollock, & Bagby, 2013) found similar results in a clinical population. They also reported McDonald's omega values for all domains: Negative Affect $\omega = .83$, Detachment $\omega = .75$, Psychoticism $\omega = .87$, Antagonism $\omega = .83$, and Disinhibition $\omega = .80$. Facet's average interitem correlations ranged from $r = 0.27$ to $r = 0.60$, with an average of $r = .47$.

As previously reported, the domains of the PID-5 are not conceptualized as conceptually or statistically independent. Therefore, it is not surprising to note that cross-domain correlations between the five PID-5 personality traits are relatively high, ranging from $r = .23$ to $.66$, with psychoticism exhibiting consistently higher cross domain correlations than other domains and disinhibition the lowest (Crego, Gore, Rojas, & Widiger, 2015). In fact, the average cross domain correlations for the PID-5 was $.57$ in a student sample, which is twice the average correlation of the NEO-PI-3 in the same sample, which was $.28$ (Crego et al., 2015). Anderson and colleagues (2013) also found higher levels of intercorrelations among the PID-5 domains ($r_s = .32$ -. 63 , $mdn. = .63$) as compared to the PSY-5 scales ($r_s = .04$ -. 49 ; $mdn. = .29$). Average PID-5 correlations across facets ranged from $.17$ (risk taking) to $.48$ (perceptual dysregulation, hostility, and perseveration; Crego et al., 2015). When correlated with other

personality inventories measuring both normative and pathological personality traits PID-5 facets showed mostly good convergent validity (Crego et al., 2015; Sleep et al., 2017; Yalch & Hopwood, 2016).

A study using latent state-trait analyses found that in average 7.7% of the score in the PID-5 scales is variance due to situational factors, which is consistent with previous estimates found in the personality traits literature (Zimmermann et al., 2017). Overall, it appears that the PID-5 is measuring stable traits, with 71% to 88% of the variance in the facets being accounted for by the personality construct. The facets that were the least affected by mood were withdrawal and manipulateness (5%) and the most affected was perseveration in which 11% of its variance was accounted for by situational factors (Zimmerman et al., 2017).

Item response theory (IRT) analysis conducted on a large primarily Caucasian sample showed that both the IPIP-NEO and the PID-5 offer a similar amount of coverage of latent personality construct, with the PID-5 offering better specificity at the highest thresholds of negative affect and detachment and the IPIP-NEO at lower levels of these traits (Suzuki, Samuel, Pahlen, & Krueger, 2015). For antagonism and disinhibition, the differences were less clear, but it appeared that the IPIP-NEO had a slight advantage over the PID-5 at both the lowest and highest levels of the trait. Interestingly, but not surprisingly, the psychoticism scale did not show much overlap with the openness to experience domain, unless all openness to experience facets were removed from the model, leaving only the imagination facet in the model.

PID-5 and other FFM Measures

Research studies looking at the external validity of the PID-5 have also shown promising results. Anderson and his team (2013) compared the PID-5 to the PSY-5 scales, which were derived from items of the MMPI-2. They found that all but one of the PID-5 domain scores were

highly correlated with conceptually similar PSY-5 scales. The exception was antagonism which showed significant correlations of .44 with both AGG-r and DSIC-r. When this relationship was further analyzed using regressions, the research team noted that both PSY-5 scales accounted for similar amounts of unique variance in Antagonism, AGG-r accounted for 24% of the variance and DISC-r for 27%. For all other scales, PSY-5 scales that were conceptually similar to the PID-5 domains clearly accounted for more variance than non-related PSY-5 scales. In the case of Negative Affect, NEGE-r on its own significantly accounted for 62% of the variance in this PID-5 domain (Anderson et al., 2013). Given that the PID-5 is designed to measure the pathological spectrum of personality we can hypothesize that when factor analyzed with other well validated measures of personality traits the PID-5 items and the other measure's items (NEO-PI-3, PSY-5, etc.) will form five factors with conceptually similar traits clustering together due to their shared variance. This hypothesis has been supported by factor analytic research using measures of normative and pathological personality traits such as the NEO-PI-3 (Suzuki et al., 2015), PSY-5 (Anderson et al., 2013), and the Dimensional Assessment of Personality Pathology – Basic questionnaire (DAPP-BQ; Van den Broeck et al., 2014)

Five-Factor Structure

There is considerable support for the five-factor structure of the PID-5. Initially, analysis of the PID-5 suggested the retention of three to six factors but given that the Krueger and colleagues (2012) were interested in identifying the highest number of interpretable factors they decided that a five-factor structure was the most appropriate. They provided evidence for their decision by noting that the sixth factor proposed by a parallel analysis only contained two facets, emotional lability and hostility (Krueger et al., 2012). Research after that first study has provided support for a five-factor structure in both college and clinical samples (Quilty et al.,

2013; Wright et al., 2012). Quilty and colleagues (2013) conducted individual EFAs on each of the facets that makeup the PID-5 and found that with exception of risk taking all of them appeared to be unidimensional, as defined by facets having a one-factor solution. Of interest is to note that the risk-taking facet had a two-factor solution that appears to be differentiated by the keying of items, with the negatively keyed items loading on one factor and the positively keyed items in a second factor. Therefore, Quilty and his team (2013) theorized that the lack of unidimensionality of this facet could be attributed to an effect in the direction of the item keying. EFAs conducted in different samples across the world have also provided support for a five-factor structure, often replicating the factors proposed by Kruger and colleagues (2012), but, with some variation in where the facets load (De Fryut et al., 2013; Gutierrez et al., 2015; Roskam et al., 2015; Somma, Markon, Krueger, & Fossati, 2019).

Support for the structural validity of the PID-5 also comes from hierarchical studies that have studied the five pathological personality domains within a general pathology model. Wright and company (2012) proposed a five-level hierarchical structure of psychopathology that begins with one general factor that is further divided into two factors reflecting the division of general psychopathology into externalizing and internalizing disorders. With each added level of analysis, the specificity of the factor increases culminating with five factors that resemble the PID-5 domains at the fifth level (Wright et al., 2012). This hierarchical structure has been replicated in French (Roskam, et al., 2015) and Italian (Fossati, Krueger, Markon, Borroni, & Maffei, 2013) samples. Krueger and Markon (2014) argue that the fit of the five PID-5 domains into this structure strengthens the argument of the relevance of personality pathology within models of general psychopathology, especially as the discipline moves towards dimensional models in the conceptualization of all psychological disorders.

Openness to Experience and Psychoticism

Four of the five psychopathological domains appear to have significant overlap with the normative five factors. The outlier appears to be the psychoticism domain. When scales measuring the Big Five are factor analyzed with the scales in the PID-5 to form one five-factor model, the openness to experience items do not appear to appropriately capture the variance of the psychotic component of personality (Watson, Clark, & Chmielewski, 2008). In fact, a study found that the items in the Openness to Experience scale only accounted for 25% of the variance in the PID-5's psychoticism scale, which was half of the variance of what other Big Five scales accounted in their PID-5 counterparts (Quilty et al., 2013). This same study also found low associations between OE and the PID-scales, with the only exception being Risk Taking which had a correlation of .32 with OE (Quilty et al., 2013). In a factor created by OE, conventionality, absorption, and psychoticism, the later one obtained the lowest factor loadings (.45) indicating a weaker association within the factor as compared with the other facets. These low associations have raised questions regarding the relationship of psychoticism in the structure of pathological personality traits, but research has shown that psychoticism has a strong association with measures of general personality dysfunction (Zimmermann et al., 2014). Therefore, even if the association between OE and psychoticism is not completely clear, psychoticism continues being a relevant trait to assess in personality pathology. It has been postulated that one of the reasons as to why OE does not have strong associations with pathological personality traits is because the scale was initially designed to measure ideal personality traits related to self-actualization and intelligence, which would not be necessarily related to the eccentricity and oddity that is observed in psychoticism (Zimmermann et al., 2014).

Psychopathology Correlates

Sleep and colleagues (2017) studied associations between the PID-5 domains and a number of externalizing and internalizing criteria. They found that antagonism was positively predictive of levels of aggression, alcohol misuse, substance use and antisocial behavior.

Disinhibition was only associated with aggression and alcohols misuse. As expected, anxiety and depression, both internalizing criteria, were positively associated with negative affect and detachment (Sleep et al., 2017). Associations between psychoticism and the internalizing and externalizing criteria were only significant when analyzed using bivariate correlations, but lost significance when multivariate analysis were used. In total, PID-5 domains appeared to account for 10% to 36% of the variance in the criteria used in the study (Sleep et al., 2017).

Interestingly, when using the IPIPI, which is a scale that measures personality traits in a similar manner as the NEO-PI-3 the range of variance accounted by the IPIP was similar to that of the PID-5, ranging from .10% to 35%, with the major differences being found in aggression and anxiety. It appears that the PID-5 accounts for more variance in the externalizing criteria and the IPIP in the internalizing criteria where there was a significant difference of 7% variance accountability when predicting anxiety scores (Sleep et al., 2017).

The PID-5 has also shown promising results in the development of personality profiles for individuals at risk on developing a gambling disorder (Carlotta et al., 2015). In an Italian sample, individuals at high risk exhibited higher elevations in the antagonism and disinhibition domains, even after controlling for substance and alcohol use. This pattern of elevations was consistent with an externalizing conceptualization of gambling disorder. Carlotta's research team (2015) also noted that high risk gamblers exhibited low levels of openness to experience and conscientiousness as measured by the Big Five Inventory (BFI), but these differences

between groups lost significance once alcohol and substance use was controlled for. These results provide some evidence that the PID-5 could potentially provide incremental validity in the study of individuals with gambling problems, but further research is needed to replicate these results and further elucidate the PID-5's ability to study etiological factors in gambling disorders. In a Romanian sample of law enforcement personnel, the PID-5 added incremental validity to the assessment of narcissism, psychopathy, and Machiavellianism beyond what was measured by the FFM (Grigoras & Wille, 2017). Antagonism accounted for 26% of the variance in Machiavellianism. Negative affect, detachment, and antagonism predicted 29% of the variance in narcissism and 20% of psychopathy was accounted for by disinhibition and psychoticism (Grigoras & Wille, 2017). The PID-5 was also able to differentiate individuals with Borderline personality disorders from individuals in clinical distress without a PD diagnosis (Calvo et al., 2016). This research provides support for the ability of the PID-5 to measure real life events and opens possibilities for the scientific understanding psychopathology from a dimensional personality trait model.

The PID-5 as a Diagnostic Tool

The PID-5 is supposed to be a support tool to aid psychologist to study personality and aid in the assessment of criterion B of the personality model included on Section III of the DSM-5. Research has shown that the PID-5 can reliably assess PD variance as per DSM-IV criteria (Hopwood et al., 2012). This has been replicated in both undergraduate and clinical samples. Few and colleagues (2013) found that the domains and facets identified by the Work Group to delineate personality disorders accounted for 37% of the variance in personality diagnoses in an outpatient sample. When the PID-5 facets are correlated with PD diagnoses, the associations are mostly consistent with those proposed by the Work Group. Associations with the facets appear

to be more complicated, exhibiting some unexpected relationships, such as risk taking having low correlations with BPD or intimacy avoidance having a small correlation of $r = .13$ with Avoidant PD (Anderson, Snider, Sellbom, Krueger, & Hopwood, 2014). Since they found unexpectedly high correlations with facets that were not identified by the Work Group, Anderson and colleagues (2013) assessed the incremental validity that could be obtained by using non-specified facets with high correlations with PDs as predictors. They found that for all diagnoses, except for schizotypal and dependent PD, there was an increase in accountability of unique variance when including non-specified facets in the predictive models. These results show that it is possible to increment the amount of personality pathology that can be assessed with the PID-5. Unfortunately, adding more facets and domains to the delineated categorical diagnosis could affect our ability to differentiate between PDs. With the overlap, there would be an increase in variance accounted for by PID-5 facets, but this would be achieved at the expense of discriminative power between PDs, which would bring back the concerns of comorbidity that were present in the categorical model. When adding all the moderately correlated facets to the predictive models the average intercorrelation was $.79$ which is almost a third more than the average intercorrelation between traits ($r = .46$) when only the facets included in the DSM-5 model were entered (Hopwood et al., 2012). This level of intercorrelation could be problematic but the strength of association between facets are to be expected given the saturation of personality pathology in the PID-5 and the high levels of intercorrelations between domains. As of now it appears that the PID-5 is able to measure personality traits mostly as prescribed in the DSM-5 model and there is evidence that it can differentiate between individuals with and without BPD (Calvo et al., 2016). The PID-5 has also been used to measure psychopathy, where

antagonism, detachment, and disinhibition accounted for 50% of the variance in psychopathy as measured by a self-report questionnaire (Fossati et al., 2013).

Overall, it appears that the PID-5 has merit and the potential to help clinicians and researchers in their understanding of personality psychopathology. However, moving forward it might be beneficial to study how different experiences and backgrounds affect scores specifically when the PID-5 is used to make decisions about diagnosis and treatment. This is particularly relevant for underserved populations such as the LGBT community, who often face discrimination and difficulties accessing mental health resources (Bogart, Revenson, Whitfield, & France, 2014). A study assessing the PID-5's ability to capture personality pathology in LGBT individuals found that lesbian and bisexual individuals had significantly higher scores in PID-5 facets and that when using a t-score ≥ 65 as a cut-off point for significantly elevated scores lesbian and bisexual women were twice as likely to be classified as avoidant, borderline, narcissistic or histrionic. Gay men were 3.19 times more likely to be classified as dependent when compared to heterosexual samples (Russell, Pocknell, & King, 2017). These results are concerning, but we should also take into account that this study did not use the full DSM-5 criteria to diagnose PDs. They only used Criterion B, leaving out Criterion A, which focuses on impairment in various social and interpersonal domains. Using this method, one can infer about elevations and the presence of a pathological trait, but not the functional impairment that the person might experience. However, considering the consequences that high scores and over diagnosing could have in the lives of individuals it is important to continue researching these effects. These positive findings of differences are important because it opens the possibilities that other populations, such as racial minorities could also be at risk of overpathologization. Unfortunately, at the moment no research exists on whether this is a problem or not.

The Big Five and Culture

There are various ways in which personality can be measured, and this will depend on the theory that is being used to conceptualize the construct of personality. In the case of personality trait theory, it includes assumptions that personality traits identified by the FFM are stable and universal. As such these five traits are supposed to be replicable across cultures and populations, making them appropriate for the study of personality across cultures. This method is what Berry (1969) would call etic, and it involves bringing in a structure that has been found in another culture and imposing it onto the culture of interest.. Due to the mainstream status of Western psychology it often means that measures or models that have been developed in Europe or the United States are translated and applied to other cultures in attempts to make assumptions of universality and generalizability (Ortiz, 2015). Etic methods have received considerable support, but there is a legitimate concern that this approach does not fully capture cultural differences. An alternative to etic methods would be to take an emic approach. Using this approach researchers develop their measures or theories within the cultures of interest (Berry, 1969). The benefit of this method is that it allows researchers to tap into specific personality dimensions or behavioral correlates that are particularly relevant to the culture studied. Unfortunately, this type of research for non-Western countries is limited and the research that is available provides mixed results. Triandis and Suh (2002) reviewed this literature and found that when indigenous personality measures are created these do not always map perfectly onto the FFM, at times obtaining four factors (minus OE) or nine in the case of an emic personality measure created in Mexico. Differences in measurement can be important, especially when researchers or clinicians' intent to use the measures as predictors of behaviors of relevance in the culture of interest. In fact, it appears that including emic components to etic methodologies can increase

the ability of inventories to measure culturally relevant traits such as, filial piety in Chinese society (Zhang & Bond, 1998). Unfortunately, the disadvantage of using emic methods is that it is much more difficult to make direct comparison between cultures and to connect findings to the international nomological network of research. This is likely why epic methods are less commonly used, but this means that researchers that are conducting cross-cultural research need to ensure that imported measures are adapted to the target culture responsibly. According to Ortiz (2015) imported measures should be adequately and equivalently translated, have good psychometric properties, proper use of norms and an awareness of how culture might be relevant in the interpretation and use of scores.

As per etic standards of measurement, the five-factor model of personality does well and its structure is often replicated in other cultures, with minor reductions in reliability coefficients (Church, 2001). Given that the PID-5 is supposed to measure the maladaptive variants of the FFM it is not surprising that translations of the PID-5 have shown acceptable reliabilities, and often replicate a five-factor structure when using EFAs and CFAs (Krueger & Markon, 2014). Up to date, the PID-5 has been translated into Danish (Bo, Bach, Mortensen, & Simonsen, 2016), Dutch (De Fryut et al., 2013), French (Roskam et al., 2015), Italian (Fossati et al., 2013), Flemish (Bastiaens, Smits, De Hert, Vanwalleghem, & Claes, 2016), German (Zimmermann et al., 2014), Persian (Soraya et al., 2017), and Spanish (Gutierrez et al., 2015). At the domain level, the five-factor model is often replicated and when reported, coefficients of congruence between the translated version being studied and Krueger's (2012) structure are adequate (Fossati et al., 2013; Markon & Krueger, 2014). At the facet level, replication is more complicated, with some facets such as risk taking and suspiciousness not loading in any factor or loading in a different factors than the one that they were assigned upon development (Krueger &

Markon, 2014). This does not appear to be a problem with translation, or a culture effect given that these variations occur even within English speaking samples within the United States (Wright et al., 2012). The reasons as to why lack of stability occurs at the facet level is not fully understood, but it is likely that it is influenced by the high levels of intercorrelations between facets and domains. Given the saturation of psychopathology in the PID-5 and its hierarchical structure, a general factor measuring negative valence or a similarly global general distress construct could also account for the high cross loadings and intercorrelations among facets (Watters & Bagby, 2018).

PID-5 in Spanish

The purpose of this study is to continue building on the psychometric properties and validity of the PID-5 with non-English speakers. For this study, the population of interest is Spanish speaking Hispanics that are currently living in the United States. However, it is expected that this translation could in the future, with adequate data, be used with other Spanish-speaking populations in Latin America. In 2015, Gutierrez and his team translated and provided initial validation data for the PID-5 into Spanish with good results. The PID-5 translation replicated the five-factor structure presented in Kruger et al. (2012) with congruent coefficients over .95, had good internal consistency as measured by alpha coefficients (.65-.93 in their community sample and .76-.95 in their clinical sample), and PID-5 scores could differentiate between a clinical and a community sample (Gutierrez et al., 2015). This translation has also been successfully used to measure personality pathology in borderline personality disorder (Calvo et al., 2016). However, this translation was done in Spain and its reliability and validity were tested using a Spanish sample. Although Spanish is spoken in both Spain and in Latin America, and Spain has a strong history of colonialism with Spanish speaking countries in Latin

America, this shared history does not necessarily translate into equivalence in language use or culture. Therefore, it was decided that it was appropriate to independently create a translation of the PID-5 for Spanish speaking Hispanic individuals living in the United States.

In a review article about cultural adaptation of health-related measures Guillemín, Bombardier, and Beaton (1993) suggest that instruments should be adapted to the version of the language in which these are supposed to be administered, especially when some cultural differences are to be expected. Psychological research in this area is limited, but a study looking at emotion words describing beverages found that the way language was used was more similar in four English speaking countries (U.S., UK, Australia, and New Zealand) than between Mexico and Spain (Van Zyl & Meiselman, 2015). Van Zyl and Meiselman (2015) also found that Mexican patterns of responses were more similar to U.S. respondents than to Spaniards, which would make sense if we account for the fact that Mexico has constant contact with the U.S. and the influence of Spain has steadily decreased in Mexico since Mexico gained its independence in 1810. At the same time the U.S. and Mexico have strengthened their cultural and trade ties. Also, when Benet-Martinez and John (1998) attempted to adapt the BFI from Castilian Spanish to the Spanish spoken in Latin America they had to modify 23 out of the 44 items that make up the BFI to fit the needs of their sample after they consulted with bilingual individuals living in the U.S. that had ancestry from different parts of Latin America. When comparing a BFI translation to an emic measure using Argentinian Spanish the measure using Argentinian language had better reliability, fit the five-factor model better, and had higher loadings than the generic direct translation used in the BFI (Ledesma, Sanchez, & Diaz-Lazaro, 2011). From this data, it can be inferred that translations can exhibit a decrease in reliability when transferred to other cultures, if some of that introduced error could be accounted for by language, creating a Latin American

Spanish specific translation could aid in the measurement of personality pathology in Hispanics living in the U.S.. This is particularly relevant because the U.S. is the second country with the most Spanish Speakers in the world, only second to Mexico. According to 2010 census data, the U.S. had 36 million Spanish speakers in 2010 and this number is projected to increase to 43 million by 2020 (Ortman & Shin, 2011). The Pew Research Center conducted a study using a representative Hispanic sample of 1,220 and they found that 38% of their sample was Spanish dominant, 38% bilingual, and 24% English dominant. Given this data, a well validated measure could potentially improve the ability of clinicians and researchers to serve Spanish dominant Hispanics who are often underserved (Sentell, Shumway, & Snowden, 2007).

Purpose

The purpose of this project is to continue testing the fitness of the PID-5 as a cross-culturally valid measure of personality traits within a five-factor model, more specifically this project will focus on translating the PID-5 into Spanish commonly spoken in the U.S. This will be accomplished by assessing the reliability, structure, and validity of the PID-5 using a sample of Spanish speaking Hispanics residing in the U.S. The hope is to begin the process of building a foundation for future research on using the PID-5 effectively in the Hispanic community. It is expected that by removing language barriers, the ability of monolingual-English-speaker clinicians and researchers to comprehend the Hispanic experience will improve.

If the Personality Inventory DSM-5- Spanish Version (PID-5-Sp) captures personality variance in a similar manner than the PID-5 and its translations, we would expect good reliability coefficients, a replication of the five-factor model, and moderate to strong correlations between the PID-5-Sp domains and FFM scales. We would also expect for Antagonism and Disinhibition

to be predictive of aggression scores and for Negative Affect and Detachment to account for a significant amount of variance in depression and anxiety scores.

Method

Translation

Consistent with previous translations and validations of English instruments to other languages, the PID-5 was adapted via the back-translation technique (Guillemin et al., 1993). The PID-5 was first translated from English to Spanish by the author, a native Spanish speaker with experience working with Spanish speaking clients from Latin America. After the translation was completed, the items were back translated by three native Spanish speakers and two non-native Spanish speakers with clinical experience working with Spanish speaking clients. The back translators represented Spanish spoken in Mexico, Venezuela, and Honduras. All back translators were naïve to the original version of the PID-5. The back translation was compared with the original English PID-5 items. Discrepancies were resolved between the main translator and the back-translators to ensure that the meaning behind the English items was retained.

In instances in which no direct translations were available, measures were taken to ensure parallelism in meaning with the English items, taking into account the cultural context of both versions. For example, item 133 (“It seems like I am always getting a “raw deal” from others.”) uses the idiom “raw deal” to ask about the respondent’s perceptions of how they are treated by others. A word by word translation of “raw deal” does not transmit the same meaning as the English idiom, and to our knowledge, there is no idiom in Spanish that embodies the same concept. Therefore, we researched synonyms and definitions of the idiom looking for the terms that would most effectively convey the meaning of the original question into the Spanish translation. The final version of item 133 asks respondents if they feel like they are always

treated unjustly by others (Pareciere que siempre estoy siendo tratado injustamente por los demás.)

The final translation of PID-5, which was coined as PID-5-Sp, was given for feedback to a Spanish speaking psychologist in the community, a PhD level clinician at a public university in Mexico, and to a monolingual Spanish speaking individual. The final translation incorporated their feedback, ensuring that PID-5-Sp items were meaningful to Spanish speakers while also remaining truthful to the original items. Changes made primarily targeted syntax and awkward wording that resulted from attempts to keep parallelism to the English items in the Spanish translation.

Data Collection

Data for the current study was collected using Amazon Mechanical Turk (MTurk) which is a crowdsourcing internet marketplace in which individuals can request human participants to complete “Human Intelligence Tasks (HITs)” in exchange of monetary compensation. To our knowledge this is the first PID-5 translation study to collect data using MTurk. However, there is some precedent of the use of data collected online in the study of the psychometric properties of the PID-5. In fact, the initial data used in the development of the PID-5 was collected using Knowledge Networks, which is a web-based survey data collection platform. Also, there has been a significant amount of research on the appropriateness of using MTurk samples in social research. Studies have shown that samples obtained through this marketplace are adequate and at times more representative than samples collected on other online platforms or at universities (Behrend, Sharek, Meade, & Wiebe, 2011; Buhrmester, Kwang, & Gosling, 2011). PID-5 validation and translation studies often use student samples which have shown reliable results, but do tend to be limited in their generalizability (Bo et al., 2016; De Fryut et al., 2013;

Zimmermann et al., 2014). MTurk samples exhibit some similarity in response patterns and levels of education as college student samples but have the added benefit of a wider age range and can provide access to a wide range of diversity in participants when the appropriated inclusion criteria (language spoken, ethnicity, etc.,) is integrated into the data collection process.

In a study designed to test the reliability and quality of data collected using MTurk, Buhrmester and colleagues (2011) found acceptable reliabilities as measured by alpha coefficients and tests-retest reliability correlations evaluated in measures of personality, psychopathology, and self-esteem. The reliability of the ratings did not appear to be affected by the amount of compensation provided, which ranged from 2 to 50 cents per task. MTurk has users in 190 countries all around the globe, with the majority of participants being fluent English speakers residing in the United States and India. However, a study conducted by Pavlick and colleagues (Pavlick, Post, Irvine, & Callison-Burch, 2014) assessing the feasibility of conducting research in a language other than English using MTurk participants concluded that there are 13 languages for which fast, reliable, quality data can be confidently obtained, including bilingual Spanish speaking individuals who are the target population of the present study. Although, MTurk samples are primarily composed of non-pathological individuals, these samples do contain individuals with mental health disorders, probably commensurate with population rates. In a sample of almost 500 participants, 21 % reported having a psychiatric diagnosis, 12.5 % reported using a psychoactive medication in the last two years and 5.6% were currently engaged in psychotherapy (Shapiro, Chandler, & Mueller, 2013). The most common disorders reported were depression, social anxiety, and having a history of exposure to traumatic events (Arditte, Cek, Shaw, & Timpano, 2016; Shapiro et al., 2013). Shapiro and colleagues (2013) also

reported that a substantial number of individuals in their sample of 500 had positive screenings for potential substance use problems.

Given the characteristics that have been reported in the literature about MTurk samples and the support for its appropriateness of use in social sciences research it was decided that a sample collected through MTurk was appropriate for the current study. It has the benefit of targeting Spanish speakers from across the U.S. and from different countries of ancestral origins, allowing for the testing of PID-5-Sp properties in a more diverse sample than what could be obtained if data were to be collected in one site. Data collected in one geographical site is often restricted by the regional makeup of the Hispanic population that inhabits the area (e.g. Puerto Ricans in New York, Mexican American's in the Texas-Mexico border, Cubans in Florida, etc.) and given that the current translation aims to be appropriate for Hispanics across different ancestral countries of origin it would be beneficial to have a diverse sample, representing Hispanics from various regions of the country. The uniqueness of the MTurk sample would likely place some limits in the external generalizability of the results, but its use is appropriate for the exploratory purposes of the current study, which attempts to explore the psychometric properties of the PID-5-Sp in a Hispanic sample.

Participants were asked to complete a survey that included demographics, the PID-5-Sp, BFI, and measures of anxiety, depression and aggression. They were told that the questionnaire would take approximately 45 minutes to complete and would ask them about their beliefs, behavioral patterns, and mental health. Those who accepted the "HIT" were redirected to Survey Monkey, where they completed the questionnaire and were assigned a unique completion code by which they demonstrated they had completed the task. Inclusion criteria included, being of legal age, Hispanic, a native Spanish speaker, to be currently living in the U.S., and having a

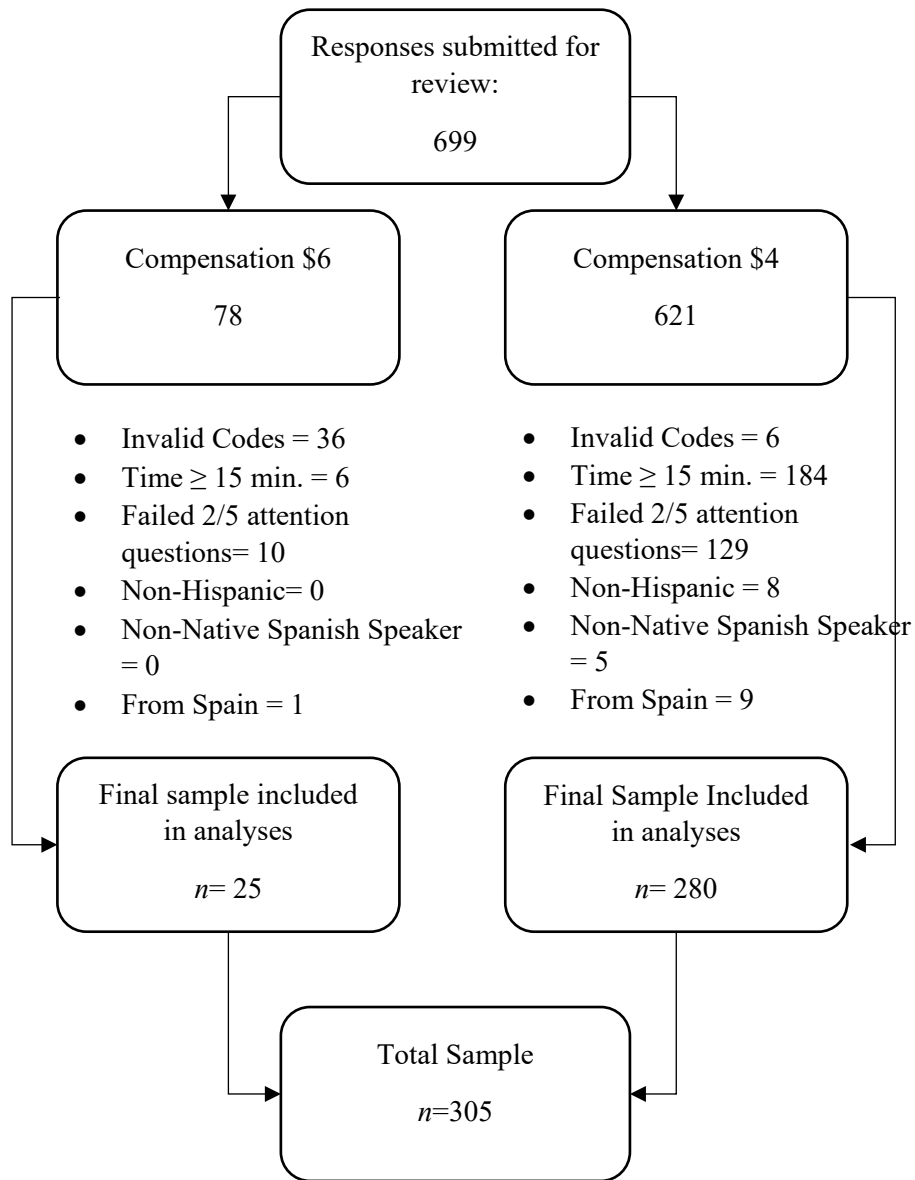
95% or higher task acceptance rate. Task acceptance rates are often used as inclusion criteria to improve data quality. A 95% task acceptance rate means that workers have completed successfully 95% or more of the past tasks that they have been assigned. Criterion information was collected via self-report, with the exception of geographical area and acceptance hit rate, which were set as restrictions for participation by the MTurk interface, not allowing individuals who did not meet the criteria to participate in the study.

Participants were initially offered \$6 dollars for their participation. However, this compensation amount was above market value and led to numerous fraudulent responses, which included participants submitting answers soon after accepting the HIT or using invalid or previously submitted completion codes. Therefore, the compensation was changed to \$4 dollars in subsequent requests for “workers.” This amount of compensations was more consistent with the market value of tasks published at the time of data collection. The final study design and the change in compensation was approved by Brigham Young University’ Institutional Review Board. Only 78 participants were offered \$6 dollars in compensation for their participation and out of those 78 individuals only 25 responses were kept in the final data set after applying exclusion and inclusion criteria. The process of inclusion and exclusion of participants has been delineated in Figure 1.

Ultimately, 699 responses were submitted for review to the MTurk interface, 78 were offered \$6 for their participation and 621 were offered \$4 after the IRB approved the change in compensation. Upon review of the quality of the data only 305 responses were retained and used for the data analyses presented in the current study. Decisions regarding which data points to retain (see Figure 1.) were based on whether they met the previously stated criteria (e.g. being Hispanic, Latin American country of origin, etc.,) and an evaluation of attention measures. The

use of attention checks has been recommended to improve the quality of data collected through MTurk (Goodman, Cryder, & Cheema, 2013). Therefore, the current questionnaire included five

Figure 1
Flow of Participants Included in Final Analyses



Note. Participants from Spain were eliminated from the sample because the focus of the study was on Spanish speakers from Latin America

questions that were scattered throughout the survey and asked participants to choose a particular answer (e.g. Conteste “Muy de acuerdo” a esta expresion.) to demonstrate that they were paying

attention. Participants who missed more than two attention questions were disqualified from participation. To ensure that participants were putting acceptable effort in the completion of the task, time limits were placed on the work that was accepted. No submissions that took less than 15 minutes to complete were accepted. This cut off point was deemed reasonable after taking into consideration that two of the back translators were able to complete the full survey in no less than 23 minutes. The added flexibility in time was included to account for any practice effects that might exist for MTurk participants, who might take surveys for compensation multiple times a day, making them faster than the average responder. This hypothesis has been supported by the literature which reported the MTurk participants show significantly faster completion rates than samples collected from colleges or other online data collection platforms (Kees, Berry, Burton, & Sheehan, 2017). The average completion time rate in the current sample was 35.67 minutes with a range of 15 to 113.7 minutes, $SD = .22.24$, $mdn. = 31.58$.

Participants

The 305 participants that were retained had an average age of 32.23 ($SD=8.60$) that ranged from 20 to 63 years of age. The sample included 151 (49.6%) males and 154 (50.5% females). Of these participants, 44.3% reported being single and 55.4% reported being married (44.3%) or in a domestic relationship (11.1%). The majority of the sample (88.5%) reported having at least some college education, with only 3 participants reporting having less than a high school degree, which is consistent with previous descriptions of MTurk samples in psychological research (Behrend et al., 2011; Buhrmester et al., 2011). Surprisingly, 40% of the sample reported struggling with mental health difficulties, primarily reporting depression (18.4%) and anxiety (14.8%). This was almost double the 21% reported by Shapiro and colleagues (2013). However, their sample was primarily Caucasian (83.5%). Also, instructions to participate in the

study warned participants that the questionnaire would ask about the state of their mental health, which could also have affected the way in which participants responded, either by allowing them to disclose their psychiatric diagnosis or setting the expectation that a mental health diagnosis was desirable. Psychiatric history was self-reported, answering to a question asking participants if at any time they had been diagnosed with a mental health disorder. All 305 participants identified as Hispanic, but only 259 (84.5%) provided information about their country of ancestral origin. Consistent with previous studies evaluating the demographic properties of Spanish speaking samples in MTurk, the group that was most heavily represented were individuals of Mexican ancestry, who represented 48.6% of those that reported a country of ancestral origin, followed by Venezuela (9.7%) and Puerto Rico (8.9%).

Measures

Personality Inventory for DSM-5, Spanish version (PID-5-Sp). The PID-5 is a personality inventory that consists of 220 items measuring maladaptive personality traits. Each question is answered on a Likert scale ranging from 0 (“*Very True* or *Often False*”) to 3 (“*Very True* or *Often True*”). The English version of the PID-5 has a five-factor structure that has been replicated in other languages, such as Italian (Fossati et al., 2013) and French (Roskam et al., 2015). Alpha coefficients for the five scales have been reported to range between .75 and .95 with a median value of .86 (De Fryut et al., 2013).

Spanish Big Five Inventory (BFI). This personality inventory measures the big five normative personality traits and consists of 44 items rated on a Likert scale from 1 to 5 from “*Strongly disagree*” to “*Strongly agree*”. This questionnaire was developed based on Goldberg’s lexical data and is meant to measure the Big Five with fewer items than the NEO-PI-3 (John, Donahue, & Kentle, 1991). In 1998, Benet-Martinez and John translated the original BFI into

Castilian and Latin American Spanish. They collected and analyzed data from college students and a community sample. The reliability of the translated version was slightly lower than that of the English version, but overall was acceptable with a mean reliability of $\alpha = 0.78$ for the Spanish translation and $\alpha = 0.83$ for the English version (Benet-Martinez & John, 1998). The five-factor model was replicated in all versions of the questionnaire, but the Spanish translation had lower factor loadings 0.69 to 0.77 compared to the English version with loadings of 0.73 to 0.80. This questionnaire was also tested in a Mexican sample in which alpha values ranged from $\alpha = 0.62$ (agreeableness) to $\alpha = 0.78$ (conscientiousness), with a full-scale reliability of $\alpha = 0.72$ (Reyes, Alvarez, Paredo, Sandoval, & Rebolledo, 2014). Validity information beyond the five-factor replicability is limited, but in this same Mexican sample the BFI exhibited significant differences between individuals with and without a psychiatric diagnosis in agreeableness, conscientiousness, and neuroticism.

Aggression Questionnaire-Refined (AG-R). The original AG-R was created by Bryant and Smith (2001). It measures aggression in four subscales: physical aggression, verbal aggression, anger, and hostility. This measure was later adapted by Gallardo-Pujol and colleagues (Gallardo-Pujol, Kramp, Garcia-Forero, Perez-Ramirez, & Andres-Pueyo, 2006) to use in a Spanish population. Their adaptation kept the original 12-items. However, they decided that a five-point Likert scale (from “*Never*” to “*Always*”) was more appropriate for their sample than the original six-point Likert scale. The four-factor structure was replicated in the Spanish sample and the reliability for the subscales and total scores ranged from .58 to .78. AG-R scales showed significant correlations (.23 - .39) with impulsiveness as measured by Barratt’s Impulsiveness Scale (BIS-10). Motor Impulsiveness had the highest correlation of all subscales,

$r = .50, p \leq .01$, indicating good conceptual congruence as theorized by Gallardo-Pujol and colleagues (2006).

Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 questionnaire is a self-report measure used to screen for depression. It consists of nine questions that are congruent with DSM-IV-TR depression criteria. Severity scores range from 0 to 27. In Spanish speaking samples, the recommended cut off score for a positive depression screening is ≥ 10 because this score has been shown to correlate with moderate to severe levels of depression. The PHQ-9 also has support as an effective measure of depression in Mexican samples (Familiar et al., 2015). Merz and colleagues (Merz, Malcarne, Roesch, Riley, & Sadler, 2011) also reported good internal ($\alpha = .84$) and structural consistency in a sample of Hispanic females.

Generalized Anxiety Disorder-7 (GAD-7). The GAD-7 is a self-report screener often used to assess for GAD. The scale consists of seven items measuring the presence of anxiety related symptom as measured by DSM-IV criteria. The items are measured on a scale from 0 (“*Not at all*”) to 4 (“*Very often*”). Total scores range from 0 to 21. It was initially tested in an English-speaking sample and found to have good psychometric properties (Spitzer, Kroenke, Williams, & Lowe, 2006). Mills and colleagues (2014) translated and reported good internal consistency values ($\alpha = .93$) for a Spanish speaking sample of Hispanics living in the U.S. The GAD-7 also exhibited good convergent validity, as measured by construct consistent correlations with the PHQ-9 ($r = .70, p \leq .01$), Physical Health ($r = -.31, p \leq .01$), and a correlation of $r = .66, p \leq .01$ with the Perceived Stress Scale (PSS; Mills et al., 2014).

Statistical Analysis

After ensuring all data points met inclusion criteria variables were evaluated for missing data and normalcy. Participants were required to respond to all items to successfully complete

that task published on MTurk, therefore there was not data missing in submitted responses.

Given that the PID-5-Sp is a measure of pathology and that the data was collected in a normative sample PID-5-Sp facets exhibited skewedness, as is typical. This is reported along with the descriptive statistics of the 25 facets and five PID-5 domains in Table 1. Cronbach's alphas were used to assess the internal consistency of all 25 facets and the five domain scores, which were calculated using Kruger et al. (2013) directives.

Exploratory factor analysis. In congruence with previous studies assessing the structure and validity of the PID-5 the current study used Exploratory Factor Analyses (EFA) to explore the structural replicability of the PID-5 in a Spanish speaking, Hispanic sample. EFA's were conducted using principal factors methods in Stata 15. This option was chosen because it is more appropriate for the measurement of latent variables and can account for data normality deviations (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Maximum Likelihood extractions were also explored to evaluate differences across methods, but no meaningful differences were noted. Guidelines for sample size in EFA methods vary, but recommendations have been made for samples to contain no less than 100 observations (Kline, 1994). Proposed ratios using sample size to number of variables have suggested ratios anywhere from three to ten observations per variable. (Cattell, 1978; Gorsuch, 1997; Nunnally, 1978). Taking these guidelines into consideration and opting for a conservative approach it was decided to include a minimum of ten observations per variable which set the minimum of 250 observations at the facet level and 2,200 at the item level. Since previous translations studies (Fossati et al., 2013; Gutierrez et al., 2015; Roskam et al., 2015) have conducted EFAs at the facet level it was decided that continuing with this methodology was appropriate for the developmental stages of the PID-5-Sp. Therefore, an EFA was conducted on the 25 PID-5-SP facets, which required a

minimum sample size of 250 subjects to achieve acceptable levels of statistical reliability.

Conclusions regarding the number of factors to extract were informed by substantive interpretability, eigen values, and parallel analyses (Buja & Eyuboglu, 1992). Parallel analysis suggests the number of factors to extract based on 1,000 random permutations of the collected PID-5-Sp data. Then the eigen values of the factor analysis of the actual data set are compared with results of eigen values of a factor analysis from the randomly generated dataset. Only factors with eigen values equal or greater than one were retained. In accordance with the literature that suggests that personality traits are better conceptualized as intercorrelated, a Promax with Kaiser rotation was applied to the extracted factors.

Confirmatory factor analysis. In order to further explore the structural model of the PID-5-Sp in the current sample, a CFA was conducted using structural equation modeling in Stata 15 to test the fit of the five-factor model reported by Krueger and colleagues (2012). The model delineates five domains with a range of seven to three facets per domain. Negative Affect is the largest domain and includes the following facets: submissiveness, restricted affectivity, separation insecurity, anxiousness, emotional lability, hostility and perseverance. Suspiciousness, depressivity, withdrawal, intimacy avoidance, and anhedonia made up Detachment and manipulativeness, deceitfulness, callousness, attention seeking, and grandiosity were assigned to Antagonism. Disinhibition was constructed with the irresponsibility, impulsivity, distractibility, rigid perfectionism, and risk-taking scales. Lastly, eccentricity, perceptual dysregulation and unusual beliefs and experiences were assigned to the Psychoticism factor. The factors were allowed to correlate, and each facet was restricted to only load in the factor assigned. Although it is likely that the error terms of the PID-5 facets are significantly intercorrelated at this time there is no data clearly delineating these interactions and therefore

were not included in the initial CFA model. Model fit was evaluated using multiple indices of model fit, which included Chi-squared test, goodness of fit (GFI), Browne and Cudeck's (1993) root mean square error of approximation (RMSEA), the Tucker-Lewis index (TLI), comparative fit statistics (CFI; Bentler, 1990) and the standardized root mean squared residual (SRMSR). In congruence with Fossati and colleagues (2013) methods modification indices of error terms were used to improve model fit.

Pearson's correlations between the five PID-5-SP domains, and normative personality traits as measured by the BFI were conducted to assess for theoretically consistent relationships between normative and maladaptive personality traits. Lastly, multiple linear regressions were used to further evaluate construct validity. Three regression models were created in which depression, anxiety, and aggression scores were independently entered as the dependent variables and were regressed onto all five PID-5-SP domains, which functioned as the independent variables in the model. Simple multiple linear regressions were deemed appropriate given the exploratory basis of the current study and the lack of definitive research on specific PID-5 domain variance accountability in psychopathological outcomes.

Results

Descriptives and Distributions

Means, standard deviations, and skewedness are reported in Table 1. In regard to skewedness, consistent with expectations of collecting data from a mostly normative sample, a number of PID-5 facets and domains exhibited mild to moderate positive skewedness, with Callousness, Depressivity, and Perceptual Dysregulation being the only facets that reached skewedness scores greater than 1. Guidelines developed using computation simulation studies of

Table 1

Descriptive Statistics for PID-5-Sp Facets, Domains and Outcome Measures

	<i>M</i>	<i>SD</i>	α	Skewedness
Anhedonia	1	.69	.86	.61*
Anxiousness (9)	1.36	.75	.88	.17
Attention Seeking (8)	.94	.69	.87	.56*
Callousness (14)	.56	.61	.92	1.31*
Deceitfulness (10)	.78	.66	.88	.83*
Depressivity (14)	.81	.66	.91	1.04*
Distractibility (9)	.99	.78	.92	.59*
Eccentricity (13)	.86	.80	.95	.76*
Emotional Lability (7)	1.21	.76	.85	.37*
Grandiosity (6)	.96	.64	.76	.61*
Hostility (10)	1.02	.68	.88	.53*
Impulsivity (6)	.95	.68	.81	.53*
Intimacy avoidance (6)	.76	.71	.82	.83*
Irresponsibility (7)	.76	.66	.81	.93*
Manipulativeness (5)	.88	.70	.80	.78*
Perceptual Dysregulation (12)	.65	.67	.92	1.18*
Perseveration (9)	1.10	.67	.86	.49*
Restricted Affectivity (7)	1.08	.64	.78	.42*
Rigid Perfectionism (10)	1.24	.69	.88	.27
Risk Taking (14)	1.20	.51	.80	-.07
Separation Insecurity (7)	.87	.77	.88	.63*
Submissiveness (4)	1	.77	.83	.43*
Suspiciousness (7)	1.25	.53	.60	.36*
Unusual Beliefs & Experiences (8)	.72	.70	.87	.92*
Withdrawal (10)	1.06	.74	.91	.43*
Negative Affect (23)	1.14	.66	.93	.33*
Detachment (24)	.94	.62	.94	.52*
Antagonism (21)	.87	.59	.92	.87*
Disinhibition (22)	.90	.62	.94	.61*
Psychoticism (33)	.74	.66	.97	.91*
PHQ-9	7.91	6.88	.92	.82*
GAD-7	6.29	5.89	.94	.86*
AQ-R	26.11	10.61	.93	.75*
BFI-Extroversion	2.90	.86	.82	-.05
BFI-Agreeableness	3.67	.66	.73	-.44*
BFI-Conscientiousness	3.64	.76	.80	-.41*
BFI- Neuroticism	2.79	.91	.84	.30*
BFI- Openness	3.55	.73	.82	-.25

Note. Patient Health Questionnaire- 9 (PHQ-9); Generalized Anxiety Disorder-7 (GAD-7); Aggression Questionnaire- Refined (AQ-R); Big Five Inventory (BFI)

* significant at $p < .05$

estimation methods suggest that skewedness values below 3.0 are acceptable for psychometric purposes (Kline, 2016).

Reliability

Reliability was measured by Cronbach's Alpha coefficients, which assess the internal consistency or relatedness of items in each scale using the number of items and their intercorrelations. As it can be observed in Table 1, most scales exhibited appropriated reliability, as defined by a coefficient of $\geq .70$ (Nunnally, 1978). Alpha coefficients ranged from .60 to .95, with an average of .85. Suspiciousness was the only facet that did not meet the $\geq .70$ cut-off. However, upon closely examining item statistics, it became apparent that deleting item 177, "I rarely feel that people I know are trying to take advantage of me" improved Cronbach's Alpha statistics to .72. Overall, reliability coefficients of 24 PID-5 facets and all five domains were deemed acceptable.

Factor Structure

The structure of the PID-5-Sp was explored using exploratory factor analysis, with Promax Kaiser rotations. The analysis was conducted at the facet level, which included the 25 facets that make up the five PID-5 domains. The Kaiser-Meyer-Olkin index was $>.90$ and Bartlett's statistic was significant at, $p < .00001$. A Parallel analysis suggested the retention of three factors. The eigen values of the first five factors were 13.28, 1.90, 1.10, 0.63, and 0.52. Given that only three factors had eigen values equal or greater than 1 and that an exploration of a fourth and fifth factor did not add meaningful interpretability, three factors were retained. All three factors accounted for 93.3% of the total variance before rotation, with the first factor accounting for the first 76.2% of the variance. After the Promax rotation, which allows for correlations between factors, was applied the first factor accounted for 65.6 % of the variance.

The second and third factor accounted for 48% and 53.2% of the shared variance, respectively. As expected, there were cross loadings, but in the final solution, which is shown in Table 2, facets were assigned to the factor in which they had the highest loading. There were no loadings below .30, which has been a criterion for exclusion in previous studies testing the factor structure of PID-5 translations (Krueger et al., 2012). The final three-factor solution included one large factor that combined the facets of the Negative Affect, Disinhibition, and Psychoticism and two smaller factors that exhibited a similar structure of previously reported domains labeled Antagonism and Detachment in the Krueger et al., (2012) development study.

EFA methods suggested a three- factor structure for the current sample. However, given that previous studies evaluating the structure of other PID-5 translations have suggested a five-factor structure it was decided that it was appropriated to test the model fit of a five-factor model in the current sample. Therefore, a CFA was added to the analyses, testing the model proposed by Krueger et al. (2012) during the development phase of the PID-5, which has been replicated in other translation studies (Bastiaens et al., 2016; Bo et al., 2016; De Fryut et al, 2013; Fossati et al., 2013; Gutierrez et al., 2015; Roskam et al., 2015; Soraya et al., 2017; Zimmermann et al., 2014). The model was the result of EFAs and is composed of five factors that are intercorrelated and facets that are constrained to only load in their designated factors. Error values were constrained to be independent. When this model was applied to the current sample it exhibited poor fit, Satorra-Bentler $\chi^2(265) = 1536.24$, $p < .001$, RMSEA = .13 (90% confidence interval = 0.12-0.13), TLI = 0.79, CFI = .82, and SRMR = .08. All indices, with exception of SRMR which suggested a marginal acceptable fit (.077), failed to meet threshold cut-off points for good model fit.

Table 2
PID-5-Sp Exploratory Factor Analysis

	Factor 1	Factor 2	Factor3
Anhedonia (8)	0.65	-0.29	.45
Anxiousness (9)	0.94	-0.21	-0.01
Attention Seeking (8)	0.19	0.78	-0.18
Callousness (14)	-0.06	0.5	0.59
Deceitfulness (10)	0.06	0.65	0.29
Depressivity (14)	0.65	0.02	0.32
Distractibility (9)	0.7	0	0.2
Eccentricity (13)	0.52	0.27	0.16
Emotional Lability (7)	0.99	0.03	-0.25
Grandiosity (6)	-0.11	0.73	0.1
Hostility (10)	0.53	0.16	0.29
Impulsivity (6)	0.48	0.37	-0.02
Intimacy avoidance (6)	0	-0.09	0.79
Irresponsibility (7)	0.38	0.24	0.35
Manipulativeness (5)	-0.06	0.79	0.16
Perceptual Dysregulation (12)	0.43	0.37	0.25
Perseveration (9)	0.7	0.13	0.14
Restricted Affectivity (7)	-0.06	0.11	0.74
Rigid Perfectionism (10)	0.54	0.21	0.05
Risk Taking (14)	-0.01	0.58	-0.22
Separation Insecurity (7)	0.64	0.28	-0.07
Submissiveness (4)	0.6	0.14	-0.01
Suspiciousness (7)	0.59	-0.02	0.18
Unusual Beliefs & Experiences (8)	0.37	0.48	0.06
Withdrawal (10)	0.31	-0.22	0.78

Note. Bold numbering denotes highest loading and factor to which facet was assigned after Promax with Kaiser rotation was applied

Modification indices were reviewed to improve model fit. Given that the goal of the study was to test a previously existing model in the current sample, only covariances between error terms within factors were considered. After careful consideration, covariances for the following pairs were added: Restricted Affectivity with Emotional Lability, Intimacy Avoidance with Withdrawal, Impulsiveness with Risk Taking, Irresponsibility and Rigid Perfectionism and Depressivity with Withdrawal. After the inclusion of these covariances the model fit indices marginally improved Satorra-Bentler $\chi^2(265) = 1263.70$, $p < .001$, RMSEA = .11 (90% confidence interval = 0.11-0.12), TLI = .83, CFI = .86, and SRMR = .07, but not meaningfully.

Construct Validity

Table 3 shows Pearson's correlations between the five PID-5-Sp domains and BFI scores, which measure the Big Five normative personality traits. As expected, the five PID-5-Sp domains exhibited moderate to strong intercorrelations, ranging from .45-.77, with an average of .65 and a median of .68. Normative BF traits were also intercorrelated, but to a lesser extent than the five PID-5-Sp domains (using absolute values, $r = .20-.53$, $p < .001$; $M = .39$; $Mdn. = .37$). Overall, the scales exhibited mostly theoretically congruent moderate to strong correlations. Negative Affect correlated negatively with four of the five normative personality traits, exhibiting a strong positive relationship with Neuroticism ($r = .78$, $p < .001$), which is supposed to measure a similar construct, within the normative Big Five model. Detachment was moderately correlated with Extroversion ($r = -.50$, $p < .001$), but also exhibited a strong correlation with Agreeableness ($r = -.61$, $p < .001$). Antagonism had a moderate correlation of $r = -.42$, $p < .001$ with Agreeableness, which was the strongest among all BFI traits, denoting congruence with the maladaptive personality model. Disinhibition also had the strongest correlation with its normative counterpart, Conscientiousness ($r = -.69$, $p < .001$). Regarding Psychoticism, consistent with the literature it had smaller correlations with normative personality traits than with the maladaptive traits and ultimately was not significantly correlated with Openness to Experience ($r = -.08$, $p > .05$).

Multiple linear regressions were used to further investigate the PID-5-Sp ability to measure variance in depression, anxiety, and aggression. Results are presented in Table 4. When depression was regressed on all five PID-5-Sp domains, the model accounted for 66% of the variance in depression. Negative Affect ($\beta = .39$, $p < .001$) and Detachment ($\beta = .39$, $p < .001$) independently accounted for the most variance when controlling for the remaining four domains.

Table 3

Two-Tailed Pearson's Correlations of PID-5-SP Domains and BFI Traits

	1	2	3	4	5	6	7	8	9	10
1.Negative Affect	-									
2.Detachment	.63**	-								
3.Antagonism	.48**	.45**	-							
4.Disinhibition	.76**	.70**	.61**	-						
5.Psychoticism	.72**	.67**	.68**	.77*	-					
6.Extroversion	-.30**	-.50**	.10	-.23**	-.13*	-				
7.Agreableness	-.46**	-.61**	-.42**	-.58**	-.51**	-.37**	-			
8.Conscientiousness	-.50**	-.49**	-.28**	-.69**	-.43**	.32**	.53**	-		
9.Neuroticism	.78**	.52**	.24**	.61**	.51**	-.38**	-.48**	-.53**	-	
10.Openness	-.21**	-.23**	-.04	-.22**	-.08	.34**	.33**	.37**	-.20**	-

Note. ** denotes significant at $p < 0.001$; * significant at $p < .05$

Anxiety was best predicted by Negative Affect ($\beta = .61, p < .001$), but also had significant relationships with Detachment ($\beta = .12, p < .01$), Antagonism ($\beta = -.13, p < .001$), and Psychoticism ($\beta = .17, p < .001$), which together accounted for 65% of the variance GAD-7 scores. All five PID-5-Sp domains accounted for 69% of the variance in aggression scores. As expected, aggression was positively associated with Antagonism ($\beta = .14, p < .001$), but it had the strongest association with Negative Affect ($\beta = .30, p < .001$). Interestingly, Antagonism was negatively associated with both depression ($\beta = -.19, p < .001$) and anxiety ($\beta = -.13, p < .001$). Overall, Negative Affect appeared to be the stronger predictor in all three outcome measures.

Discussion

The purpose of the current study was to continue expanding the understanding of the PID-5 and its applicability to other cultural groups, particularly Spanish speaking Hispanics living in the United States, who are often underserved. Given that the PID-5 is theoretically founded in personality trait theory it was important to test its structure replicability as it is transferred to another population. This is particularly relevant because Krueger and colleagues (2012) proposed it to be used in the future diagnosis of personality disorders and as a building block for personality research.

Table 4
Regressions between PID-5 Domains and Outcomes

PID-5-Sp Domains	Depression (PHQ-9)			Anxiety (GAD-7)			Aggression (AQ-R)		
	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β	<i>B</i>	<i>SE B</i>	β
Negative Affect	4.11**	.58	.39**	5.47**	.50	.61**	4.83**	.85	.30**
Detachment	3.36**	.55	.30**	1.15*	.48	.12*	2.79**	.81	.16**
Antagonism	-2.18**	.55	-.19**	-1.27**	.48	-.13**	2.48**	.81	.14**
Disinhibition	1.38**	.72	.20**	.42	.63	.04	4.27**	1.06	.25**
Psychoticism	1.01*	.65	.14*	1.48**	.57	.17**	1.82	.96	.12
R ²	.66			.65			.69		

Note. ** denotes significant at $p < 0.001$; * significant at $p < .01$

The current study replicated, in part, findings of past PID-5 translations studies, but failed to replicate the five-factor structure that has been found in the development article and other translations studies (Bastiaens et al., 2016; Bo et al., 2016; De Fryut et al., 2013; Fossati et al., 2013; Gutierrez et al., 2015; Roskam et al., 2015; Soraya et al., 2017; Zimmermann et al., 2014). In congruence with previous studies, EFA methods were used and extraction analyses suggested a three factors structure, which has been suggested in other samples. In fact, in the case of the original article (Krueger et al., 2012) MAP analyses suggested retaining three factors, parallel analysis suggested six, but only five factors had eigen values that were significantly greater than one. Gutierrez and colleagues (2015) confronted similar results when deciding the number of factors to retain in the PID-5 Spanish translation study. However, as the five-factor structure was explored in their samples, five factors appeared to represent the data most consistently. In the current sample a fourth and fifth factor did not add meaningful information.

Therefore, the most appropriate factorial structure for the current sample was a three-factor-solution, which upon close inspection resembled factors reported in other PID-5 solutions. More specifically, factor two and three mapped onto PID-5 factors that have been previously conceptualized as Antagonism and Detachment, respectively. The interpretation of the first factor was more complicated because it lumped facets that normally load onto Negative Affect, Disinhibition, and Psychoticism into one large factor. In this factor, the facets of Negative Affect which denote a tendency towards experiencing a wide range of intense negative emotions showed the higher loadings, followed by Disinhibition facets that measure impulsiveness and thoughtless reactivity to the environment. The Psychoticism facets, which measure incongruity and a separation from normative experiences had the lowest loadings in this factor. The information combined in factor 1 create a factor that could be conceptualized as a measure of intense negative emotionality combined with poor resourcefulness to reacting and coping to internal experiences that lead to significant social impairment.

The large amount of variance accounted for by factor 1 and its breadth of conceptual coverage is reminiscent of the demoralization factor that was extracted from the MMPI-2 Clinical scales in creation of the Restructured Scales (RC) (Tellegen et al., 2003). Hoelzle and Meyer (2008) found that the RC scales, which sought to improve the clinical specificity of the Clinical scales, had a five-factor structure, but the clinical scales themselves were better suited for a three-factor structure with one large complex factor accounting for most of the variance in the model (47%) and two smaller factors measuring somatic complains (11%) and a impulsivity/heightened energy (8%) accounting for a smaller percentage of the remaining variance. Although the content of the PID-5-Sp factors does not completely map onto the Clinical scales, the variance partition is similar, which might speak to the high pathological

saturation in the PID-5 scales. Using this model as a reference, factor 1 could also be conceptualized as one large global factor measuring general distress in the presence of significant personality psychopathology.

Interestingly the three factors that emerged in the current sample are conceptually similar to those presented in Eysenck's personality model, which included three factors: Neuroticism, Extroversion/ Introversion, and Psychoticism (Eysenck & Eysenck, 1975). The first two are conceptually similar to their respective traits in the FFM and would map onto factor one and three in the current sample, but in a maladaptive pathological presentation. Psychoticism as defined by Eysenck (1975) has a broader definition, both capturing what the field currently defines as psychotic (reality testing), while also adding a component of psychopathy, which in this case would be conceptually related to the second factor in the PID-5-SP which measures Manipulativeness, Grandiosity, Deceitfulness, and Risk Taking. The overlap of the current factor solution with "Big Five" and "Big Three" personality trait theories speaks to the convergence of theories in personality assessment and the importance of creating clearly defined models of personality that can communicate information consistently across populations.

Consistency and replicability can be difficult in personality research due to the complexity of personality models that often include cross loadings, high intercorrelations among factors, and statistical variability across samples (Watters & Bagby, 2018). Although attempts were made to test model fit of a five-factor structure using CFA methods, our efforts were hindered by the lack of definitive knowledge about the expected covariances between error terms at the facet level. Fossati and colleagues (2013) attempted to conduct CFAs in a sample testing the Italian translation of the PID-5. They tested three different models in which they varied the independence of cross loadings of facets and the covariance of error terms. The CFA model that

exhibited the best fit among the tested models was the one in which all facets were only allowed to load in their assigned factor and included the correlation of error terms based on modification indices, Satorra-Bentler $\chi^2(192) = 564.63$, $p < .001$, RMSEA = .06 (90% confidence interval = 0.05-0.07), TLI = .98, CFI = .99, and SRMR = .06, AIC = 830.63. Due to high intercorrelations in the PID-5 domains, it is expected that there will be cross loadings of facets and correlations among error terms. Therefore, it follows that the model with the best fit indices will include covariances of error terms, but due to some instability in the PID-5 at the facet level these have not been defined in the literature. This means that researchers wanting to conduct CFAs testing the more complex PID-5 models are left to the mercy of modification indexes. This is a good starting point, but ideally, covariances added to CFA models using modification indexes should be made thoughtfully and be theory driven as to not artificially inflate the fitness of the model with non-relevant covariances (Hermida, 2015). Unfortunately, Fossati et al. (2013) did not report the covariances in error term that they included in their PID-5 model. So, it was impossible to replicate or review their error covariances. However, even after allowing for the covariance of error terms that appeared conceptually logical the five-factor structure remained a poor fit for the current sample. It is likely that allowing for more covariances of error terms and the deletion of items would have improved the model fit, but the focus of the current study was of an exploratory nature, so a conservative approach was taken to corrections made using modification indices.

A factor that must be considered while interpreting the current factor solution is that this is the first PID-5 translation study to use MTurk to collect data. This gives the current sample unique qualities when compared to the other studies that make up the PID-5's body of research. In consistency with previous research using MTurk samples, the reliability of scales in the

current sample was adequate and the demographic presentation of the sample was similar to what has been observed in other MTurk samples (Behrend et al., 2011; Buhrmester et al., 2011; Kees et al., 2017). However, this was a 100% Hispanic, Spanish speaking sample responding to a relatively long survey, which might have affected how participants responded to the questions in the study. Research has shown that MTurk participants tend to multitask more than other samples, which could potentially lead to less attention to detail (Kees et al., 2017), reducing the amount of variance available to parse out using EFA methods making it more likely that results would show a large factor composed of the presence or absence of general distress. Therefore, it is likely that the three-factor solution is a combination of the uniqueness of the diverse Hispanic sample and some of the traits that are part of the individuals that make up the MTurk community.

Ultimately, even though the five-factor structure was not replicated by the PID-5-Sp, PID-5-SP domains still exhibited expected interactions with normative personality traits and psychopathology outcomes. When assessing the interaction between PID-5 domains and normative personality traits Sleep and colleagues (2018) found that PID-5 domains were significantly correlated with most FFM traits, as measured by 60 items of the IPIP, which is an open access normative personality inventory that has demonstrated good congruence with the NEO-PI. Therefore, multiple correlations between PID-5-Sp domains and FFM traits were expected in the current study, particularly because both the PID-5-Sp and the BFI were created to measure personality traits. Domains that belong to the same dimension are expected to correlate to a higher extent than traits that belong to different dimensions. This hypothesis has been supported by previous studies in which all four FFM traits with exception of Openness to Experience correlated with their theoretically congruent PID-5 domains (Sleep et al., 2018). Similar results were obtained in the current study, with the main difference being that

Detachment had the highest correlation ($r = -.61, p < .001$) with Agreeableness and not Extroversion, which is its theoretical counterpart. However, Detachment still exhibited a moderate correlation of $r = -.50, p < .001$ with Extroversion, indicating that a moderately strong relationship exists between these two traits.

The predictive ability of PID-5 domains was explored by regressing psychopathology outcomes onto PID-5-SP domains. The PID-5 has been theorized to have a hierarchical structure (Wright et al., 2012), which has been replicated in various translations studies (Fossati et al., 2013; Krueger & Markon, 2014; Roskam et al., 20015). At the second level, the hierarchy breaks into internalizing and externalizing factors, which are further divided into the PID-5 domains, with all five facets emerging at the fifth level. Based on this structure it would be expected that internalizing disorders such as depression and anxiety would be better predicted by the two internalizing domains, Detachment and Negative Affect (Wright et al., 2012). The remaining PID-5 domains should then be better predictors of externalizing traits, which are often found in substance use disorders, aggression, and other disorders that include overt dysfunctional behaviors. In the case of depression, the current study supported this theoretical conjecture, Negative Affect and Detachment were the best predictors of PHQ-9 scores when holding all other domains constant. Zimmermann and associates (2014) found similar results with the added benefit of conducting their analyses at the facet level. They found that anxiousness and emotional lability, which are both central facets of Negative Affect had incremental associations (using part correlations) with depressive disorders. The Detachment anhedonia and depressivity facets also exhibited incremental association with depression scores. As for GAD-7 scores, Negative Affect was by far the most effective predictor of anxiety scores in the current sample with $\beta = .61, p < .001$. Theoretical congruence was less clear for aggression scores,

Antagonism, and Disinhibition did exhibit significant predictive power, but Negative Affect was the strongest predictor ($\beta = .30, p = < .001$). Other studies examining the relationships between PID-5 domains and psychopathology outcomes have found similar results (Few et al., 2013; Sleep et al., 2018; Zimmermann et al., 2014). Interestingly, in all of these studies Negative Affect consistently emerged as a significant predictor for aggression and other externalizing disorders. Although this could be in part attributed to the high shared variance of PID-5 domains it would be an interesting relationship to continue exploring, especially since one of the reasons as to why the PID-5 has a large number of facets is to increase specificity in the measurement of personality pathology.

Consistent with the literature, the relationship between Psychoticism and Openness to Experience continues to be elusive. Psychoticism correlated significantly with all FFM traits, with exception of OE, but this lack of correlational relationship has been a consistent finding in the literature (Few et al, 2013; Sleep et al., 2018; Zimmermann et al., 2014). In regard to its relevance as a predictor of pathological disorders, results from previous studies contrast with the current findings, often reporting significant bivariate Pearson's correlations with both internalizing and externalizing psychopathology factors, that become non-significant once the shared variance is accounted for by including all five PID-5 domains in regression models. However, in the current sample Psychoticism continues to be a significant predictor of depression and anxiety scores, even after controlling for the variance accounted for by Detachment, Disinhibition, Negative Affect, and Antagonism. A study looking at the incremental associations of Psychoticism facets in depression and anxiety scores found that perceptual dysregulation and eccentricity added significant predictive value (Zimmermann et al., 2014). These findings are indicative that even if OE and Psychoticism do not directly map onto

each other and might even measure different aspects of a construct, Psychoticism remains an important construct in the measurement of psychopathological traits (Quilty et al., 2013).

Overall, PID-5-Sp domains show theoretically congruent relationships with FFM normative personality traits and psychopathological outcomes. The five-factor structure was not replicated in the current sample, but the factors that did emerge showed some consistency to previously defined maladaptive personality models, which speaks to the complexity of personality as a construct, especially when taking into account cultural factors, which impact how individuals interact with their environment.

The five-factor PID-5 model has been replicated multiple times across different U.S. and European samples (Somma et al., 2019). These studies have relied primarily on EFA methods due to the exploratory nature of translation studies, and the general complexity of personality models. This has allowed for flexibility in the interpretation of the number of factors to be extracted, with a number of studies suggesting anywhere between 3 to 6 factors (Bastiaens et al., 2016; Bo et al., 2016; De Fryut et al., 2013; Fossati et al., 2013; Krueger et al., 2012; Gutierrez et al., 2015; Roskam et al., 2015; Soraya et al., 2017; Zimmermann et al., 2014). Researchers ultimately arrive to the number of factors to be retained using both statistical analysis and preconceptions derived from the FFM. The FFM gives researchers a structure from which they can interpret research findings, but it can also bias researchers into interpreting their data according to the etic model that they are importing rather than mindfully capturing the essence of the studied cultural group (De Raad et al., 2010). In fact, De Raad and colleagues (2010) reviewed 14 taxonomies of psycholexical studies measuring personality traits across cultures that included European, Asian, and North American samples. Using congruency coefficients, they concluded that when assessing the replicability of the five factors structure it was only the first

three factors that actually achieved strong congruency across all cross-cultural samples. They identified extroversion, agreeableness, and conscientiousness as the most replicable traits. Surprisingly, neuroticism did not emerge as an independent cross-culturally replicable factor. The authors acknowledge that this was likely due to that nature of psycholexical studies that focus on attributes that are most often talked about within the social context of the culture of study. This approach would limit the availability of negative valence available for measurement because most cultures do not use terms denoting psychopathology in their day-to-day interactions. Their criticism of the replicability of the five-factor model across cultures brings up a relevant caveat in the development of the FFM, which was initially developed from psycholexical studies in European samples (Digman, 1990) and as such might show measurement limitations in non-Western samples. In fact, the congruency coefficients of the Filipino sample included in their study exhibited the lowest values when compared with the European and North American samples, indicating a lower conceptual replicability.

Culturally reliable and valid measures include both statistical congruence, and integration of culturally relevant phenomena (Dana, 2015). This is particularly relevant for groups in which one would expect significant cultural differences to exist. Unfortunately, research on the cultural effects of importing etic measures into non-western cultures is limited. However, we do know that when emic measures are used in non-Western and Latin samples measurement reliability and validity increases (Ledesma et al., 2011; Zhang & Bond, 1998). Therefore, even though the FFM might replicate appropriately and provide useful information about other cultures, it might not paint the full picture, ignoring cultural identity and societal concepts that are relevant to the culture of study. This becomes particularly relevant for Hispanic samples which are not only a

heterogeneous group in their nationalities, but also exhibit different levels of acculturation (Dana, 2015).

Research in Mexico and in Mexican Americans living in the United States has shown that retention of Mexican culture (as measured by the Acculturation Rating Scale for Mexican Americans) significantly affects MMPI-2 scores, often leading to elevations in validity scales and scale 9 or the hypomania scale (Whitney, 2002). The differences are of less than a standard deviation, but little to no research exists about the clinical relevance, if any, of these differences.

Ultimately, the fact that these inconsistencies exist in the most commonly used measure of personality psychopathology, which has 70 years of existence and research, should serve as a word of caution in the development of the PID-5, which has shown good structural replicability across U.S. and European samples (Somma et al., 2019), but lacks research regarding its replicability in non-Western samples and its sensitivity to cultural effects. Taking into account the results of the current study, it would be beneficial to continue exploring the structure of maladaptive personality traits in Latin America, and other areas of the world in which cultural identity has been forged from the interaction of European colonizers and native civilizations. This would ensure that the PID-5 is used to the best of its ability in a culturally competent manner.

Limitations

Results presented in this study are promising for the use of the PID-5-Sp. However, there are some considerations and limitations that must be taken into account as the findings are integrated into the body of the PID-5 research literature. Although it is common practice to test the reliability and validity of translations of previously developed instruments using normal samples, an argument can be made that it would be most ideal to test the measure in a sample of

the population that it is intended for, in this case, Spanish speaking individuals living in the U.S. who exhibit maladaptive personality traits or have a history of psychopathology. This would likely improve the skewedness of the distributions and add more theoretically relevant variance to scale scores, improving the study's ability to measure the statistical properties of the measure in the manner that it is intended to be used. However, these samples tend to be less accessible, and the use of normal samples provides researchers with acceptable foundational knowledge from which one can draw hypotheses that can later be tested in samples with the desired specificity. Given the early stages of development of the translation of the PID-5-Sp and the exploratory nature of the study the use of this sample was adequate. Results provided information about the characteristics of the PID-5-Sp in a non-clinical sample, information that can later be used as a comparison point against its performance in clinical samples.

The current sample was collected using MTurk, which is a platform in which participants complete tasks with the purpose of compensation. Previous studies have shown that data collected in this site is comparable to college samples and other online sources (Behrend et al., 2011; Buhrmester et al., 2011), but similar to any data collection methods using self-report measures one must take into consideration impression management and the validity of responses. Another area in which improvements could have been made is in the specificity of the population studied. The target population of the current study were Spanish speaking Hispanics living in the U.S. However, this ethnic group is heterogeneous in both countries of ancestral origin and levels of acculturation. It is therefore, likely that the current study would have benefitted from identifying a smaller target group, such as Mexican Americans or Puerto Ricans. This would have allowed for more specificity in the vocabulary used in the translation presented in the PID-5-Sp. But, ultimately the PID-5-Sp has shown promising results in its ability to predict

psychopathological outcomes, which builds hope that one translation might be adequate for a diverse sample of Hispanics. These limitations ultimately open a sea of possibilities for future research.

Future Research and Conclusions

This translation of the PID-5 is a starting point for the development of a Latin American Spanish translation of the PID-5, as such, future research should focus on testing its properties in clinical samples and samples collected in Latin American countries. As that research base is expanded the use of CFA and IRT methods could further clarify the structure of the PID-5-Sp. Along with a move towards confirmatory analyses, the study of the PID-5-Sp would benefit from being tested in its ability to predict scores of psychopathological concepts that are emic to the culture of study. This would expand not only on the relevance of the PID-5 as cross-cultural measure of personality traits, but also add information on its ability to capture native culturally relevant concepts. Lastly, as a measure for Hispanics in the U.S., future studies should research the effects of acculturation in PID-5-Sp scores.

In conclusion, the PID-5 has the potential to become a powerful tool to measure, diagnose, and aid in the development of treatments of mental health disorders across different cultural groups, especially since translations in European and U.S. samples have shown promising results of replicability (Somma et al., 2019). However, further research is needed to further understand the effects of using the PID-5 to measure personality psychopathology in Latin American countries. A solid understanding of the PID-5-Sp as a measure of psychopathology has the potential to improve research and treatment of Spanish speaking Hispanic populations, aiding in the reduction of mental health racial disparities in the provision of mental health services to underserved populations.

References

- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47*.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders: DSM-III*. Washington, D.C.: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, D.C.: American Psychiatric Association.
- Anderson, A., Snider, S., Sellbom, M., Krueger, R., & Hopwood, C. (2014). A comparison of the DSM-5 section II and section III personality disorder structures. *Psychiatry Research, 216*, 363-372.
- Anderson, J. L., Sellbom, M., Bagby, R. M., Quilty, L. C., Veltri, C. O., Markon, K. E., & Krueger, R. F. (2013). On the convergence between PSY-5 domains and PID-5 domains and facets: Implication for assessment of DSM-5 personality traits. *Assessment, 20*(3), 286-294.
- Arditte, K. A., Cek, D., Shaw, M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment, 28*(6), 684-691.
- Arnau, R. C., Handel, R. W., Archer, R. P., Bisconer, S., & Gross, D. (2004). External correlates for the MMPI-2 personality psychopathology five (PSY-5) facet subscales in an adult inpatient psychiatric sample. *Annual Symposium on Recent Developments in the MMPI-2/MMPI-A*. Minneapolis.

- Bastiaens, T., Smits, D., De Hert, M., Vanwalleghem, D., & Claes, L. (2016). DSM-5 section III personality traits and section II personality disorders in a Flemish community sample. *Psychiatry Research, 238*, 290-298.
- Behrend, T. S., Sharek, D., Meade, A., & Wiebe, E. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods, 43*(3), 800-813.
- Benet-Martinez, V., & John, O. P. (1998). Los cinco grandes across cultures and ethnic groups: Multitrait multimethod analyses of the big five in Spanish and English. *Journal of Personality and Social Psychology, 75*(3), 729-750.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246.
- Berry, J. W. (1969). On cross-cultural comparability. *International Journal of Psychology, 4*, 119-128.
- Bo, S., Bach, B., Mortensen, E. L., & Simonsen, E. (2016). Reliability and hierarchical structure of DSM-5 pathological traits in a Danish mixed sample. *Journal of Personality Disorders, 30*(1), 112-129.
- Bogart, L. M., Revenson, T. A., Whitfield, K. E., & France, C. R. (2014). Introduction to the special section on lesbian, gay, bisexual, and transgender (LGBT) health disparities: Where we are and where we're going. *Annals of Behavioral Medicine, 47*(1), 1-4.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression. A measurement model for the Buss-Perry aggression questionnaire. *Journal of Research in Personality, 35*, 138-167.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? In A. E. Kazdin, *Methodological issues and strategies in clinical research*, (pp. 133-139). Washington, D.C.: American Psychiatric Association.
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27*, 509-540.
- Butcher, J. N., Dahlstrom, W. G., Tellegen, A. M., & Kreammer, B. (1989). *The Minnesota Multiphasic Personality Inventory-3 (MMPI-2) Manual of administration and scoring*. Minneapolis, MN: University of Minneapolis Press.
- Calvo, N., Valero, S., Saez-Francas, N., Gutierrez, F., Casas, M., & Ferrer, M. (2016). Borderline personality disorder and Personality Inventory of DSM-5 (PID-5): Dimensional personality assessment with DSM-5. *Comprehensive Psychiatry, 70*, 105-111.
- Carlotta, D., Krueger, R. F., Markon, K. E., Borroni, S., Frera, F., Somma, A., . . . Fossati, A. (2015). Adaptive and maladaptive personality traits in high-risk gamblers. *Journal of Personality Disorders, 29*(3), 378-392.
- Cattell, R. (1978). *The scientific use of factor analysis*. New York: Plenum.

- Cattell, H. E., & Mead, A. D. (2008). The sixteen personality factor questionnaire (16PF). In G. J. Boyle, G. Matthews, & D. H. Saklofske, *The SAGE handbook of personality theory and assessment: Personality measurement and testing*. (pp. 135-159). London, UK: SAGE Publications Ltd.
- Church, A. T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality, 69*(6), 979-1006.
- Costa, P. T., & McCrae, R. R. (1976). Age differences in personality structure: A cluster analytic approach. *Journal of Gerontology, 31*(5), 564-570.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., McCrae, R. R., & David, A. D. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences, 12*(9), 887-898.
- Crego, C., Gore, W. L., Rojas, S. L., & Widiger, T. A. (2015). The discriminant (and convergent) validity of the Personality Inventory for DSM-5. *Personality Disorders: Theory, Research, and Treatment, 6*(4), 321-335.
- Cupani, M., Garrido, S., & Tavella, J. (2013). El modelo de los cinco factores de personalidad: Contribucion predictiva al rendimiento academico. *Revista de Psicologia, 9*(17), 67-86.

- Dana, R. H. (2015). A personality approach to testing Hispanics. In K. F. Geisinger, *Psychological testing of Hispanics: Clinical, cultural, and intellectual issues* (2nd ed., pp. 189-213). American Psychological Association.
- De Fryut, F., De Clercq, B., De Bolle, M., Wille, B., Markon, K., & Krueger, R. F. (2013). General and maladaptive traits in a five-factor framework for DSM-5 in an university student sample. *Assessment, 20*(3), 295-307.
- De Raad, B., Barelds, D.P., Levert, E., Mlacic, B., Hrebickova, M., Szarota, P, ... Katigbak, M. S. (2010). Only three factors of personality description are fully replicable across languages. A comparison of 14 trait taxonomies. *Journal of Personality and Social Psychology, 98*(1), 160-173.
- Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology, 41*, 417-440.
- Digman, J. M. (2002). Historical antecedents of the Five-Factor Model. In P. T. Costa, & T. A. Widiger, *Personality disorders and the Five-Factor Model of personality* (pp. 17-22). Washington, DC: American Psychological Association.
- Eysenck, H. J., & Eysenck, S. B. (1975). *Psychoticism as a dimension of personality*. New York: Crane, Russak & Co.
- Fabrigar, L., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 271-299.

- Familiar, I., Ortiz-Panozo, E., Hall, B., Vieitez, I., Romieu, I., Lopez-Ridaura, R., & Lajous, M. (2015). Factor structure of the Spanish version of the Patient Health Questionnaire -9 in Mexican women. *International Journal of Methods in Psychiatric Research, 24*(1), 74-82.
- Few, L. R., Miller, J. D., Rothbaum, A. O., Meller, S., Maples, J., Terry, D. P., . . . MacKillop, J. (2013). Examination of the section III DSM-5 diagnostic system for personality disorders in an outpatient clinical sample. *Journal of Abnormal Psychology, 122*(4), 1057-1069.
- First, M. B., Bell, C. C., Cuthbert, B., Krystal, J. H., Malison, R., Offord, D. R., . . . Wisner, K. L. (2002). Personality disorders and relational disorders: A research agenda for addressing crucial gaps in DSM. In D. J. Kupfer, M. B. First, & D. A. Regier, *A Research agenda for DSM-V* (pp. 123-200). Washington, D.C.: American Psychiatric Association.
- Fossati, A., Krueger, R. F., Markon, K. E., Borroni, S., & Maffei, C. (2013). Reliability and validity of the Personality Inventory for DSM-5 (PID-5): Predicting DSM-IV personality disorders and psychopathy in community-dwelling Italian adults. *Assessment, 20*(6), 689-708.
- Fournier, M. A., Di Domenico, S. I., Weststrate, N. M., Quitasol, M. N., & Dong, M. (2015). Toward unified science of personality coherence. *Canadian Psychology, 56*(2), 253-262.
- Gallardo-Pujol, D., Kramp, U., Garcia-Forero, C., Perez-Ramirez, M., & Andres-Pueyo, A. (2006). Assessing aggressiveness quickly and efficiently: The Spanish adaptation of Aggression Questionnaire-Refined version. *European Psychiatry, 21*, 487-494.
- Godfray, H. C. (2002). Challenges for taxonomy. *Nature, 417*, 17-19.

- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology, 2*, 141-165.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213-224.
- Gorsuch, R. L. (1997). Exploratory factor analysis. Its role in item analysis. *Journal of Personality Assessment, 68*(3), 532-560.
- Graham, J. R. (2012). *MMPI-2 assessing personality and psychopathology*. New York, NY: Oxford University Press.
- Grigoras, M., & Wille, B. (2017). Shedding light on the dark side: Associations between the dark triad and the DSM-5 maladaptive trait model. *Personality and Individual Differences, 104*, 516-521.
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures. Literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46*(12), 1417-1432.
- Gutierrez, F., Aluja, A., Peri, J. M., Clavo, N., Ferrer, M., Bailles, E., . . . Krueger, R. F. (2015). Psychometric properties of the Spanish PID-5 in a clinical and community sample. *Assessment, 24*(3), 326-336.
- Haigler, E. D., & Widiger, T. A. (2001). Experimental manipulation of NEO-PI-R items. *Journal of Personality Assessment, 77*(2), 339-358.

- Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The personality psychopathology five (PSY-5): Constructs and MMPI-2 scales. *Psychological Assessment, 7*(1), 101-114.
- Hermida, R. (2015). The problem of allowing correlated errors in Structural Equation Modeling: Concerns and considerations. *Computational Methods in Social Sciences, 3*(1), 25-42.
- Hoelzle, J. B., & Meyer, G. J. (2008). The factor structure of the MMPI-2 Restructured Clinical (RC) scales. *Journal of Personality Assessment, 90*(5), 443-455.
- Hopwood, C. J., Thomas, K. M., Markon, A. G., Markon, K. E., Wright, A. G., & Krueger, R. F. (2012). DSM-5 personality traits and DSM-IV personality disorders. *Journal of Abnormal Psychology, 121*(2), 424-432.
- John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of research of trait taxonomic research. *European Journal of Personality, 2*, 171-203.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The "Big Five" inventory - versions 4a and 54*. University of California, Berkeley: Institute of Personality and Social Research.
- Kees, J., Berry, C., Burton, S., & Sheehan, K. (2017). An analysis of data quality. Professional panels, student subject pools, and Amazon's Mechanical Turk. *Journal of Advertising, 46*(1), 141-155.
- Kline, P. (1994). *An easy guide to factor analysis*. New York, NY: Routledge.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*, (4th Ed). New York, NY: The Guilford Press.

- Krueger, R. F., & Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology, 10*, 477-501.
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*, 1879-1890.
- Ledesma, R. D., Sanchez, R., & Diaz-Lazaro, C. M. (2011). Adjective checklist to assess the Big Five personality factors in the Argentine population. *Journal of Personality Assessment, 93*(1), 46-55.
- McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy:" Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*(3), 710-721.
- McDougall, W. (1926). The chemical theory of temperament applied to introversion and extroversion. *Journal of Abnormal and Social Psychology, 24*(3), 293-309.
- McDougall, W. (1932). Of the words character and personality. *Journal of Personality, 1*(1), 3-16.
- Merz, E. L., Malcarne, V. L., Roesch, S. C., Riley, N., & Sadler, G. R. (2011). A multigroup confirmatory factor analysis of the Patient Health Questionnaire-9 among English-and Spanish-speaking Latinas. *Cultural Diversity and Ethnic Minority Psychology, 17*(-3), 309-316.

- Mills, S. D., Fox, R. S., Malcarne, V. L., Roesch, S. C., Champagne, B. R., & Sadler, G. R. (2014). The psychometric properties of the Generalized Anxiety Disorder -7 scale in Hispanic Americans with English or Spanish language preference. *Cultural Diversity and Ethnic Minority Psychology, 20*(3), 463-468.
- Nielson, M. B., Glaso, L., & Einarsen, S. (2017). Exposure to workplace harassment and the Five Factor Model of personality: A meta-analysis. *Personality and Individual Differences, 104*, 195-206.
- Norman, T. (1967). *2,800 personality trait descriptors: Normative operating characteristics for a university population*. Department of Psychology, University of Michigan.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality traits. *Journal of Abnormal and Social Psychology, 66*(6), 574-583.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill
- Ortiz, F. A. (2015). Psychological testing of Hispanics: Clinical, cultural, and intellectual issues. In F. A. Ortiz (Ed.), *Evaluating and measuring Mexican personality: Etic and cross-cultural perspective*. (2nd ed., pp. 29-50). DC, Washington, United States: American Psychological Association.
- Ortman, J. M., & Shin, H. B. (2011). *Language projections: 2010 to 2020*. Population Division U.S. Census Bureau, Las Vegas.

- Pavlick, E., Post, M., Irvine, A., & Callison-Burch, C. (2014). The language demographic of Amazon Mechanical Turk. *Transactions of the Association for Computational Linguistics, 2*, 79-92.
- Pohling, R., Bzdok, D., Eigenstetter, M., Stumpf, S., & Strobel, A. (2016). What is ethical competence? The role of empathy, personal values, and the five-factor model of personality in ethical decision-making. *Journal of Business Ethics, 137*, 449-474.
- Quilty, L. C., & Bagby, R. M. (2007). Psychometric and structural analysis of the MMPI-2 personality psychopathology (PSY-5) facet subscales. *Assessment, 14*(4), 375-384.
- Quilty, L. C., Ayearst, L., Chmielewski, M., Pollock, B. G., & Bagby, R. M. (2013). The psychometric properties of the Personality Inventory for DSM-5 in an APA DSM-5 field trial sample. *Assessment, 20*(3), 362-369.
- Reyes, E., Alvarez, C., Paredo, A., Sandoval, A. M., & Rebolledo, I. M. (2014). Psychometric properties of the Big Five inventory in a Mexican sample. *Salud Mental, 37*(6), 491-497.
- Robinson, D. N. (1995). *An intellectual history of psychology*. (3rd ed.). London: The University of Wisconsin.
- Roskam, I., Galdiolo, S., Hansenne, M., Massoudi, K., Rossier, J., Gicquel, L., & Rolland, J. P. (2015). The psychometric properties of the French version of the Personality Inventory for DSM-5. *PLoS ONE, 10*(7), 1-13.
- Russell, T. D., Pocknell, V., & King, A. R. (2017). Lesbians and bisexual women and men have higher scores on the Personality Inventory for the DSM-5 (PID-5) than heterosexual counterparts. *Personality and Individual Differences, 110*, 119-124.

- Samuel, D. B., Hopwood, C. J., Krueger, R. F., Thomas, K. M., & Ruggero, C. J. (2013). Comparing methods for scoring personality disorder types using maladaptive traits in DSM-5. *Assessment, 20*(3), 353-361.
- Sentell, T., Shumway, M., & Snowden, L. (2007). Access to mental health treatment by English language proficiency and race/ethnicity. *Journal of General Internal Medicine, 22*(2), 289-293.
- Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science, 1*(2), 213-220.
- Skodol, A. E., Bender, D. S., Morey, L. C., Clark, L. A., Oldham, J. M., Alarcon, R. D., . . . Siever, L. J. (2011). Personality disorder types proposed for DSM-5. *Journal of Personality Disorders, 25*(2), 136-169.
- Sleep, C. E., Hyatt, C. S., Lamkin, J., Maples-Keller, J. L., & Miller, J. D. (2017). Examining the relations among the DSM-5 alternative model of personality, the Five-Factor Model, and externalizing and internalizing behavior. *Personality Disorders: Theory, Research, and Treatment, 1-6*.
- Somma, A., Markon, K. E., Krueger, R. F., & Fossati, A. (2019). The replicability of the Personality Inventory for DSM-5 domain scale factor structure in U.S. and non-U.S. samples: A quantitative review of the published literature. *Psychological Assessment, 1-17*.
- Soraya, S., Kamalzadeh, L., Nayer, V., Bayat, E., Alavi, K., & Shariat, S. V. (2017). Factor structure of Personality Inventory for DSM-5 (PID-5) in an Iranian sample. *Iranian Journal of Psychiatry and Clinical Psychology, 22*(4), 308-317.

- Spitzer, R.L., Kroenke, K., Williams, J.B., & Lowe, B. (2006). A brief measure for assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092-1097.
- StataCorp. (2017). *Stata Statistical Software: Release 15*. College Station, TX: Stata Corp LLC.
- Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.
- Sutin, A. R., Stephan, Y., Luchetti, M., Artese, A., Oshio, A., & Terracciano, A. (2016). The five-factor model of personality and physical inactivity: A meta-analysis of 16 samples. *Journal of Research in Personality*, 63, 22-28.
- Suzuki, T., Samuel, D. B., Pahlen, S., & Krueger, R. F. (2015). DSM-5 Alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: An item-reponse theory analysis. *Journal of Abnormal Psychology*, 124(2), 243-354.
- Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *The MMPI-Restructured Clinical (RC) scales: Development, validation and interpretation*. Minneapolis, MN: University of Minnesota Press.
- Triandis, H. C., & Suh, E. M. (2002). Cultural influences on personality. *Annual Review of Psychology*, 53, 133-160.
- Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings*. USAF ASD Tech. Rep. No. 61-97, TX: Lack-land Air Force Base.
- Van den Broeck, J., Bastiaansen, L., Rossi, G., Dierckx, E., De Clercq, B., & Hofmans, J. (2014). Hierarchical structure of maladaptive personality traits in older adults: Joint

- factor analysis of the PID-5 and the DAPP-BQ. *Journal of Personality Disorders*, 28(2), 198-211.
- Van Zyl, H., & Meiselman, H. L. (2015). The roles of culture and language in designing emotion lists: Comparing the same language in different English and Spanish speaking countries. *Food Quality and Preference*, 41, 201-213.
- Verheul, R. (2005). Clinical utility of dimensional models for personality pathology. *Journal of Personality Disorders*, 19(3), 283-302.
- Watson, D., Clark, L. A., & Chmielewski, M. (2008). Structures of personality and their relevance to psychopathology: II. Further articulation of a comprehensive unified trait structure. *Journal of Personality*, 76(6), 1545-1586.
- Watters, C. A., & Bagby, M. (2018). A meta-analysis of the five-factor internal structure of the Personality Inventory for DSM-5. *Psychological Assessment*, 30(9), 1255-1260.
- Whitney, J. M. (2002). *Using the MMPI/MMPI-2 with the Hispanic/Latino population*. Ohio: U.S. Department of Education.
- Widiger, T. A. (2011). The DSM-5 dimensional model of personality disorder: Rationale and empirical support. *Journal of Personality Disorders*, 25(2), 222-234.
- Widiger, T. A., & Costa, P. T. (1994). Personality and personality disorders. *Journal of Abnormal Psychology*, 103(1), 78-91.
- Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual of Mental Disorders-fifth edition. *Journal of Abnormal Psychology*, 114(4), 494-504.

- Widiger, T. A., & Simonsen, E. (2005). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders, 9*(2), 110-130.
- Wright, A. G., Thomas, K. M., Hopwood, C. J., Markon, K. E., Pincus, A. L., & Krueger, R. F. (2012). The hierarchical structure of DSM-5 pathological personality traits. *Journal of Abnormal Psychology, 121*(4), 951-957.
- Yalch, M. M., & Hopwood, C. J. (2016). Convergent, discriminant, and criterion validity of the DSM-5 traits. *Personality Disorders: Theory, Research, and Treatment, 7*(4), 394-404.
- Zhang, J., & Bond, M. H. (1998). Personality and filial piety among college students in two Chinese societies: The added value of indigenous constructs. *Journal of Cross-Cultural Psychology, 29*(3), 402-417.
- Zimmermann, J., Alternstein, D., Krieger, T., Holtforth, M. G., Pretsch, J., Alexopoulos, J., . . . Leising, D. (2014). The structure and correlates of self-reported DSM-5 maladaptive personality traits: Findings from two German-speaking samples. *Journal of Personality Disorders, 28*(4), 518-540.
- Zimmermann, J., Mayer, A., Leising, D., Krieger, T., Holtforth, M. G., & Pretsch, J. (2017). Exploring occasion specificity in the assessment of DSM-5 maladaptive personality traits: A latent state-trait analysis. *European Journal of Psychological Assessment, 33*(1), 47-54.