*Archives of Assessment Psychology,* Vol. 13, No. 1, (25-36)
Printed in U.S.A. All rights reserved

© 2023 American Board of Assessment Psychology

# Construct Validity of the CAPQ and PAI in a Sample of College Students

*Tiffany Grezmak, M.A., Jennifer B. Levin, Ph.D. Amir M. Poreh, Ph.D.,*

**Abstract**

*This study assessed the normative equivalence and construct validity of the Cleveland Adaptive Personality Questionnaire (CAP-Q, Poreh and Levin 2019), a relatively brief multi-scale personality inventory developed to assess personality traits and psychopathological states. The CAP-Q was administered concurrently with the Personality Assessment Inventory (PAI; Morey, 1991) to a sample of 109 college students. The standard scores for matching scales, computed using non-aged-corrected norms, were equivalent, with the CAP-Q age-corrected norms producing better fitting data. Additional analyses showed adequate convergent validity with highly correlated matching scales. Additionally, the multi-scale profiles were comparable across both non-elevated and elevated profiles. Overall, this study shows that the CAP-Q and PAI have similar psychometric properties, with the former being more consistent with prevailing diagnostic models. Recommendations for future studies of the CAP-Q are discussed, including the development of factorial-based subscales.*

## Introduction

Psychologists have developed various multi-scale personality questionnaires for the screening of psychopathology. The earliest measures include the Humm and Wadsworth *Personality Inventory* (1935) and McKinley and Hathaway's (1944; 1948) 567-item true-false *Minnesota Multiphasic Personality Inventory* (MMPI). In the late 1980s, Morey (1991) used prevailing diagnostic criteria (American Psychiatric Association, 1987) to develop the *Personality Assessment Inventory* (PAI). With the MMPI-2 serving as the criterion, follow-up studies confirmed the convergent validity of this new scale (Kurtz et al., 1993; McDevitt-Murphy et al., 2007; Walters & Geyer, 2005).

The PAI, much like other lengthy multi-scale personality measures, is often used for the screening of law enforcement applicants (Weiss et al., 2004), personal injury claimants (Blanchard et al., 2003; Cheng et al., 2010; Whiteside et al., 2020), and in forensic settings (Morey & Quigley, 2002). However, due to its prohibitive length, it is rarely utilized in routine clinical practice or research settings, and is not practical for assessing older adults (Camara et al., 2000; Ingram et al., 2020).

The *Cleveland Adaptive Personality Questionnaire* (CAP-Q, Poreh & Levin, 2019) was developed as a relatively brief measure for screening psychopathology. This new multi-scale measure relies on the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; DSM–5; American Psychiatric Association, 2013) personality disorders nosology and includes scales addressing Obsessive-Compulsive Personality and Avoidant personality traits. Unlike existing

multi-scale measures, the CAP-Q also includes a scale for assessing bipolar features rather than existing hypomanic states. Additionally, unlike existing lengthy personality inventories, which utilize a lengthy fixed set of questions, the CAP-Q employs a flexible semi-adaptive approach. This methodology mirrors the flexible approach to neuropsychological assessment. Namely, the inventory's questions can be tailored to the referral question utilizing a two-tiered system. The first tier includes a static core inventory, while the second-tier addresses domains that usually require the administration of standalone measures, such as scales for assessing ADHD, pain, PTSD, and ASD. The clinician relies on the referral question to determine which second-tier scales to administer, if any. Both the core and supplemental scales utilize indirect questions and age-corrected norms. The CAP-Q also includes brief social desirability, defensiveness, and response bias indices, much like those developed for the lengthier measures.

Following Morey (1991), we used Cronbach and Meehl's (1955) and Loevinger's (1957) recommendations to assess the construct validity of the CAP-Q core inventory. We only examined the overlapping scales of the two measures. Initially, we compared the norms produced by the two measures using both general and college student norms. Following Ben-Porath and Butcher (1989) and Tarescavage and Menton (2020), we also assessed the profile consistency generated by the two measures.

## Method

### *Measures*

**The PAI (Morey, 1996).** The PAI is a 344-item Likert-based self-report questionnaire consisting of 22 scales. This study focuses on the 11 primary clinical scales, associated subscales, and four validity scales (see Table 1). According to the test publisher, the normative sample is comprised of 1000 community-dwelling adults (matched based on gender, race, and age), a sample of 1265 patients from 69 clinical sites, and a college sample of 1051 students. The internal consistency coefficient (Cronbach's *alpha,* α) for the full clinical scales ranged from .82 to .94; clinical subscales, .63 to .87; and validity scales, .29 to .81. Test-retest reliability was assessed for two to four weeks, and median alpha and test-retest correlations exceeded .80 for the 22 scales (Boone et al., 1998; Morey, 1991; Wise et al., 2010; Morey, 1991). Boyle et al. (1994) report that in a sample of 70 non-psychiatric patients, the test-retest interval of 28 days resulted in a median coefficient of .70, indicating less than optimal stability.

**The CAP-Q (Poreh & Levin, 2019).** The CAP-Q self-report measure consists of four validity scales and 10 clinical scales. Four of the scales assess for primarily internalizing factors such as depression and anxiety, six of the scales match DSM-5 personality disorder diagnostic entities combined, and three are validity scales (Table 1). The normative data for the CAP-Q is based on a sample of 1835 community-dwelling adults (matched based on gender, race, and age) and a sample of 2568 community-dwelling adults with a mental health history. Age-based norms consist of three age groups; 18 to 35 years ($n$ = 734); 36 to 64 years ($n$ = 925) and 65 to 90 ($n$ = 184). The internal consistency coefficients for the clinical scales ranged from .70 to .87, and .77 to .84 for the validity scales. The test-retest correlations ranged from .74 to .90 across two to five months ($n$ = 831), .67 to .89 between 6 to 12 months ($n$ = 420), and .58 to .82 for over a year ($n$ = 250). Since CAP-Q publication, the authors have collected additional normative data, and are developing new non-linear age and education, regression-based norms. The CAP-Q is available in German, Czech, Hebrew, French, French-Canadian, Spanish, Norwegian, and Arabic translations.

**Table 1**

***The CAPQ and PAI Clinical and Validity Scales***

| CAPQ | | | PAI | | |
|---|---|---|---|---|---|
| Clinical Scales | | # Items | Clinical Scales | | # Items |
| Somatization | SOM | 10 | Somatic concerns | SOM | 24 |
| Depressive Mood | DEP | 11 | Depression | DEP | 24 |
| Anxiety | ANX | 11 | Anxiety | ANX | 24 |
| Bipolar | BIP | 10 | Mania | MAN | 24 |
| Paranoid Personality Traits | PAR | 10 | Paranoia | PAR | 24 |
| Schizotypal Personality | SCIZ | 10 | Schizophrenia | SCZ | 24 |
| Obsessive Compulsive Personality Traits | OCP | 7 | Anxiety Related Disorders | ARD | 24 |
| Sociopathic Personality Traits | SOC | 10 | Antisocial features | ANT | 24 |
| Avoidant Personality Traits | AVD | 10 | | | |
| Borderline Personality Traits | BPT | 11 | Borderline features | BPT | 24 |
| Substance Abuse | SUB | 11 | | | |
| Alcohol Abuse | AA | 6 | Alcohol Problems | ALC | 12 |
| Drug Use | DU | 5 | Drug Problems | DRG | 12 |
| Validity Scales | | | Validity Scales | | |
| Naiveté | NVA | 6 | Positive Impression | PIM | 9 |
| Inconsistency | CON | 17 * | Inconsistency | ICN | 10 |
| Infrequency Scale | INF | 16 * | Infrequency | INF | 8 |
| | | | Negative Impression | NIM | 9 |

* Adapted from Poreh and Levin 2019. Items on these scales appear throughout the questionnaire.

## Procedure

### *Participants*

The data were collected as part of the first author's master's thesis. Participants were contacted via email by the researcher with information about the study. Those who agreed to participate completed the IRB-approved informed consent form. They were then individually scheduled to complete the study through a one-on-one Zoom meeting (due to COVID-19, meeting the participants in person was not feasible). After completing a brief demographic questionnaire, the volunteers were administered the two questionnaires. They then underwent a debriefing, and then mailed back the forms. It took about one to one-and-a-half hours for participants to complete

both questionnaires.

### *Analyses*

The data from PAI and CAP-Q were manually scored and converted to standard scores using published norms. Following Ben-Porath & Butcher (1989) and Wygant et al. (2009), we conducted paired t-test analyses to investigate the comparability of general population norms across both measures. The effect size was assessed using Cohen's d (Cohen et al., 2013; Hall et al., 2021), guided by the interpretive recommendations of Wygant et al. (2009). We then went on to examine the internal consistency of the primary scales of the two questionnaires. After the aforementioned analyses, we evaluated the correlation between the scales and compared the profiles based on the criteria proposed by Morey (1999). Following Ben-Porath and Butcher (1989) and Tarescavage and Menton (2020), we also examined the profile consistency using three criteria: (a) the percentage of participants classified as having a profile within normal limits, defined by T-scores less than 65; (b) the level of similarity for the top two-point configuration on both measures; and (c) the application of the chi-square statistic to assess differences and similarities in the three-point configuration profiles.

## Results

### *Adequacy of Normative Data*

Table 2 shows the mean standardized T-scores for the PAI and CAP-Q matched scales. Independent *t*-tests showed statistically significant differences between the matched PAI Mania (MAN) and CAP-Q Bipolar (BIP) scales and the PAI Antisocial (ANT) and CAP-Q Sociopathy (SOC) scales. The correlation between the PAI and CAP-Q matching clinical scales ranged from .62 to .87 ($p < .001$). The following PAI subscales evidenced weaker correlation coefficients: the CAP-Q Schizotypal Personality Traits (SCIZ) scale moderately correlated with the PAI Schizophrenia (SCZ) Social Detachment subscale (.35, $p < .001$); the CAP-Q BIP scale marginally correlated with the PAI MAN Grandiosity subscale (.21, $p = .03$); and the CAP-Q Borderline Personality Traits (BPT) scale correlated with the PAI Borderline (BOR) Self-Harm subscale (.42, $p < .001$).

Additional exploratory analyses showed high intercorrelation between related constructs. For example, scales that assess the DSM-5 cluster B personality traits, such as the CAP-Q SOC scale and PAI BOR scales (.78, $p < .001$), and more specifically, the BOR Irritability subscale (.73, $p < .001$), were highly correlated. The CAP-Q BIP and PAI BOR scales were also positively correlated (.68, $p < .001$), reflecting the shared symptoms assessed by these two scales. Finally, as seen in Table 2, the matching validity scales, the PAI Negative Impression (NIM) and Positive Impression (PIM) and the CAP-Q Infrequency (INF) and Naiveté (NAV) validity scales were highly correlated, whereas the PAI inconsistency (INC) and CAP-Q Consistency (CON) scale were not.

**Table 2**

***Norms and Statistical Data for Matched PAI and CAPQ Clinical and Validity Scores***

*Note.* Bolded values indicate medium effect size (.5 < Cohen's *d* < 0.8). The PAI ARD and CAPQ

| | PAI Norms | | | | | CAPQ Norms | | | | Comparisons | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | General | | College Student* | | | General | | Age Corrected* | | | | | |
| Scale | M | SD | M | SD | Scale | M | SD | M | SD | *d* | *t* | *p* | *r* |
| SOM | 55.1 | 11.1 | 60.3 | 14.8 | SOM | 56.3 | 10.5 | 54.4 | 9.6 | 0.11 | 1.2 | .23 | .80 |
| DEP | 59.6 | 15.2 | 61.6 | 16.5 | DEP | 58.3 | 12.2 | 53.6 | 10.7 | 0.19 | 2.0 | .05 | .86 |
| ANX | 62.6 | 15.1 | 61.9 | 15.3 | ANX | 61.2 | 11.4 | 56.4 | 9.9 | 0.14 | 1.5 | .23 | .87 |
| MAN | 51.9 | 10.9 | 48.0 | 10.9 | BIP | 58.1 | 11.1 | 54.5 | 9.6 | **0.67** | 7.0 | < .001 | .62 |
| ANT | 52.1 | 10.3 | 46.7 | 9.2 | SOC | 55.4 | 12.0 | 51.2 | 9.6 | 0.38 | 4.1 | < .001 | .80 |
| PAR | 55.4 | 12.2 | 56.0 | 12.5 | PAR | 58.3 | 12.2 | 54.3 | 9.7 | 0.25 | 2.6 | .90 | .78 |
| SCZ | 55.3 | 14.4 | 56.8 | 15.3 | SCIZ | 58.7 | 14.0 | 58.7 | 11.4 | 0.29 | 3.1 | .38 | .68 |
| BOR | 58.3 | 13.6 | 54.5 | 13.3 | BPT | 57.5 | 12.3 | 52.2 | 9.8 | 0.06 | 1.0 | .30 | .82 |
| ALC | 48.8 | 10.5 | 46.6 | 9.1 | ALC | 51.2 | 11.1 | 50.8 | 11.2 | 0.22 | 4.1 | < .01 | .73 |
| DRG | 50.6 | 12.0 | 53.9 | 15.5 | DRG | 52.2 | 13.7 | 53.2 | 17.7 | 0.12 | 4.6 | < .01 | .66 |
| INC | 50.9 | 8.6 | 52.7 | 10.7 | CON | 45.8 | 9.1 | 45.2 | 8.8 | 0.20 | 3.9 | < .01 | .37 |
| NIM | 55.1 | 13.4 | 57.2 | 16.5 | INF | 59.9 | 13.7 | 53.8 | 10.6 | 0.35 | 1.5 | .13 | .79 |
| PIM | 44.8 | 13.4 | 49.7 | 12.5 | NAV | 35.1 | 10.2 | 35.0 | 10.1 | **0.54** | 0.9 | < .01 | .34 |

COPD and AVD scales were not included in the analyses.

Table 2 also depicts the PAI and the CAP-Q T-scores and paired *t*-tests for the non-age - corrected norms. Most of the scales showed no (or marginal) differences, aside from the PAI MAN and CAP-Q BIP ($t = 4.2$, $df = 218$, $p < .001$), the PAI ANT and CAP-Q SOC ($t = 2.2$, $df = 218$, $p < .01$), and the PAI Drug (DRG) and Alcohol (ALC) and the CAP-Q Drug Use (DU) and Alcohol Abuse (AA) scales ($t = 2.2$, $df = 218$, $p < .01$) and ($\underline{t} = 2.2$, $df = 218$, $p < .01$), respectively. Cohen's *d* was medium for the MAN and BIP scales, and small for the ANT and SOC, and ALC and DRG, scales. Glass's Delta (Δ), a measure that adjusts for the standard deviation and allows assessment of the practical significance, shows that aside from the MAN, none of the effects were meaningful, as depicted by the paired *t*-tests.

Table 2 also shows that the college student norms for the PAI did not significantly impact the sample elevations. In contrast, the use of the CAP-Q age-based norms significantly reduced the elevations of the Depression (DEP; $t = 3.0$, $df = 218$, $p < .001$), Anxiety (ANX; $t = 2.2$, $df = 218$, $p < .001$), BIP ($t = 2.5$, $df = 218$, $p < .01$), SOC ($\underline{t} = 2.9$, $df = 218$, $p < .001$), Paranoia (PAR; $\underline{t} = 2.7$, $df = 218$, $p < .001$), BIP ($\underline{t} = 3.0$, $df = 218$, $p = .003$), Avoidant (AVD) ($\underline{t} = 3.7$, $df = 218$, $p < .001$), and to a lesser extent Obsessive Compulsive Personality Traits (OCP) ($\underline{t} = 2.3$, $df = 218$, $p = .02$). The SCIZ, ALC, DRG, CON, and NAV scales T-scores remained relatively stable when using age-corrected norms. Assuming that this was a random sample, one sees that the mean T-Scores are closer to the anticipated mean. Table 3 shows the internal consistency of matching clinical scales. One sees that the PAI exhibits somewhat higher internal consistency, as would be expected, given that the CAP-Q has fewer items.

**Table 3**

***Internal Consistency of Matching Clinical Scales***

| PAI | # Items | α | α* | CAPQ | # Items | α | α ** |
|---|---|---|---|---|---|---|---|
| SOM | 24 | .91 | .85 | SOM | 11 | .82 | .84 |
| DEP | 24 | .93 | .91 | DEP | 11 | .88 | .88 |
| ANX | 24 | .94 | .93 | ANX | 11 | .87 | .87 |
| MAN | 24 | .83 | .78 | BIP | 14 | .75 | .80 |
| ANT | 24 | .83 | .84 | SOC | 14 | .77 | .69 |
| PAR | 24 | .88 | .80 | PAR | 10 | .84 | .86 |
| SCZ | 24 | .89 | .88 | SCIZ | 11 | .77 | .78 |
| BOR | 24 | .92 | .88 | BPT | 11 | .85 | .86 |
| ALC | 12 | .84 | .92 | ALC | 3 | .72 | .84 |
| DRG | 12 | .80 | .89 | DRG | 3 | .79 | .73 |

* Boone 1998 (psychiatric patients).
** Standardization group.

Profile consistency was assessed using the previously outlined criteria. Table 4 shows the sensitivity and specificity of the same high point on both measures. One sees that all of the scales, aside from the MAN and BIP, had comparable diagnostic agreement. The first criterion relates to the proportion of participants on each scale that produced normal profiles. The results showed that

35.5 and 30.0 % of the participants, respectively, produced normal profiles on both measures ($X_1^2$ = 50, $p$ < .001; $\eta^2$ = .68).

**Table 4**

***Diagnostic Agreement***

| PAI Scale | CAPQ Scale | Sensitivity % (T < 65) | Specificity % (T > 65) | $\eta^2$ | $p$ |
|---|---|---|---|---|---|
| SOM | SOM | 94.8 | 61.5 | .56 | < .001 |
| ANX | ANX | 81.0 | 73.1 | .49 | < .001 |
| DEP | DEP | 82.6 | 90.8 | .69 | < .001 |
| MAN | BIP | 40.0 | 96.7 | .10 | .169 |
| ANT | SOC | 98.8 | 56.0 | .37 | < .001 |
| PAR | PAR | 58.3 | 82.7 | .31 | .004 |
| SCZ | SCIZ | 58.8 | 83.9 | .37 | < .001 |
| BOR | BPT | 85.9 | 52.0 | .38 | < .001 |
| ALC | ALC | 99.0 | 80.0 | .34 | < .001 |
| DRG | DRG | 95.9 | 63.3 | .56 | < .001 |

Table 5 provides a parametric analysis of the same data. A comparison of the two-point and three-point configurations was also statistically significant ($X_{1225}^2$ = 1888, $p$ < .001, and $X_{2215}^2$ = 3125, $p$ < .001, respectively).

**Discussion**

This is the first study to evaluate the norm equivalency and construct validity of the CAP-Q using a commonly used multi-scale measure. Initial analyses showed that the general-population normative data of the CAP-Q and PAI are statistically equivalent on matched clinical scales. It also confirms previous PAI studies showing that young adults produce significant elevations on specific scales (De Moor, et al., 2009; Trofimova, 2015). When we corrected the skewed scores with available age-based norms, the CAP-Q scales adjusted to lower, whereas the norms published by the PAI test publishers remained high. This finding confirms the importance of adjusting for age to avoid pathologizing young adults (Osberg & Poland, 2002). Nonetheless, the implications of these findings for clinical practice need to be further investigated. Namely, although several clinical studies confirm a significant correlation between age and PAI scale elevations, Butcher et al. (1991) argued that age-based norms have limited clinical significance. Kennedy et al. (2015) argued that by correcting for elevations on scales, the clinician might disregard emotional difficulties that, although common in specific age groups, impact day-to-day functioning. Therefore, much like in clinical neuropsychological practice, we recommended that personality measures include both age-corrected and uncorrected norms when presenting the resulting profile. Such a solution would limit the pathologizing of young adults and the underdiagnosis of older adults, who tend to be more defensive.

**Table 5**

***Profile Comparison Between CAPQ and PAI Clinical and Validity Scales (n = 110)***

| PAI Scale | % Elevated Cases (T > 65) | CAPQ Scale | % Elevated Cases (T < 65) | *p* |
|---|---|---|---|---|
| SOM | 12 | SOM | 12 | < .001 |
| DEP | 24 | DEP | 21 | < .001 |
| ANX | 32 | ANX | 24 | < .001 |
| MAN | 5 | BIP | 17 | .170 |
| ANT | 6 | SOC | 12 | < .001 |
| PAR | 1 | PAR | 22 | .001 |
| SCZ | 15 | SCIZ | 23 | < .001 |
| BOR | 23 | BPT | 23 | < .001 |
| ARD | 25 | AVD | 14 | .070 |
| NAV | [a] | OCD | 16 | .001 |
| ALC | 4 | ALC | 13 | < .010 |
| DRG | 10 | DRG | 11 | < .001 |
| NIM | 36 | INF | 19 | < .001 |
| PIM | 6 | NAV | 1 | < .001 |
| INC | 7 | CON | 2 | .710 |

[a] No statistic could be computed because NAV did not contain any significant elevations.

As previously mentioned, both the PAI clinical scales and subscales and the matching CAP-Q clinical scales were highly correlated. These correlations were reported during the validation of other measures, such as the PICTS *Thinking Style Scales,* PAI (Walters & Geyer, 2005), and *Personality Inventory for ICD 11* (PiCD) (Tarescavage & Menton, 2020).

It is noteworthy that the PAI-MAN and CAP-Q BIP scales were only moderately correlated. These findings reflect a fundamental difference between the two measures. Namely, while the PAI MAN scale was designed to detect current elevated mood or hypomanic states, it is less helpful in identifying the presence of an underlying bipolar disorder (Mullen-Magbalon, 2008). The CAP-Q BIP scale aimed to improve the classification of individuals with bipolar disorder by relying on life-long (trait) questions.

As was previously mentioned, the CAP-Q and PAI Inconsistency scales did not correlate. However, this finding is not unexpected, as these scales measure random error (Morey, 2007). Moreover, prior research indicates that the PAI Inconsistency scale has more difficulty identifying participants who responded inconsistently than other measures (Nikolova, 2012). With that said, additional research is underway to assess the CAP-Q CON scale ability to detect simulated psychiatric conditions. Another related finding was the inconsistency between the PAI PIM scale and the CAP-Q NAV elevations. This finding might result from the fact that the NAV has fewer items than the PAI PIM scale. The CAP-Q authors are considering lengthening this scale in future editions of the test by adding previously removed items.

In sum, the present study provides robust support for the concurrent validity of the CAP-Q. With that said, it is recommended that future studies continue to examine the construct validity of the CAP-Q. As suggested by Peter (1981), a single study provides evidence but does not provide conclusive proof of construct validity. Instead, it serves as one more step in an ongoing reevaluation, refinement, and continuous development process of scales.

While the present study included a relatively small sample of clinical participants, it is predicted that our results will be replicated with a larger sample of clinical participants. Additional studies should also be conducted to correlate the CAP-Q with other personality and psychopathology measures, including the MMPI-2, to assess convergent and discriminant validity further.

The present study had several limitations. First, this study took place over 2019-2020, during the Covid-19 pandemic. As such, the personality assessment administration was done online using Zoom, an internet-based conferencing service. This method of administering assessments is not ideal, and some participants may not have treated the situation the same as they would have in person.

Although controversial, online recruitment is a reliable and valid way to conduct psychological research (Paolacci et al., 2010). Finally, a portion of the data was not returned due to package delivery failures. Another limitation relates to the fact that the assessments were not randomized, with the CAP-Q being completed first followed by the PAI. Due to the length of the questionnaires, some participants may have become fatigued by the time they were finishing up the PAI, which may have impacted their results. Finally, no information was obtained from participants regarding existing mental health disorders. Future studies should evaluate the consistency between the CAP-Q, PAI, elevations, and mental health diagnoses.

## Conclusions

This study provides robust evidence for the construct validity of the *Cleveland Adaptive Personality Questionnaire.* It shows that this new brief measure has sound normative properties and could be used to replace lengthier existing personality inventories to screen common mental health conditions. With the collection of extensive data, it is also believed that artificial intelligence and deep learning methods could be used to predict treatment adherence and length of hospital stay, and improve the diagnostic classification. We hope that other clinicians and researchers will utilize the CAP-Q in research and as a screener for both normal and clinical populations, and independently develop interpretative software, much like was done during the early years of the MMPI.

## Compliance with Ethical Standards

*Conflict of Interest:* The first author is the author of the CAP-Q. There are no other potential conflicts of interest with respect to the research, authorship, and publication of this article.

*Informed Consent:* Informed consent was obtained from all participants included in the study.

## About the Authors

**Tiffany Grezmak, MA** is a graduate of the Clinical Psychology Program at Cleveland State University.

**Amir M. Poreh, PhD** is a Full Professor at Cleveland State University, Department of Psychology, and an Associate Clinical Professor at Case Western Reserve University School of Medicine, Department of Psychiatry. He is the author of the Cleveland Adaptive Personality Questionnaire.

**Jennifer B. Levin, Ph.D.,** is a Full Professor at Case Western Reserve University School of Medicine, Department of Psychiatry, and a clinical psychologist at University Hospitals Cleveland Medical Center. She is a co-author of the Cleveland Adaptive Personality Questionnaire.

Correspondence concerning this article should be addressed to Dr. Amir Poreh at the following email: a.poreh(at)csuohio.edu

## References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.).

Ben-Porath, Y. S., & Butcher, J. N. (1989). The comparability of MMPI and MMPI-2 scales and profiles. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*(4), 345-347. https://doi.org/10.1037/1040-3590.1.4.345

Blanchard, D. D., McGrath, R. E., Pogge, D. L., & Khadivi, A. (2003). A comparison of the PAI and MMPI-2 as predictors of faking bad in college students. *Journal of Personality Assessment, 80*(2), 197-205. https://doi.org/10.1207/S15327752JPA8002_08

Boone, K. B., Pontón, M. O., Gorsuch, R. L., González, J. J., & Miller, B. L. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology, 13*(7), 585-595. https://doi.org/10.1016/S0887-6177(97)00074-7

Boyle, G. J., & Lennon, T. J. (1994). Examination of the reliability and validity of the Personality Assessment Inventory. *Journal of Psychopathology and Behavioral Assessment, 16*(3), 173-187. https://doi.org/10.1007/BF02229206

Butcher, J. N., Aldwin, C., M., Levenson, M. R., Ben-Porath, Y. S., Spiro, A., & Bossé, R. (1991). Personality and aging: A study of the MMPI-2 among older men. *Psychology and Aging*, 361-370. https://doi.org/10.1037/0882-7974.6.3.361

Cheng, M. K., Frank, J. B., & Hopwood, C. J. (2010). Assessment of motor vehicle accident claimants with the PAI. In *Clinical Applications of the Personality Assessment Inventory.* (pp. 177-193). Routledge/Taylor & Francis Group.

Cohen, R. J., Swerdlik, M. E., & Sturman, E. (2013). *Psychological testing and assessment: An introduction to tests and measurement* (8th ed.). McGraw-Hill.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. https://doi.org/10.1037/h0040957

De Moor, M. H., Distel, M. A., Trull, T. J., & Boomsma, D. I. (2009). Assessment of borderline personality features in population samples: Is the Personality Assessment Inventory-Borderline Features scale measurement invariant across sex and age? *Psychological Assessment, 21*(1), 125-130. https://doi.org/10.1037/a0014502

Hall, J. T., Menton, W. H., & Ben-Porath, Y. S. (2021). Examining the psychometric equivalency of MMPI-3 scale scores derived from the MMPI-3 and the MMPI-2-RF-EX. *Assessment, 29(4),* 842–853. https://doi.org/10.1177/1073191121991921

Humm, D. G., & Wadsworth, G. W. (1935). The Humm-Wadsworth Temperament Scale. *American Journal of Psychiatry, 92*(1), 163-200. https://doi.org/10.1176/ajp.92.1.163

Kennedy, J. E., Cooper, D. B., Reid, M. W., Tate, D. F., & Lange, R. T. (2015). Profile analyses of the Personality Assessment Inventory following military-related traumatic brain injury. *Archives of Clinical Neuropsychology, 30*(3), 236-247. https://doi.org/10.1093/arclin/acv014

Kurtz, J. E., Morey, L. C., & Tomarken, A. J. (1993). The concurrent validity of three self-report measures of borderline personality. *Journal of Psychopathology and Behavioral Assessment, 15*(3), 255-266. https://doi.org/10.1007/BF01371382

Liu, H. Y., & Weng, L. J. (2009). An effect size index for comparing two independent alpha coefficients. *British Journal of Mathematics and Statistical Psychology, 62*(Pt 2), 385-400. https://doi.org/10.1348/000711008X315518

Lenzenweger, M.F., Johnson, M.D., Willett, J.B., 2004. Individual growth curve analysis illuminates stability and change in personality disorder features. *Archives of General Psychiatry, 61*, 1015. https://doi.org/10.1001/archpsyc.61.10.1015

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*(3), 635-694. https://doi.org/10.2466/pr0.1957.3.3.635

McDevitt-Murphy, M. E., Weathers, F. W., Flood, A. M., Eakin, D. E., & Benson, T. A. (2007). The utility of the PAI and the MMPI-2 for discriminating PTSD, depression, and social phobia in trauma-exposed college students. *Assessment, 14*(2), 181-195. https://doi.org/10.1177/1073191106295914

McKinley, J. C., & Hathaway, S. R. (1944). The Minnesota Multiphasic Personality Inventory. V. Hysteria, hypomania and psychopathic deviate. *Journal of Applied Psychology, 28*(2), 153-174. https://doi.org/10.1037/H0059245

McKinley, J. C., Hathaway, S. R., & Meehl, P. E. (1948). The Minnesota Multiphasic Personality Inventory: VI. The K Scale. *Journal of Consulting Psychology, 12*(1), 20-31. https://doi.org/10.1037/h0061377

Morey, L. C. (1991, 2007). *The Personality Assessment Inventory: professional manual*. PAR.

Morey, L. C., & Quigley, B. D. (2002). The use of the Personality Assessment Inventory (PAI) in assessing offenders. *International Journal of Offender Therapy and Comparative Criminology, 46*(3), 333-349. https://doi.org/10.1177/0306624x02463007

Mullen-Magbalon, S. D. (2008). *Using the Personality Assessment Inventory to discriminate among borderline personality disorder, bipolar disorder, and post-traumatic stress disorder.* (Ph.D. Dissertation). University of Tennessee, Knoxville.

Nikolova, N. L., Hendry, M., Douglas, K., Edens, J., & Lilienfeld, S. (2012). The inconsistency of inconsistency scales: A comparison of two widely used measures. *Behavioral sciences & the law, 30 1*, 16-27. https://doi.org/10.1002/bsl.1996

Osberg, T. M., & Poland, D. L. (2002). Comparative accuracy of the MMPI-2 and the MMPI-A in the diagnosis of psychopathology in 18-year-olds. *Psychol Assess, 14*(2), 164-169. https://doi.org/10.1037//1040-3590.14.2.164

Paolacci, G., Chandler, J., & Ipeirotis. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*(5), 411 - 419. https://doi.org/10.1017/S1930297500002205

Peter, J. P. (1981). Construct validity: A review of basic issues and marketing practices. *Journal of Marketing Research, 18*(2), 133-145. https://doi.org/10.2307/3150948

Tarescavage, A. M., & Menton, W. H. (2020). Construct validity of the Personality Inventory for ICD-11 (PiCD): Evidence from the MMPI-2-RF and CAT-PD-SF. *Psychol Assess, 32*(9), 889-895. https://doi.org/10.1037/pas0000914

Trofimova, I. (2015). Age and temperament differences in mental illness: Results from three studies. *European Psychiatry, 30*, 672. https://doi.org/10.1016/S0924-9338(15)30533-2

Walters, G. D., & Geyer, M. D. (2005). Construct validity of the psychological inventory of criminal thinking styles in relationship to the PAI, disciplinary adjustment, and program completion. *J Pers Assess, 84*(3), 252-260. https://doi.org/10.1207/s15327752jpa8403_04

Weiss, W. U., Rostow, C., Davis, R., & DeCoster-Martin, E. (2004). The Personality Assessment Inventory as a selection device for law enforcement personnel. *Journal of Police and Criminal Psychology, 19*(2), 23-29. https://doi.org/10.1007/BF02813870

Whiteside, D. M., Hunt, I., Choate, A., Caraher, K., & Basso, M. R. (2020). Stratified performance on the Test of Memory Malingering (TOMM) is associated with differential responding on the Personality Assessment Inventory (PAI). *Journal of Clinical and Experimental Neuropsychology, 42*(2), 131-141. https://doi.org/10.1080/13803395.2019.1695749

Wise, E. A., Streiner, D. L., & Walfish, S. (2010). A review and comparison of the reliabilities of the MMPI-2, MCMI-III- and PAI presented in their respective test manuals. *Measurement and Evaluation in Counseling and Development, 42*(4), 246-254. https://doi.org/10.1177/0748175609354594

Wygant, D. B., Ben-Porath, Y. S., Arbisi, P. A., Berry, D. T. R., Freeman, D. B., & Heilbronner, R. L. (2009). Examination of the MMPI-2 Restructured Form (MMPI-2-RF) validity scales in civil forensic settings: Findings from simulation and known group samples. *Archives of Clinical Neuropsychology, 24*(7), 671-680. https://doi.org/10.1093/arclin/acp073