Strengths and Limitations of the Personality Inventory for Youth (PIY) in Juvenile Delinquency Assessments

Robert A. Semel

Abstract
This article highlights and discusses the usefulness of the Personality Inventory for Youth (PIY) in juvenile delinquency assessments. Psychiatric disorders have high prevalence rates among youths in the juvenile justice system. The PIY was developed to evaluate a broad range of behavioral and psychological characteristics, which may make it useful in juvenile delinquency assessment contexts. Practical and psychometric strengths and limitations of the PIY in the juvenile delinquency assessment context are presented, with reference to relevant research literature. The effectiveness of this instrument in detecting response bias, particularly under-reporting, and for identifying problems associated with delinquency is discussed. The issue of item overlap and spurious influences on scale correlations, especially between the PIY Defensiveness and Delinquency scales, is also addressed. A comparison of findings with the PIY and the Minnesota Multiphasic Personality Inventory–Adolescent (MMPI-A) in juvenile justice samples helps to identify broader considerations about how youths in juvenile justice settings respond to self-report psychological inventories. Finally, the author offers some practical considerations for evaluators when using the PIY in juvenile delinquency assessments, and suggestions for future research.

Keywords
juvenile delinquency, multi-method assessment, PIY, response style, item overlap, MMPI-A

Use of Standardized Personality Tests in Juvenile Delinquency Assessment

While forensic mental health assessments in juvenile delinquency cases may be ordered by the court at various phases in the judicial process, and sometimes to address specific legal issues (e.g., competence to stand trial, capacity to waive Miranda rights, waiver of jurisdiction from juvenile to criminal court), evaluations in juvenile courts are most often ordered to aid in dispositive decisions through the assessment of mental health needs and the identification of the most appropriate interventions and settings to address existing mental health needs (Archer, 2005; Grisso & Vincent, 2005).

A second focus of assessments to aid the court in dispositive decisions may be an assessment of the risk of aggression that a youth may pose, which has implications for the degree of security that may be needed to minimize the risk of harm to others (Grisso & Vincent, 2005).

High prevalence rates of psychiatric disorders among youths in the juvenile justice system have led respectively to a need for multi-method, multi-informant assessment approaches to adequately address the needs of these youths. The use of such an assessment approach can begin with an interview of the youth and collateral data from sources such as parents, schools, court-affiliated programs, detention facility records, therapists, and mental health records. A significant number of these youths have a history of misusing substances, and therefore a measure to assess substance misuse is also an integral part of the assessment process.

Following a standard multi-method, multi-informant approach to the clinical assessment of the youth, evaluators typically collect collateral data from various sources such as parents, schools, court-affiliated programs, detention facility records, therapists, mental health records, and conduct a clinical interview(s) of the respondent youth. The objective of the evaluation is to provide information that helps inform treatment decision-making and identify the most effective and efficient interventions and settings for the youth based on their mental health needs.
comorbid psychiatric disorders in delinquent adolescents have been reported (Abram, Teplin, McClelland, & Dulcan, 2003; Adams et al., 2013; Teplin, Abram, McClelland, Dulcan, & Mericle, 2002; Wasserman, McReynolds, Schwalbe, Keating, & Jones, 2010), with more than 60% of youths in the juvenile justice system found to meet diagnostic criteria for one or more psychiatric disorders, excluding conduct disorder (Teplin et al., 2002). The National Comorbidity Survey–Adolescent Supplement (Coker, Smith, Westphal, Zonana, & McKee, 2014) reported that youth with three or more psychiatric diagnoses accounted for 54.1% of those who reported being arrested for violent crime. Adolescents with conduct disorder, alcohol use disorders, and drug use disorders had the greatest odds of being arrested for violent crime. A prospective longitudinal study by Abram et al. (2013) found that more than 1 in 10 juvenile detainees experienced posttraumatic stress disorder (PTSD) in the year prior to interview. Some standardized clinical assessment instruments that were developed to assess personality and psychopathology in youths in the general population or in clinical settings have been studied to examine the usefulness of their application in assessing mental health disorders in samples of youths involved in the juvenile justice system. Researchers have taken note of the substantial overlap between the juvenile justice and community mental health youth populations (Grisso, 2005). Grisso (2005) commented that “the ‘clinical’ samples of youths on which some instruments are based included many youths—perhaps a majority of them—who had juvenile justice contact at one time or another” (p. 83).

The Personality Inventory for Youth (PIY; Lachar & Gruber, 1995) was included by Hoge and Andrews (2010) among several standardized personality tests that are useful in forensic assessments. Although designed as a general diagnostic aid in the assessment of youth adjustment rather than as a measure for a specific forensic or juvenile justice purpose, the PIY has been studied and applied in the evaluation of youths from a variety of juvenile justice settings and is used routinely in certain juvenile justice facilities as an evaluative tool (Lachar, Hammer, & Hammer, 2006). An advantage of standardized self-report personality test instruments such as the PIY is that they incorporate measures of the test taker’s response style, which is of utmost importance in forensic assessment contexts. Denial of problematic behaviors, such as aggressive, antisocial, disruptive behaviors, is not unexpected in the evaluation of youths in a juvenile justice context in which the evaluation might have bearing on legal decisions with potential for restriction of freedom (Devieux et al., 2002; Lachar et al., 2006; Penney & Skilling, 2012; Smith, 2007). The validity of test profiles may also be affected by factors such as reading skills and language comprehension, inadequate attention or concentration, poor motivation or cooperation, and arbitrary or random responding. The PIY contains four validity scales to assess for accuracy of test responses. A related advantage of using the PIY in juvenile justice assessment contexts is that the test items are written at a third- to fourth-grade reading level. Administration time typically is between 30 and 60 min.

The PIY was standardized using a national normative sample of 2,327 students in regular education (i.e., mainstream, non-special education) classes, and a second sample of 1,178 clinically referred students recruited from 64 different sites that included in- and outpatient clinics, school special service clinics, hospitals, and a wide variety of private practice settings. The test contains 9 non-overlapping clinical scales, 24 non-overlapping clinical subscales, and 4 validity scales. Gender-specific linear T-scores are based on the scores of the regular education student sample. Although statistical differences were found for at least one scale on each of the demographic variables assessed, that is, gender, age, ethnicity, socioeconomic status (SES), geographic region, and guardianship status, the test authors concluded that the effects of demographic variables on standardized test scores were unsubstantial and separate norm conversions were deemed necessary only for gender. With specific regard to ethnicity effects, statistically significant differences were found on only one of the nine clinical scales, that is, Asian Americans tended to report more difficulties with peer adjustment on the Social Skill Deficits (SSK) scale. On the 9 clinical scales, scores reaching at least 60T are interpreted to be of clinical significance, whereas for the 24 subscales scores elevated at 65T or above are interpreted to be of clinical significance. For the regular education student sample, internal consistency coefficients on the clinical scales ranged from .71 to .90, with a median of .82. Test–retest correlations ranged from .81 to .91, with a median of .85. Reliability estimates for the 24 clinical subscales ranged from .40 to .79, with a median of .70. Test–retest correlations for the clinical subscales ranged from .66 to .90, with a median of .80. Internal consistency coefficients were slightly higher for the clinically referred student sample. Considerable evidence supports the content, construct, concurrent, and criterion-related validity of the PIY profile. As reported in the test manual, correlations between the PIY scales and subscales and the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943) basic clinical scales and the Wiggins Content Scales support convergent and discriminant validity and substantiate the concurrent validity of the PIY. In a review of the PIY by Destefano (1998) for the Buros Institute’s Mental Measurements Yearbook, the test is described as a psychometrically sound instrument, albeit Destefano indicates that the clinical subscales are technically limited based on internal consistency estimates.

Application of the PIY, and its companion instruments, the Personality Inventory for Children, Second Edition (PIC-2; Lachar & Gruber, 2001), an inventory completed by parents, and the Student Behavior Survey (SBS; Lachar, Wingenfeld, Kline, & Gruber, 2000), a teacher report, has been discussed by Lachar and Boyd (2005), and by
Lachar et al. (2006), whose writings include the application of these instruments in forensic assessments. The book chapter by Lachar et al. (2006) is especially informative about the usefulness of this family of measures in the psychological evaluation of children and adolescents across multiple settings. However, those previous writings do not focus on certain issues relevant in the forensic context, such as psychometric considerations pertaining to item overlap between the validity scales and the clinical scales of the PIY. In addition, previous writings have not discussed how studies using the PIY with juvenile justice samples have demonstrated certain similar findings when compared with studies of juvenile delinquent samples with the MMPI–Adolescent (MMPI-A; Butcher et al., 1992). Such comparisons between these two instruments may be helpful in identifying broader considerations about how youths in juvenile justice settings respond to self-report psychological test inventories.

**Purpose of the Current Report**

The current report aims to identify and highlight for the reader some of the advantages as well as some limitations of the PIY when utilized in a juvenile justice context. The following sections will first review several studies of the PIY in juvenile justice settings. Another section will focus on issues pertaining to item overlap on the PIY which have not been previously discussed in the literature. Subsequently, a comparison of research findings with the PIY and the MMPI-A will be discussed, as it offers both a perspective from which to consider strengths and limitations of the PIY in delinquency assessments and a broader perspective on the use of self-report personality measures in the juvenile justice population. Finally, the writer will offer some practical considerations for evaluators when using the PIY in the juvenile delinquency assessment context and suggestions for further research.

**Application of the PIY in Forensic Contexts**

In a study by Negy, Lachar, Gruber, and Garza (2001), which compared English and Spanish versions of the PIY among bilingual Mexican American high school students and incarcerated bilingual Mexican American males, the incarcerated males “obtained significantly higher scores than regular education males on all 9 clinical and 19 out of 24 clinical subscales” (p. 250). Mean $T$-scores for the incarcerated youths and for the regular education youths on the Delinquency (DLQ) scale were 66.0 and 47.3, respectively, with an effect size of 1.26. The significant difference between the two groups on the DLQ scale and the elevated score pattern of the incarcerated youths were consistent with the type of behavior typically seen in youths involved in the juvenile justice system, such as conflicts with general expectations and specific requests of adults and violation of established rules, as may be seen in running away from home, school truancy and/or suspension, illegal acts, association with peers who engage in similar behavior, limited self-control and poorly modulated behavioral expression of anger. Differences between the two groups on scale elevation patterns also reflected higher percentages of incarcerated youths who described themselves as having poor school performance with limited intellectual abilities or learning problems, as feeling different or alienated from others, as experiencing psychological distress through somatic complaints, and as experiencing social isolation or conflict with peers. This study also found that “neither a sex effect in regular education students, nor a language effect for males in regular education classrooms versus a juvenile justice facility obtained significance” (p. 258).

In another study of incarcerated youths (Marsh, 2002) that included 193 adolescents (151 boys, 42 girls), the PIY DLQ scale was found to have made a unique contribution, beyond demographic and offense history information, to the prediction of moderate and major offenses of juvenile offenders during their first 6 months of incarceration. The PIY added 10% variance beyond clinical diagnoses, demographic information, and offense history in predicting major offenses. For females, the PIY identified those who committed major offenses with 86% accuracy. The DLQ scale was the highest of all clinical scales in both males and females (mean $T$-scores of 56.1 and 61.5, respectively), followed by Psychological Discomfort (DIS) in males and by Family Dysfunction (FAM) and DIS in females. Significant differences between the clinical scale scores of White and non-White youths were not found. As noted by Marsh (2002), the results of this study, overall, demonstrate the potential value of the PIY as a screening instrument for incarcerated youths in juvenile justice placements.

In a study of 122 female adolescent offenders who were incarcerated in a correctional facility, Aikman (2000) found that the PIY Impulsivity and Distractibility (ADH) scale predicted the number and severity of recorded disciplinary incidents, and the DLQ scale predicted the severity of recorded disciplinary incidents, within the first 3 months of incarceration. In a discriminant analysis, the PIY correctly classified 79% of the adolescents who were assigned to minimum security-level group versus medium/maximum security-level group (based on ratings by case manager and treatment team). The PIY was able to discriminate a maximum from a combined minimum/medium security group with 86% classification accuracy. Based upon the variables that contributed significantly in discriminating between security levels (i.e., ADH, DLQ, and Reality Distortion [RLT]), maximum security youths were characterized as more aggressive, impulsive, overactive, and distrustful than medium and minimum security-level youths. The maximum security youths were characterized also as more likely to confuse internal experience with reality. The mean DLQ scale $T$-score for the total group of female offenders was 64.9 ($SD = 13.3$). Mean
**Item Overlap on the PIY**

Whereas the PIY has non-overlapping clinical scales and subscales, there is considerable item overlap between both the Dissimulation (FB) and the DEF validity scales with some of the clinical scales. As is the case with other self-report inventories such as the MMPI-2 (Butcher et al., 2001) and the MMPI-A, which contain substantial item overlap, item overlap between scales creates built-in, spurious correlations between such scales (Budescu & Rodgers, 1981; Helmes & Reddon, 1993; Simms, Casillas, Clark, Watson, & Doebbeling, 2005), with increased statistical correlation corresponding with increased item overlap (Ben-Porath, 2012; Weiner & Greene, 2008). Item overlap resulting in spurious correlations between scales weakens the distinctiveness of the scales (Helmes & Reddon, 1993; Tellegen et al., 2003) and can lead to a factor structure that may be partly artificial (Retzlaff & Gibertini, 1987). Item redundancy results in a lack of experimental independence on item scores as each test-taker gives a single response to each item regardless of the fact that the same item is included on more than one scale (Hsu, 1994). Although item overlap on the PIY is much more circumscribed in comparison with item overlap on the MMPI-A, it is important to recognize the potential effects of item overlap between two of the validity scales and the clinical scales on the PIY.

On the PIY DLQ scale, 9 of its 42 items (11%) are contained within the DEF scale, keyed in opposite direction. All of these 9 overlapping items are contained within one Delinquency subscale, DLQ3, an 11-item subscale which is
a measure of non-compliance (e.g., breaking rules, testing limits, manipulation of others). Thus, 9 of the 11 items (81%) on DLQ3 are contained within the DEF scale. Considering the items on the DEF scale, 9 of its 24 items (37.5%) are contained within the DLQ scale, specifically within the DLQ3 subscale. The DEF scale also shares three items with the COG scale.

The following correlation coefficients are reported in the PIY technical manual between the validity scales and the clinical scales and subscales. The correlation between DLQ and DEF is −.73 for the regular education group (n = 2,327) and −.74 for the clinically referred group (n = 1,178). The correlation between DLQ3 and DEF is −.82 for the regular education group, and −.86 for the clinically referred group. Thus, there is approximately 53% shared variance between the DLQ and DEF scales and approximately 67% to 74% shared variance between DLQ3 and DEF (for the regular education and clinically referred groups, respectively).

A number of item overlap correlation or item overlap coefficient (IOC) formulas have been proposed as ways of measuring correlation of scales due to item overlap, taking into account the subset of shared items. These IOC formulas estimate the amount of overlap between scales with shared items, although the IOC does not necessarily equate with or determine the minimum possible value of the correlation between two overlapping scales (Hsu, 1994). Hsu (1994) states that his IOC formula (Equation 7) is in accord with Budescu and Rogers’ (1981) formula and is applicable to all pairs of scales, X and Y, which share items, including pairs of scales all of whose shared items are keyed in the opposite direction. Applying this equation in which Na = # of items unique to Scale X, Nb = # of items unique to Scale Y, Nc = # of items shared and keyed in the same direction, Nd = # of items shared and keyed in the opposite direction, and (Nc + Nd) = # of items common to both scales, yields a phi coefficient measuring the association for two binary variables, defined by Hsu as the inter-subset inter-item correlations. The phi coefficient is calculated by dividing Nc − Nd by the square root of the product of (Na + Nc + Nd) and (Nb + Nc + Nd).

Applying Equation 7 (Hsu, 1994) for scales whose shared items are all keyed in the opposite direction, for the relationship between the DLQ and DEF scales, yields an IOC coefficient of −.28, a magnitude of small to medium effect size. However, applying the same equation for the relationship between DLQ3 and DEF yields an IOC coefficient of −.56, a magnitude of large effect size. Thus, it appears that as a result of item overlap, the DLQ3 subscale and the DEF scale have a baseline (negative) correlation at the magnitude of a large effect size. Such substantial shared variance and lack of distinctiveness between these scales raise several questions. First, what is being measured distinctly by the DLQ3 subscale? Second, does the DLQ3 subscale have a built-in confusion of response style? Third, can the DLQ3 subscale be considered as a proxy measure of admission versus denial of problems in the area of compliance with rules and limits expected by adults? Still, the correlations between DEF and DLQ1 (−.50 and −.53 for the regular education and clinically referred samples, respectively), and between DEF and DLQ2 (−.57 and −.59 for the regular education and clinically referred groups, respectively), while these two DLQ subscales share no items with DEF, support the construct validity of the DEF scale. It is also noted that DEF has negative correlations with all of the clinical scales and subscales, which is consistent with the construct being measured by the DEF scale. The DLQ scale has differential correlations with the other clinical scales at magnitudes that are consistent with the constructs that are theoretically measured by those scales. Therefore, it appears to be that it is particularly the substantial covariance between the DEF scale and DLQ3 that obscures the distinctiveness of what is being measured by each scale.

Although it is unusual for youths who are being evaluated for legal dispositional purposes (rather than for purposes of emergency mental health evaluation, competence to stand trial, or waiver to criminal court) to over-report psychological and behavioral problems (Lachar et al., 2006; Pinsoneault, 1996), it is nevertheless important to examine test profiles for possible dissimulation, exaggeration, or over-reporting of problems. Exaggerated or distorted over-reporting of problems may result from an effort to draw attention to current fear, emotional pain, or psychological distress (Friedman, Lewak, Nichols, & Webb, 2001), and it may reflect a person’s experiential style to “overreact and to be traumatized” (Greene, 2011, p. 78). The FB validity scale on the PIY consists of 42 items, 37 of which overlap with, and are distributed fairly evenly among the clinical subscales. However, on the shorter scales and subscales, such overlap may contribute to spuriously high shared variance. In particular, the Social Skill Deficits subscale 2 (SSK2) has 11 items, 5 of which (45%) are shared with the FB scale. The SSK2 subscale has correlations of .69 with the FB scale in both the regular education sample and the clinically referred sample, indicating that SSK2 has 47% shared variance with the FB scale. Similarly, the Reality Distortion subscale 2 (RLT2) is composed of 11 items, 5 of which (45%) are shared with the FB scale. The correlations between RLT2 and FB are .64 for the regular education sample and .70 for the clinically referred sample, respectively, indicating 41% shared variance and 49% shared variance between the two scales, respectively. Thus, when scores on both the FB and RLT scales are elevated, the distinctions between exaggeration of problems and severe maladjustment may become more difficult to discern.

Effectiveness of the PIY in Assessing Delinquency/Conduct Problems in Juvenile Delinquents and a Comparative Examination of Research With the MMPI-A

As indicated earlier, some investigators have questioned the sensitivity of the PIY DEF scale in detecting under-reporting in juvenile delinquents, or whether the PIY is effective in
obtaining accurate self-reports of the type of serious delinquent behavior problems that are typically displayed by youths who are adjudicated delinquent. Branson and Cornell (2008) found the low number of profiles in their study that had elevated DLQ scores to be unexpected and surprising. They further noted that the PIY was not designed as an instrument to assess juvenile delinquents. Marsh (2002) similarly stated that the DLQ scale items mostly sample behaviors which do not characterize severe delinquent offense behaviors. Indeed, Marsh considered the possibility that females in her sample scored higher than males on the DLQ scale in part because DLQ items tend to represent status offense behaviors which occur more frequently among female offenders. Branson and Cornell nevertheless concluded that the DLQ scale appears to have clinical utility in juvenile justice samples.

It appears that defensive response style and accurate reporting of adjustment problems on the PIY are salient, inter-related issues. As noted previously, it is anticipated that youths undergoing mental health evaluation in a forensic context, especially in which the outcome of the evaluation will have bearing on legal consequences, with potential for restriction of their freedom, may attempt to present themselves as free of psychological or behavioral problems (Devieux et al., 2002; Lachar et al., 2006; Penney & Skilling, 2012; Smith, 2007). Also, in forensic settings one may encounter a higher base rate of youths who exhibit psychopathic personality traits, such as dishonesty, superficiality, deceptive self-presentation, and who may produce greater response biases on self-report measures of psychopathology (Murrie & Cornell, 2002; Penney & Skilling, 2012; Pinsoneault, 1996). In addition, self-report scales generally consist of items that are transparent in the endorsement of socially undesirable qualities which can further lead to under-reporting or denial of problems (Murrie & Cornell, 2002). Thus, the contextual demand characteristics of a court-ordered evaluation in conjunction with personality issues of many delinquent adolescents and test item transparency may result in many youths attempting to portray an unrealistically favorable image of themselves, while denying problems in response to various assessment methodologies. A defensive or uncooperative attitude on the part of a youth being evaluated for the court affects the validity and usefulness of information whether it is obtained through clinical interview or through self-report measures. On the PIY in particular, there is also a “built-in” inverse relationship between defensiveness and endorsement of delinquent-type test items due to item overlap as was discussed earlier. How this affects the sensitivity of one or both of these scales in any specific case needs to be examined within the fuller assessment context.

A broader view of the juvenile delinquency assessment literature finds that the absence of elevations on both the DEF and the DLQ scales of the PIY in the majority of youths studied in some samples appears to be consistent with findings on the MMPI-A in juvenile justice samples. In a review of the MMPI-A and the Millon Adolescent Clinical Inventory (MACI; Millon, 1993) in juvenile justice samples, Baum, Archer, Forbey, and Handel (2009) reported mean and standard deviation scores from 20 or more samples. With respect to measurement of systematic response bias, particularly under-reporting, the mean T-scores on the MMPI-A L and K scales across multiple studies (52.9 and 51.5, respectively) were comparable with the average MMPI-A scores of non-referred, non-forensic youths, typically 2 or 3 points above the mean of 50, but within the standard error of the mean. Thus, the absence of elevations on the DEF scale of the PIY in the majority of delinquent youths studied across different samples is in line with findings on the MMPI-A. With respect to those MMPI-A basic clinical scales that are most highly associated with acting out, impulsive, non-conforming, aggressive, disobedient, resentful, argumentative, interpersonally distrustful behavior patterns, Baum et al.’s review found that the average T-scores on the clinical scales assessing these features were all below T60 but were mildly elevated (in the range of T55-T60) and were the most elevated scales in the profile. Based on 22 samples, the highest mean T-score was on Scale 4 (Pd: Psychopathic Deviate; M = 58.8, SD = 9.8), followed by a mean T-score of 57.4 (SD = 10.2) on Scale 6 (Pa: Paranoia), and a mean T-score of 55.7 (SD = 11.3) on Scale 9 (Ma: Hypomania). Thus, those MMPI-A scales which have conceptual or behavioral correlate overlap with the PIY DLQ scale were comparable in elevation to T-scores on the DLQ scale in several studies using the PIY. Also, in the studies that were cited earlier, the DLQ scale, while not significantly elevated in some studies, was the highest scale in the profile, as would be expected in juvenile justice samples, analogous to Scale 4 of the MMPI-A being the scale with the highest mean score in the review by Baum et al. The foregoing review by Baum et al. helps to put into perspective findings with the PIY from the same population of youths.

Furthermore, in consideration of Tyndall’s (2002) concern about the sensitivity of the PIY DEF scale in detecting under-reporting in juvenile justice samples, difficulty in accurately detecting “fake good” profiles has been reported concerning the use of the MMPI-A (Archer, 2005; Herkov, Archer, & Gordon, 1991; Pinsoneault, 1996; Stein & Graham, 1999). Stein and Graham (1999) suggest that when considering the presence of a fake-good, or under-reporting response style in youths placed in juvenile justice settings, cutoff scores lower than T ≥ 65 on both the L and K scales as used in standard clinical practice may be more effective. Limitations as noted by Pinsoneault (1996) when attempting to classify delinquent youths into response style groups based on MMPI-A validity scale criteria underscores the point that cutoff criteria developed to detect “faking” sets, while they may serve as good heuristics based on empirical evidence, are not perfectly sensitive or accurate.
It is most relevant and instructive to note that what has been termed a Within Normal Limits (WNL) profile on the MMPI-A, that is, a MMPI-A profile in which neither the validity scales nor the clinical scales are elevated, has been described by Archer, Bolinskey, Morton, and Farris (2003) as a "ubiquitous feature for adolescents evaluated with the MMPI-A across a wide variety of psychiatric, substance abuse, and juvenile delinquency settings (Archer, 1997)" (p. 408). This observation by Archer et al. (2003) may well have relevance to other standardized, self-report personality assessment inventories, including the PIY. In this light, the absence of significant elevation on the DLQ scale in some studies of adjudicated delinquents is not unexpected, even while most of the PIY profiles for the same youths did not demonstrate elevated scores on the DEF scale, a finding which is comparable with what has been found in many studies with the MMPI-A.

One additional salient point concerning the utility of the PIY in juvenile delinquency assessments may be drawn from studies of the MMPI-A. In consideration that some studies of juvenile justice samples with the MMPI-A found non-clinically elevated scores on scales that assess externalizing psychopathology, Pena, Megargee, and Brody (1996) asserted that "the critical question . . . is not the absolute elevation of delinquents’ MMPI-A profiles, but whether MMPI-A scales can differentiate delinquents from nondelinquents" (p. 390). In this regard, while studies such as those conducted by Branson and Cornell (2008) and by Tyndall (2002) have pointed to some of the limitations with use of the PIY in juvenile justice contexts, Tyndall did find that the DLQ scale discriminated between the juvenile forensic study sample and the normative regular education sample of males, which in essence satisfies the criteria for usefulness of a test as identified by Pena et al. (1996).

Conclusions, Suggestions, and Directions for Future Research

The PIY has been studied and applied in the evaluation of youths from a variety of juvenile justice settings and is used routinely in certain juvenile justice facilities as an evaluative tool (Lachar et al., 2006). This review cited some studies that have supported the validity of this instrument in identifying adolescents at risk or in need of treatment interventions for antisocial, delinquent behavior as well as for emotional/internalizing psychopathology. Among criticisms pertaining to the use of the PIY in juvenile justice contexts that were cited in this article were the findings of high percentages of what has been termed WNL profiles, given that all youths in the study samples had been adjudicated delinquent. Similarly, the failure of the PIY DLQ scale to distinguish between a delinquent sample and the clinically referred standardization sample was a criticism noted by Tyndall (2002). However, a review of studies in which the MMPI-A was used in the assessment of juvenile justice youths found an overall similar pattern of WNL profiles. In studies of both the PIY and the MMPI-A, large percentages of test profiles had both non-elevated validity scale scores and non-elevated clinical scale scores. Yet, these same studies also revealed a recurrent finding of mildly elevated scores (in the range of T55-T60) on those scales associated with externalizing behavior and delinquency problems. Nevertheless, such frequent findings of WNL profiles with various self-report measures have led some researchers to seriously question the validity of self-report measures of externalizing and internalizing psychopathology and substance abuse administered to youths in the juvenile delinquent population (Breuk, Clauser, Stams, Slot, & Doreleijers, 2007; Henggeler et al., 1993; Mieczkowski, Newel, & Wraith, 1998; Vreugdenhil, Van den Brink, Ferdinand, Wouters, & Doreleijers, 2006). Other investigators obtained results that support the utility of youth self-report measures in the delinquency assessment context (e.g., Aikman, 2000; Marsh, 2002; Negy et al., 2001; Robertson, Dill, Husain, & Undesser, 2004). Cashel (2003) obtained results with the YSR, suggesting that self-report may provide critical information in the identification of treatment needs for court-probated youth and also that obtaining multiple sources of data is needed for identifying treatment needs and interventions. The validity of self-report measures with juvenile delinquents is an area that remains open for further study. One particular direction for future research with the PIY and with other test instruments in the juvenile delinquency context is the possible identification of youth characteristics that are associated with closed, defensive response styles, and open, candid response styles. Response style might be conceptualized as a dimensional rather than a categorical construct, with no clear point of distinction between deliberate distortions and characterological response styles, or between one response style and another (Meyer, 1999; Stokes, Pogge, & Zaccario, 2013).

The sensitivity of the DEF scale on the PIY in the detection of under-reporting in the juvenile delinquency context was also questioned by Tyndall (2002). Although the PIY manual states that T-scores ≥ 60 on the DEF scale suggest defensiveness and an effort to portray exceptional adjustment, under-reporting or minimization of problems might be suggested by scores that are slightly lower (Lachar & Boyd, 2005). This may be the case especially when there are other indications of minimization of problems, for example, during interview or with other assessment instruments. In this vein, it should be noted that a difference of 1 raw score point on the DEF scale is associated with a T-score of 58 versus 60 in the male norms and a T-score of 59 versus 61 in the female norms. When considering also the Standard Error of Measure (SEM) for the DEF scale (4.5 and 4.6 T-score points in the male norms and a T-score of 58 versus 60 on the DEF scale is associated with a T-score of 56 cannot technically be considered invalid, and there currently is no empirical basis
for lowering cutoff scores on validity scales of the PIY, the evaluator might note the increasing probability of underreporting. This, especially if finding a notably suppressed profile on the clinical scales, with scores below the average scores in comparison with the regular education and the clinically referred normative groups. While caution is warranted to avoid false positive inferences derived from test scores, it may be conceded that many false negative test results appear in the delinquency evaluation context. The preceding guidelines provide a basis for this discussion in the forensic report.

In addition, it may be helpful to compare scores on the DLQ1 and the DLQ3 subscales when conducting an evaluation, particularly when DEF is at least mildly to moderately elevated—DLQ3, but not necessarily DLQ1 or DLQ2 likely will be at least mildly to moderately suppressed. It is encouraging to note that even with an elevated DEF T-score of 64, and a T-score of 54 on DLQ, including a T-score of 43 on DLQ3, one can find a T-score of 64 on DLQ1 (seen in an actual forensic case profile). Further research may examine whether a lower cutoff score on the DEF scale or some other configuration of scores might improve the detection of under-reporting on the PIY.

The PIY is a psychometrically sound measure for evaluating a broad range of behavioral and psychological characteristics relevant in clinical but also forensic contexts. Along with comorbid disruptive behavior disorders, mood and anxiety disorders, PTSD, and substance abuse disorders, many adolescents in juvenile justice settings also have histories of learning or academic problems, problems with attention, concentration, hyperactivity, dysregulation of affect and impulses, parent–child/family conflicts, social skill deficits, peer conflicts, alienation, anger, and distrust of adults. With the exception of trauma-specific symptoms and substance use problems, all of the other preceding problem areas are addressed in the PIY scales and subscales. In this vein, it has been asserted that identification of particular treatment needs may be the greatest value of the PIY in its application in the juvenile justice system (Lachar et al., 2006). Given the finding of greatly increased risk for young adult recidivism in juveniles who at baseline presented with comorbid internalizing and disruptive behavior disorders, Hoeve, McReynolds, and Wasserman (2013) concluded that “the mental health needs of juvenile justice youths’ internalizing and externalizing problems should be addressed” (p. 1368). In this connection, the PIY may be a very useful instrument to be employed in juvenile justice assessments. As noted in some of the studies using the PIY in juvenile justice contexts (i.e., Branson & Cornell, 2008; Marsh, 2002), the DIS scale was the second most highly elevated scale, following the DLQ scale. In the study by Branson and Cornell (2008), the DIS scale displayed moderate correspondence with clinical diagnoses of depressive disorders.

The identification of treatment needs for youths with substance abuse problems is an important consideration in the forensic evaluation. Adolescent drug abuse is often associated with co-occurring mental health problems, including ADHD, oppositional-defiant disorder, conduct disorder, depressive and anxiety disorders (U.S. Department of Health and Human Services, National Institute on Drug Abuse, 2012). Substance abuse also has been shown to be a risk factor in criminal recidivism in juveniles (Cottle, Lee, & Heilbrun, 2001). Youths involved in the juvenile justice system who have a substance abuse disorder, with or without co-occurring disorders, have been found to be at greater risk for escalation in offense seriousness over time during adolescence (Hoeve, McReynolds, Wasserman, & McMillan, 2013). On the PIY, substance abuse problems are addressed on two DLQ scale items and on two non-scored “New Development Items”; however, there is no scale that measures alcohol or other drug use. Future research with the PIY might examine the utility of the four substance use items on the PIY as a subscale for identifying youths with substance abuse problems. Without minimizing the importance of assessing possible problems of substance abuse, the absence of scales to assess for alcohol or other drug problems on the PIY may clearly be viewed as a limitation, but not necessarily as a weakness, per se, of this measure. Some form of screening for substance use should be conducted routinely during delinquency evaluations (at least through interview, in which youths may or may not be truthful), and, in addition to collateral sources of information, there are self-report measures available that specifically assess for potential problems of alcohol or substance abuse, such as the Substance Abuse Subtle Screening Inventory–Adolescent 2 (SASSI-A2; Miller & Lazowski, 2001) and the Drug Abuse Screening Test for Adolescents (DAST-A; Martino, Grilo, & Fehon, 2000). It should also be considered empirically desirable to employ multiple methods of assessment of a particular problem area. Indeed, robust support for the valid interpretation of test results can be best demonstrated through multi-trait–multi-method (MTMM) procedures (Campbell & Fiske, 1959). Thus, self-report of substance use problems contained within the same self-report inventory that assesses other problems, in the absence of heteromethod assessment methodology, would be considered limited due to shared method variance (Furr, 2011).

The absence of scales on the PIY to assess for possible trauma-specific symptoms may also be considered as a limitation in juvenile delinquency contexts. As reported by Abram et al. (2013), 11.2% of a sample of juvenile detainees experienced PTSD within the previous year. Among youths who experienced PTSD, 93% had at least one comorbid psychiatric disorder. The forensic evaluator in juvenile delinquency contexts should be aware that the PIY does not include scales that assess trauma-specific symptoms. Particularly when there are indications of a history of trauma exposure or potentially traumatic experiences, self-report by the adolescent during interview, which may or may not include specific structured diagnostic interview for PTSD symptoms, in conjunction with report by parent and other
sources of historical information, and perhaps the administration of a self-report inventory such as the Trauma Symptom Checklist for Children (TSCC; Briere, 1996), may be useful. Whereas the administration of multiple self-report inventories in addition to clinical interview introduces a good deal of shared method variance, and might not always be feasible, some adolescents are more open to disclosing certain problem areas through indirect, questionnaire format rather than through interview (Lachar et al., 2006). Further research may also help determine whether youths in the juvenile justice system who have confirmed histories of traumatic experiences or those who have diagnoses of PTSD display particular test score configurations on the PIY. A search of the currently available research literature found one study of sexually abused children (Fricker & Smith, 2001), who were not involved in the juvenile justice system, in which the TSCC and the PIY were compared for their assessment of trauma-related symptoms and validity of self-reports. The TSCC was found to be more sensitive than was the PIY to PTSD status, whereas the PIY validity scales were more effective in detecting under-reporting and over-reporting response styles.

This review finds that there is a need for more current research of the PIY in juvenile justice contexts. There have been a limited number of studies using the PIY in samples of juvenile delinquents and most of these studies were published within the first decade following publication of the PIY. An online search by this author identified a recent dissertation (Martin, 2013) which studied juvenile sex offenders and which included the PIY along with other test measures. The PIC-2, from which the PIY was developed, has been studied in recent years in association with callous unemotional (CU) traits (Baron, 2010); however, it is not known whether such study has been extended to the PIY.

In addition, further study of ethnic differences on the PIY in samples of delinquent adolescents is warranted. As noted earlier, the test authors reported unsubstantial differences with respect to most demographic variables, including ethnicity. In the previously cited study by Aikman (2000), incarcerated African American female juvenile offenders obtained higher scores than Caucasian female juvenile offenders on some validity and clinical scales. In contrast, in the study by Branson and Cornell (2008), White youths scored higher than Black youths on several clinical scales, whereas the Black youths scored higher than the White youths on the DEF scale. In a study of prevalence rates of psychiatric disorders in detained youth, Teplin et al. (2002) found that rates of many disorders were higher in females, non-Hispanic Whites, and older adolescents. Similarly, in a large sample of incarcerated adolescents ($n = 5,964$, 95% male), Karnik, Jones, Campanaro, Haapanen, and Steiner (2006) found lower levels of self-reported psychiatric problems in ethnic minorities (i.e., African Americans and Hispanics) than in Caucasians, as assessed with the YSR. In another large-scale study of juvenile justice youth ($n = 9,819$), Wasserman et al. (2010) found that Caucasian youth reported higher rates of most psychiatric disorders compared with African American youth. As noted by Colins et al. (2010), considering the over-representation of ethnic minorities in the juvenile justice system, there is a need for further study of whether clinically relevant differences exist on the basis of race or ethnicity.

Some final thoughts are mentioned concerning the validity of the PIY for use in juvenile forensic evaluations. It has been noted that “validity is an ongoing process wherein one provides evidence to support the appropriateness, meaningfulness and usefulness of the specific inferences made from scores about individuals from a given sample and in a given context” (Zumbo, 2007, p. 48). Furthermore, the validity of test scores should be qualified as a matter of degree that can be weighed in terms of the strength of the evidence (Furr, 2011; Zumbo, 2007). Validity of test scales and scores should not be regarded as absolute, or as all or none, and there is no unique value that determines validity (Furr, 2011; Zumbo, 2007). It also becomes understood that “invalidity is something that distorts the meaning of test results for some groups of examinees in some contexts for some purposes” (Zumbo, 2007, p. 48). An appropriate question to ask might be what inferences drawn from what tests have been consistently validated in samples of youths in juvenile justice settings. With respect to the PIY, while further study of this instrument in juvenile justice samples is warranted, it is noteworthy that in the studies reviewed here the scales with the highest elevations were the DLQ, DIS, and FAM scales, corresponding with externalizing and antisocial behavior problems, internalizing and depressive-type problems, and family conflicts and dysfunction, which have high base rates in the juvenile delinquency population. When conducting individual assessments, test results will be invalid in some cases, for example, false negative findings of no psychopathology based on test scores alone. When such data are inconsistent with historical, clinical, or other data, alternate inferences should be considered. However, a test score indicating no psychopathology which is inconsistent with other assessment data does not necessarily reflect an invalid test or scale(s). Rather, it may be that content-based or non-content-based indications of test score invalidity stem from the possibilities that a particular youth might not have adequately read and/or comprehended the content of test items. She or he may have responded carelessly to test items or she or he may have opted to respond randomly or in a partially random manner. She or he may have consciously and/or unconsciously denied or minimized problems, so as to portray a consistent, but inaccurate picture of good, if not excellent adjustment. It may be that the youth was characterized by demographic features that were very discrepant from those of the normative sample, etc. Following from the above, instruments such as the PIY, but also including the MMPI-A, may be considered to be undergoing an “ongoing process” of validity with the juvenile delinquency population.

Semel
As was stated at the outset of this article, psychological testing may be a useful component of an overall comprehensive, multi-method, multi-informant forensic evaluation. The PIY was developed as one of three companion multi-informant instruments in the assessment of a broad range of psychological and behavioral adjustments of youths. When considering whether to use the PIY, or any other standardized test instrument in the juvenile delinquency assessment context, the evaluator may consider the following statement: “Forensic assessment of emotional and behavioral adjustment relies on the assumption that psychometric evaluation improves upon the accuracy of subjective assessments, such as unstructured interviews” (Lachar et al., 2006, p. 274). Whereas the forensic evaluator cannot be assured a priori of incremental validity in the assessment through the use of a standardized test measure, it is known that a certain percentage of youths being evaluated in the legal context provide valid reports that are useful in identifying internalizing and externalizing problems and treatment needs, which in turn can assist the court in its decisions for youths.

Acknowledgment
The author expresses appreciation to Dr. Marc J. Diener for his consultation concerning interscale item overlap.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research and/or authorship of this article.

Notes
1. From Hsu (1994, p. 129, p. 130, p. 140). Taylor & Francis LLC (http://www.tandfonline.com). Adapted with permission.
2. The conversion of raw scores to T-scores does not include a T-score of 55.

References
Abraham, K. M., Teplin, L. A., King, D. C., Longworth, S. L., Emanuel, K. M., Romero, E. G., . . . Olson, N. D. (2013). PTSD, trauma, and comorbid psychiatric disorders in detained youth. Available from ojjdp.gov
Abraham, K. M., Teplin, L. A., McClelland, G. M., & Dulcan, M. K. (2003). Comorbid psychiatric disorders in youth in juvenile detention. Archives of General Psychiatry, 60, 1097-1108.
Achenbach, T. M. (1991). Manual for the youth self-report and 1991 profile. Burlington: Department of Psychiatry, University of Vermont.
Adams, Z. W., McCart, M. R., Zajac, K., Danielson, C. K., Sawyer, G. K., Saunders, B. E., & Kilpatrick, D. G. (2013). Psychiatric problems and trauma exposure in nondetained delinquent and nondelinquent adolescents. Journal of Clinical Child & Adolescent Psychology, 42, 323-331. doi:10.1080/15374416.2012.749786
Aikman, G. G. (2000). Predicting adjustment to incarceration in female juvenile delinquents: The Personality Inventory for Youth and the Youth Self-Report (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9980100)
Archer, R. P. (2005). MMPI-A: Assessing adolescent psychopathology (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
Archer, R. P., Bolinsky, P. K., Morton, T. L., & Farris, K. L. (2003). MMPI-A characteristics of male adolescents in juvenile justice and clinical treatment settings. Assessment, 10, 400-410. doi:10.1177/1073191103256128
Baron, E. M. (2010). Capturing callous unemotional (CU) traits and the behavioral inhibition/activation systems (BIS/BAS) in children: Development of three construct-oriented scales in the Personality Inventory for Children–Second Edition (PIC-2). Retrieved from http://digitalcommons.pace.edu/dissertations/AAI3428924
Baum, L. J., Archer, R. P., Forbey, J. D., & Handel, R. W. (2009). A review of the Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A) and the Millon Adolescent Clinical Inventory (MACI) with an emphasis on juvenile justice samples. Assessment, 16, 384-400. doi:10.1177/107319110938264
Ben-Porath, Y. S. (2012). Interpreting the MMPI-2-RF. Minneapolis: University of Minnesota Press.
Branson, C. E., & Cornell, D. G. (2008). Correspondence of the Millon Adolescent Clinical Inventory (MACI) and the Personality Inventory for Youth (PIY) in juvenile offenders. Journal of Forensic Psychology Practice, 7, 65-100. doi:10.1300/J158v07n04_02
Breuk, R. E., Clauser, C. A. C., Stams, G. J. M., Slot, N. W., & Doreleijers, T. A. H. (2007). The validity of questionnaire self-report of psychopathology and parent-child relationship quality in juvenile delinquents with psychiatric disorders. Journal of Adolescence, 30, 761-771.
Briere, J. (1996). Trauma Symptom Checklist for Children (TSCC): Professional manual. Odessa, FL: Psychological Assessment Resources.
Budescu, D. V., & Rodgers, J. L. (1981). Corrections for spurious influences on correlations between MMPI scales. Multivariate Behavioral Research, 16, 483-497. doi:10.1207/s15327906-mbr1604_4
Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A. M., Dahlstrom, W. G., & Kaemmer, B. (2001). MMPI-2: Manual for administration and scoring (Rev. ed.). Minneapolis: University of Minnesota Press.
Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R. P., Tellegen, A., Ben-Porath, Y. S., & Kaemmer, B. (1992). Minnesota Multiphasic Personality Inventory-Adolescent (MMPI-A): Manual for administration, scoring, and interpretation. Minneapolis: University of Minnesota Press.
Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.
Cashel, M. L. (2003). Validity of self-reports of delinquency and socio-emotional functioning among youth on probation. Journal of Offender Rehabilitation, 37, 11-23. doi:10.1300/J076v37n01_02
Coker, K. L., Smith, P. H., Westphal, A., Zonana, H. V., & McKee, S. A. (2014). Crime and psychiatric disorders among youth in the US population: An analysis of the National Comorbidity Survey-Adolescent Supplement. Journal of the American
Colins, O., Vermeiren, R., Vreugdenhil, C., Van den Brink, W., Doreleijers, T., & Broekaert, E. (2010). Psychiatric disorders in detained male adolescents: A systematic literature review. *The Canadian Journal of Psychiatry, 55*, 255-263.

Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior, 40*, 57-79.

Cruise, K. R. (2006). Special issues in juvenile justice. *Applied Psychology in Criminal Justice*, 2, 177-204. Retrieved from http://fordham.bepress.com/psych_facultypubs/12

Destaefano, L. (1998). Review of the Personality Inventory for Youth. In J. C. Impara & B. S. Plake (Eds.), *The thirteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements. Retrieved from Mental Measurements Yearbook with Tests in Print database.

Devieux, J., Malow, R., Stein, J. A., Jennings, T. E., Lucenko, B. A., Averhart, C., & Kalichman, S. (2002). Impulsivity and HIV risk among adjudicated alcohol- and other drug-abusing adolescent offenders. *AIDS Education and Prevention, 14*, 24-35.

Fricker, A. E., & Smith, D. W. (2001). Trauma specific versus generic measurement of distress and the validity of self-reported symptoms in sexually abused children. *Journal of Child Sexual Abuse, 10*, 51-66.

Friedman, A. F., Lewak, R., Nichols, D. S., & Webb, J. T. (2001). *Psychological assessment with the MMPI-2*. Mahwah, NJ: Lawrence Erlbaum.

Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. Thousand Oakes, CA: SAGE.

Greene, R. L. (2011). The MMPI-2/MMPI-2-RF: An interpretive manual (3rd ed.). Boston, MA: Allyn & Bacon.

Grisso, T. (2005). Evaluating the properties of instruments for screening and assessment. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 71-94). New York, NY: Guilford Press.

Grisso, T. (2013). *Forensic evaluation of juveniles* (2nd ed.). Sarasota, FL: Professional Resource Press.

Grisso, T., & Vincent, G. (2005). The context for mental health screening and assessment. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 44-70). New York, NY: Guilford Press.

Hathaway, S. R., & McKinley, J. C. (1943). The *Minnesota Multiphasic Personality Inventory manual*. New York, NY: Psychological Corporation.

Heilbrun, K., Marczyk, G. R., DeMatteo, D., Zillmer, E. A., Harris, J., & Jennings, T. (2003). Principles of forensic mental health assessment: Implications for neuropsychological assessment in forensic contexts. *Assessment, 10*, 329-343.

Helmes, E., & Reddon, J. R. (1993). A perspective on developments in assessing psychopathology: A critical review of the MMPI and MMPI-2. *Psychological Bulletin, 113*, 453-471. doi:10.1037/0033-2909.113.3.453

Henggeler, S. W., Melton, G. B., Smith, L. A., Foster, S. L., Hanley, J. H., & Hutchinson, C. M. (1993). Assessing violent offending in serious juvenile offenders. *Journal of Abnormal Child Psychology, 21*, 233-243.

Herkov, M. J., Archer, R. P., & Gordon, R. A. (1991). MMPI response sets among adolescents: An evaluation of the limitations of the subtle-obvious subscales. *Psychological Assessment, 3*, 424-426. doi:10.1037/1040-3590.3.3.424

Hoeve, M., McReynolds, L. S., & Wasserman, G. A. (2013). The influence of adolescent psychiatric disorder on young adult recidivism. *Criminal Justice and Behavior, 40*, 1368-1382.

Hoeve, M., McReynolds, L. S., Wasserman, G. A., & McMillan, C. (2013). The influence of mental health disorders on severity of reoffending in juveniles. *Criminal Justice and Behavior, 40*, 289-301.

Hoge, R. D., & Andrews, D. A. (2010). *Evaluation for risk of violence in juveniles*. New York, NY: Oxford University Press.

Hsu, L. M. (1994). Item overlap correlations: Definitions, interpretations, and implications. *Multivariate Behavioral Research, 29*, 127-140.

Karnik, N. S., Jones, P. A., Campanaro, A. E., Haapanen, R., & Steiner, H. (2006). Ethnic variation of self-reported psychopathology among incarcerated youth. *Community Mental Health Journal, 42*, 477-486. doi:10.1007/s10597-006-9056-5

Lachar, D., & Boyd, J. (2005). Personality Inventory for Children. Second Edition; Personality Inventory for Youth; and Student Behavior Survey. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 205-223). New York, NY: Guilford Press.

Lachar, D., & Gruber, C. P. (1995). *Personality Inventory for Youth (PIY) manual: Administration and interpretation guide: Technical guide*. Los Angeles, CA: Western Psychological Services.

Lachar, D., & Gruber, C. P. (2001). *Personality Inventory for Children, Second Edition (PIC-2): Standard form and behavioral summary manual*. Los Angeles, CA: Western Psychological Services.

Lachar, D., Hammer, B. A., & Hammer, J. H. (2006). The Personality Inventory for Youth, the Personality Inventory for Children, Second Edition, and the Student Behavior Survey. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 263-296). Mahwah, NJ: Lawrence Erlbaum.

Lachar, D., Wingenfeld, S. A., Kline, R. B., & Gruber, C. P. (2000). *Student Behavior Survey: Manual*. Los Angeles, CA: Western Psychological Services.

Marsh, T. Y. (2002). *Prediction of institutional misbehavior among juvenile offenders with the Personality Inventory for Youth* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3044902)

Martin, R. (2013). *Male and female sex offenders: A characterization comparison* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3565666)

Martino, S., Grilo, C. M., & Fehon, D. C. (2000). Development of the Drug Abuse Screening Test for Adolescents (DAST-A). *Addictive Behaviors, 25*, 57-70. doi:10.1016/S0306-4603(99)00303-1

Melton, G. B., Petrila, J., Poythress, N. G., & Slobochin, C. (2007). *Psychological evaluations for the courts: A handbook for mental health professionals and lawyers* (3rd ed.). New York, NY: Guilford Press.

Meyer, G. (1999). The convergent validity of MMPI and Rorschach scales: An extension using profile scores to define response and character styles on both methods and a re-examination of simple Rorschach response frequency. *Journal of Personality Assessment, 72*, 1-35.

Mieczkowski, T., Newell, R., & Wright, B. (1998). Using hair analysis, urinalysis, and self-reports to estimate drug use in
a sample of detained juveniles. Substance Use & Misuse, 33, 1547-1567.

Miller, F. G., & Lazowski, L. E. (2001). The Adolescent Substance Abuse Subtle Screening Inventory-A2 (SASSI-A2) manual. Springfield, IN: SASSI Institute.

Millon, T. (1993). Millon Adolescent Clinical Inventory. Minneapolis, MN: National Computer Systems.

Murrie, D. C., & Cornell, D. G. (2002). Psychopathy screening of incarcerated juveniles: A comparison of measures. Psychological Assessment, 14, 390-396.

Negy, C., Lachar, D., Gruber, C. P., & Garza, N. D. (2001). The Personality Inventory for Youth (PIY): Validity and comparability of English and Spanish versions for regular education and juvenile justice samples. Journal of Personality Assessment, 76, 250-263.

Pena, L. M., Megargee, E. I., & Brody, E. (1996). MMPI-A patterns of male juvenile delinquents. Psychological Assessment, 8, 388-397.

Penney, S. R., & Skilling, T. A. (2012). Moderators of informant agreement in the assessment of adolescent psychopathology: Extension to a forensic sample. Psychological Assessment, 24, 386-400.

Pinsoneault, T. B. (1996). Rationally developed fake-good and fake-bad scales for the Jesness Inventory. Journal of Psychopathology and Behavioral Assessment, 18, 255-273.

Retzlaff, P. D., & Gibertini, M. (1987). Factor structure of the MCMI basic personality scales and common-item artifact. Journal of Personality Assessment, 51, 588-594.

Robertson, A. A., Dill, P. L., Husain, J., & Undesser, C. (2004). Prevalence of mental illness and substance abuse disorders among incarcerated juvenile offenders in Mississippi. Child Psychiatry & Human Development, 35, 55-74.

Simms, L. J., Casillas, A., Clark, L. A., Watson, D., & Doebbeling, B. N. (2005). Psychometric evaluation of the restructured clinical scales of the MMPI-2. Psychological Assessment, 17, 345-358. doi:10.1037/1040-3590.17.3.345

Smith, S. R. (2007). Making sense of multiple informants in child and adolescent psychopathology: A guide for clinicians. Journal of Psychoeducational Assessment, 25, 139-149.

Stein, L. A. R., & Graham, J. R. (1999). Detecting fake-good MMPI-A profiles in a correctional facility. Psychological Assessment, 11, 386-395.

Stokes, J. M., Pogge, D. L., & Zaccario, M. (2013). Response character styles in adolescents: A replication of convergent validity between the MMPI-A and the Rorschach. Journal of Personality Assessment, 95, 159-173. doi:10.1080/00223891.2012.730084

Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). The MMPI-2 restructured clinical (RC) scales: Development, validation, and interpretation. Minneapolis: University of Minnesota Press.

Teplin, L. A., Abram, K. M., McClelland, G. M., Dulcan, M. K., & Mericle, A. A. (2002). Psychiatric disorders in youth in juvenile detention. Archives of General Psychiatry, 59, 1133-1143.

Tyndall, M. L. (2002). Validation of the Personality Inventory for Youth (PIY) on a juvenile delinquent population (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3049750)

U.S. Department of Health and Human Services, National Institute on Drug Abuse. (2012). Principles of drug addiction treatment: A research-based guide. Retrieved from www.drugabuse.gov/sites/default/files/podat_1.pdf

Viljoen, J. L., McLachlan, K., & Vincent, G. M. (2010). Assessing violence risk and psychopathy in juvenile and adult offenders: A survey of clinical practices. Assessment, 17, 377-395. doi:10.1177/1073191110359587

Vreugdenhil, C., Van den Brink, W., Ferdinand, R., Wouters, L., & Doreleijers, T. (2006). The ability of YSR scales to predict DSM/DISC-C psychiatric disorders among incarcerated male adolescents. European Child & Adolescent Psychiatry, 15, 88-96. doi:10.1007/s00787-006-0497-8

Wasserman, G. A., McReynolds, L. S., Schwalbe, C. S., Keating, J. M., & Jones, S. A. (2010). Psychiatric disorder, comorbidity, and suicidal behavior in juvenile justice youth. Criminal Justice and Behavior, 37, 1361-1376. doi:10.1177/0093854810382751

Weiner, I. B., & Greene, R. L. (2008). Handbook of personality assessment. Hoboken, NJ: John Wiley & Sons.

Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), Handbook of statistics (Vol. 26, pp. 45-79). Amsterdam, The Netherlands: Elsevier.

**Author Biography**

**Robert A. Semel, Psy.D.** is a licensed psychologist (NY State) who conducts juvenile delinquency and child custody evaluations under the auspices of the Staten Island Family Court Services, New York State Office of Mental Health.