Morey, L.C. (1996). An Interpretive Guide to the Personality Assessment Inventory (PAI). Odessa: Psychological Assessment Resources, Inc.

CHAPTER 1
GENERAL INTRODUCTION AND OVERVIEW

The *Personality Assessment Inventory* (PAI; Morey, 1991) is a self-administered, objective test of personality and psychopathology designed to provide information on critical client variables in professional settings. From its inception, the PAI was developed to provide measures of constructs that are central in treatment planning, implementation, and evaluation. Although it was introduced fairly recently, the PAI already has generated considerable attention from clinicians and researchers, and the test has been described as a "substantial improvement from a psychometric perspective over the existing standard in the area" (Helmes, 1993, p. 417) and as "one of the most exciting new personality tests" (Schlosser, 1992, p. 12). The various applications of the test have generated findings that are important considerations in the interpretation of the test. The purpose of this interpretive guide is to integrate this recent work, and in doing so, to provide specific interpretive information about the use of the PAI in addressing questions central to the clinician and the researcher. This chapter will provide a summary of basic psychometric information about the test, including reliability and validity studies. Subsequent chapters will be devoted to the use of the PAI in addressing specific clinical issues.

The PAI: Rationale and Development

The development of the PAI was based upon a construct validation framework that emphasized a rational as well as quantitative method of scale development. This framework placed a strong emphasis on a theoretically informed approach to the development and selection of items and on the assessment of their stability and correlates. The theoretical articulation of the constructs to be measured was assumed to be critical, because this articulation had to serve as a guide to the content of information sampled and to the subsequent assessment of content validity. In this process, both the conceptual nature and empirical adequacy of the items played an important role in their inclusion in the final version of the inventory. The development of the test went through four iterations in a sequential construct

validation strategy similar to that described by Loevinger (1957) and Jackson (1971), although a number of item parameters were considered in addition to those described by these authors. Of paramount importance in the development of the test was the assumption that no single quantitative item parameter should be used as the sole criterion for item selection. An overreliance on a single parameter in item selection typically leads to a scale with one desirable psychometric property and numerous undesirable ones.

As an example, each PAI scale was constructed to include items addressing the full range of severity of the construct, including both its milder as well as most severe forms. Such coverage would not be possible if a single item selection criterion was applied; "milder" items would be most effective in distinguishing clinical subjects from normal respondents, while items reflecting more severe pathology would be more useful in discriminating among different clinical groups. Also, item-total correlations for such different items would be expected to vary as a composition of the sample, due to restriction of range considerations; milder items would display higher biserial correlations in a community sample, whereas more severe items would do so in an inpatient psychiatric sample. Thus, items selected according to a single criterion (e.g., discrimination between groups or item-total correlation) are doomed to provide limited coverage of the full range of symptomatology and/or severity of a clinical construct. The PAI sought to include items that struck a balance between different desirable item parameters, including content coverage as well as empirical characteristics, so that the scales could be useful across a number of different applications.

The clinical syndromes assessed by the PAI were selected on the basis of two criteria: the stability of their importance within the nosology of mental disorder, and their significance in contemporary diagnostic practice. These criteria were assessed through a review of the historical and contemporary literature as well as through a survey of practicing diagnosticians. In generating items for these syndromes, the literature on each clinical syndrome was examined to identify those components most central to the definition of the disorder, and items were written directed at providing an assessment of each component of the syndrome in question.

The test itself contains 344 items that are answered on a four-alternative scale, with the anchors *Totally False, Slightly True, Mainly True,* and *Very True*. Each response is weighted according to the intensity of the feature that the different alternatives represent. Thus, a client who answers *Very True* to the item "Sometimes I think I'm worthless" adds 3 points to his or her raw score on the Depression scale, whereas a client who responds *Slightly True* to the same item adds only 1 point. The use of this four-alternative scaling is justified psychometrically in that it allows

a scale to capture more true variance per item, meaning that even scales of modest length can achieve satisfactory reliability. It is also justified clinically, because sometimes even a *Slightly True* response to some constructs (e.g., as suicidal ideation) may merit clinical attention. Furthermore, clients themselves often express dissatisfaction with forced choice alternatives, expressing the belief that the true state of affairs lies somewhere "in the middle" of the two extremes presented.

The 344 items of the PAI comprise 22 nonoverlapping full scales: 4 validity, 11 clinical, 5 treatment consideration, and 2 interpersonal scales. Ten of the full scales contain conceptually derived subscales that were designed into the test to facilitate interpretation and coverage of the full breadth of complex clinical constructs. A brief description of the PAI full scales is provided in Table 1-1; Table 1-2 presents a description of the PAI subscales.

Table 1-1
PAI Full Scales and Their Descriptions

Scale (designation)	Description
Validity Scales	
Inconsistency (*ICN*)	Determines if client is answering consistently throughout inventory. Each pair consists of highly correlated (positively or negatively) items.
Infrequency (*INF*)	Determines if client is responding carelessly or randomly. Items are neutral with respect to psychopathology and have extremely high or low endorsement rates.
Negative Impression (*NIM*)	Suggests an exaggerated unfavorable impression or malingering. Items have relatively low endorsement rates among respondents in clinical settings.
Positive Impression (*PIM*)	Suggests the presentation of a very favorable impression or reluctance to admit minor flaws.
Clinical Scales	
Somatic Complaints (*SOM*)	Focuses on preoccupation with health matters and somatic complaints associated with somatization and conversion disorders.
Anxiety (*ANX*)	Focuses on phenomenology and observable signs of anxiety with an emphasis on assessment across different response modalities.
Anxiety-Related Disorders (*ARD*)	Focuses on symptoms and behaviors related to specific anxiety disorders, particularly phobias, traumatic stress, and obsessive-compulsive symptoms.
Depression (*DEP*)	Focuses on symptoms and phenomenology of depressive disorders.

(continued)

Table 1-1 (continued)
PAI Full Scales and Their Descriptions

Scale (designation)	Description
Clinical Scales (continued)	
Mania (*MAN*)	Focuses on affective, cognitive, and behavioral symptoms of mania and hypomania.
Paranoia (*PAR*)	Focuses on symptoms of paranoid disorders and more enduring characteristics of paranoid personality.
Schizophrenia (*SCZ*)	Focuses on symptoms relevant to the broad spectrum of schizophrenic disorders.
Borderline Features (*BOR*)	Focuses on attributes indicative of a borderline level of personality functioning, including unstable and fluctuating interpersonal relations, impulsivity, affective lability and instability, and uncontrolled anger.
Antisocial Features (*ANT*)	Focuses on history of illegal acts and authority problems, egocentrism, lack of empathy and loyalty, instability, and excitement-seeking.
Alcohol Problems (*ALC*)	Focuses on problematic consequences of alcohol use and features of alcohol dependence.
Drug Problems (*DRG*)	Focuses on problematic consequences of drug use (both prescription and illicit) and features of drug dependence.
Treatment Scales	
Aggression (*AGG*)	Focuses on characteristics and attitudes related to anger, assertiveness, hostility, and aggression.
Suicidal Ideation (*SUI*)	Focuses on suicidal ideation, ranging from hopelessness to thoughts and plans for the suicidal act.
Stress (*STR*)	Measures the impact of recent stressors in major life areas.
Nonsupport (*NON*)	Measures a lack of perceived social support, considering both the level and quality of available support.
Treatment Rejection (*RXR*)	Focuses on attributes and attitudes theoretically predictive of interest and motivation in making personal changes of a psychological or emotional nature.
Interpersonal Scales	
Dominance (*DOM*)	Assesses the extent to which a person is controlling and independent in personal relationships. A bipolar dimension with a dominant style at the high end and a submissive style at the low end.
Warmth (*WRM*)	Assesses the extent to which a person is interested in supportive and empathic personal relationships. A bipolar dimension with a warm, outgoing style at the high end and a cold, rejecting style at the low end.

Table 1-2
PAI Subscales and Their Descriptions

Subscale (designation)	Description
Somatic Complaints	
Conversion (*SOM-C*)	Focuses on symptoms associated with conversion disorder, particularly sensory or motor dysfunctions.
Somatization (*SOM-S*)	Focuses on the frequent occurrence of various common physical symptoms and vague complaints of ill health and fatigue.
Health Concerns (*SOM-H*)	Focuses on a preoccupation with health status and physical problems.
Anxiety	
Cognitive (*ANX-C*)	Focuses on ruminative worry and concern about current issues that result in impaired concentration and attention.
Affective (*ANX-A*)	Focuses on the experience of tension, difficulty in relaxing, and the presence of fatigue as a result of high perceived stress.
Physiological (*ANX-P*)	Focuses on overt physical signs of tension and stress, such as sweaty palms, trembling hands, complaints of irregular heartbeats, and shortness of breath.
Anxiety-Related Disorders	
Obsessive-Compulsive (*ARD-O*)	Focuses on intrusive thoughts or behaviors, rigidity, indecision, perfectionism, and affective constriction.
Phobias (*ARD-P*)	Focuses on common phobic fears, such as social situations, public transportation, heights, enclosed spaces, or other specific objects.
Traumatic Stress (*ARD-T*)	Focuses on the experience of traumatic events that cause continuing distress and that are experienced as having left the client changed or damaged in some fundamental way.
Depression	
Cognitive (*DEP-C*)	Focuses on thoughts of worthlessness, hopelessness, and personal failure, as well as indecisiveness and difficulties in concentration.
Affective (*DEP-A*)	Focuses on feeling of sadness, loss of interest in normal activities, and anhedonia.
Physiological (*DEP-P*)	Focuses on level of physical functioning, activity, and energy, including disturbance in sleep pattern and changes in appetite and/or weight loss.
Mania	
Activity Level (*MAN-A*)	Focuses on overinvolvement in a wide variety of activities in a somewhat disorganized manner and the experience of accelerated thought processes and behavior.
Grandiosity (*MAN-G*)	Focuses on inflated self-esteem, expansiveness, and the belief that one has special and unique skills or talents.

(continued)

Since it
Person
(PAI®) i
most ir
of clini
Morey
clinical
a volu
art in
Guide t
tory™
inatior

- Th
sca
sul
- Th
po
sc
- Th
so
P
at
- T
re
m
st
- T
n
o
- T
p
b
-
i
l
-

Thi
her
int
ass
pra
to
ins

Table 1-2 (continued)
PAI Subscales and Their Descriptions

Subscale (designation)	Description
Mania (continued)	
Irritability (*MAN-I*)	Focuses on the presence of strained relationships due to the respondent's frustration with the inability or unwillingness of others to keep up with their plans, demands, and possibly unrealistic ideas.
Paranoia	
Hypervigilance (*PAR-H*)	Focuses on suspiciousness and the tendency to monitor the environment for real or imagined slights by others.
Persecution (*PAR-P*)	Focuses on the belief that one has been treated inequitably and that there is a concerted effort among others to undermine one's interests.
Resentment (*PAR-R*)	Focuses on a bitterness and cynicism in interpersonal relationships, and a tendency to hold grudges and externalize blame for any misfortunes.
Schizophrenia	
Psychotic Experiences (*SCZ-P*)	Focuses on the experience of unusual perceptions and sensations, magical thinking, and/or other unusual ideas that may involve delusional beliefs.
Social Detachment (*SCZ-S*)	Focuses on social isolation, discomfort and awkwardness in social interactions.
Thought Disorder (*SCZ-T*)	Focuses on confusion, concentration problems, and disorganization of thought processes.
Borderline Features	
Affective Instability (*BOR-A*)	Focuses on emotional responsiveness, rapid mood changes, and poor emotional control.
Identity Problems (*BOR-I*)	Focuses on uncertainty about major life issues and feelings of emptiness, unfulfillment, and an absence of purpose.
Negative Relationships (*BOR-N*)	Focuses on a history of ambivalent, intense relationships in which one has felt exploited and betrayed.
Self-Harm (*BOR-S*)	Focuses on impulsivity in areas that have high potential for negative consequences.
Antisocial Features	
Antisocial Behaviors (*ANT-A*)	Focuses on a history of antisocial acts and involvement in illegal activities.
Egocentricity (*ANT-E*)	Focuses on a lack of empathy or remorse and a generally exploitive approach to interpersonal relationships.
Stimulus-Seeking (*ANT-S*)	Focuses on a craving for excitement and sensation, a low tolerance for boredom, and a tendency to be reckless and risk-taking.

(continued)

Table 1-2 (continued)
PAI Subscales and Their Descriptions

Subscale (designation)	Description
Aggression	
Aggressive Attitude (*AGG-A*)	Focuses on hostility, poor control over anger expression, and a belief in the instrumental utility of aggression.
Verbal Aggression (*AGG-V*)	Focuses on verbal expressions of anger ranging from assertiveness to abusiveness, and a readiness to express anger to others.
Physical Aggression (*AGG-P*)	Focuses on a tendency to physical displays of anger, including damage to property, physical fights, and threats of violence.

Normative Data

The PAI was developed and standardized for use in the clinical assessment of individuals in the age range of 18 through adulthood. The initial reading level analyses of the PAI test items indicated that reading ability at the fourth-grade level was necessary to complete the inventory. Subsequent studies of this issue (e.g., Schinka & Borum, 1993) have supported the conclusion that the PAI items are written at a grade equivalent lower than estimates for comparable instruments.

PAI scale and subscale raw scores are transformed to *T* scores in order to provide interpretation relative to a standardization sample of 1,000 community-dwelling adults. This sample was carefully selected to match 1995 U.S. census projections on the basis of gender, race, and age; the educational level of the standardization sample was selected to be representative given the required fourth-grade reading level. The only stipulation for inclusion in the standardization sample (other than stratification fit) was that the respondent had to endorse more than 90% of PAI items (i.e., no more than 33 items could be left blank). No other restrictions based upon PAI data were applied in creating the census-matched standardization sample.

The PAI *T* scores are calibrated to have a mean of 50 and a standard deviation of 10, using a standard linear transformation from the community sample norms. Thus, a *T*-score value greater than 50 lies above the mean in comparison to the scores of respondents in the standardization sample. Roughly 84% of nonclinical respondents will have a *T* score below 60 (i.e., 1 *SD* above the mean) on most scales, whereas 98% of nonclinical respondents will have scores below 70 (i.e., 2 *SD* above the mean). Thus, a *T* score at or above 70 represents a pronounced deviation from the typical responses of adults living in the community.

Since it
Person
(PAI®)
most ir
of clini
Morey
clinical
a volu
art in
Guide t
tory™
ination

- The sca sub
- The poi sca
- The sou PAI atic
- Th res ma sub
- Th ma of
- Th po bel
- Th int PA
- Th mc

This I
hensiv
interp
assess
practi
to ma
instru

For each scale and subscale, the *T* scores were linearly transformed from the means and standard deviations derived from the census-matched standardization sample. Unlike many other similar instruments, the PAI does not calculate *T* scores differently for men and women; instead, the same (combined) norms are used for both genders. This is because separate norms distort natural epidemiological differences between genders. For example, women are less likely than men to receive a diagnosis of antisocial personality, and this is reflected in lower mean scores for women on the Antisocial Features (*ANT*) scale. A separate normative procedure for men and women would result in similar numbers of each gender scoring in the clinically significant range, a result that does not reflect the established gender ratio for this disorder. The PAI development included several procedures designed to eliminate items that might be biased due to demographic features (e.g., race, gender, or age), and items that displayed any signs of being interpreted differently as a function of these features were eliminated in the course of selecting the final test items. As it turns out, with relatively few exceptions, differences as a function of demography were negligible in the community sample. Table 1-3 lists all PAI variables for which any of three demographic variables (i.e., race, gender, or age) accounted for more than 5% of the variance in the PAI score and the resulting effect (in terms of *T*-score units) of that variable.

Table 1-3
Summary of Significant Gender, Race, and Age Influences on PAI Scale Scores

PAI Scale	Demographic influences		Primary subscales affected
PAR	Non-White:	+ 6*T*	*PAR-H*
	18-29 years:	+ 5*T*	*PAR-P*
	60+ years:	– 4*T*	*PAR-R*
BOR	18-29 years:	+ 6*T*	*BOR-I*
	60+ years:	– 4*T*	*BOR-I*
ANT	Male:	+ 3*T*	*ANT-A*
	18-29 years:	+ 7*T*	*ANT-S*
	60+ years:	– 4*T*	*ANT-A*
AGG	18-29 years:	+ 5*T*	*AGG-V*
	60+ years:	– 4*T*	*AGG-P*
STR	18-29 years:	+ 4*T*	(no subscales)
	60+ years:	– 4*T*	

T scores are derived from a representative community sample; therefore, they provide a useful means for determining whether certain problems are clinically significant, because relatively few normal adults will obtain markedly elevated scores. However, other comparisons are often of equal importance in clinical decision-making. For example, nearly all patients report depression at their initial evaluation; the question confronting the clinician considering a diagnosis of major depression is one of *relative* severity of symptomatology. Knowing that an individual's score on the PAI Depression scale is elevated in comparison to the standardization sample is of value, but a comparison of the elevation relative to a clinical sample may be more critical in forming diagnostic hypotheses.

To facilitate these comparisons, the PAI profile form (shown in Figure 1-1) also indicates the *T* scores that correspond to marked elevations when referenced against a representative *clinical* sample. The profile "skyline" indicates the score for each scale and subscale that represents the raw score that is 2 standard deviations above the mean for a clinical sample of 1,246 patients selected from a wide variety of different professional settings. Thus, roughly 98% of clinical patients will obtain scores below the skyline on the profile form. Therefore, scores above this

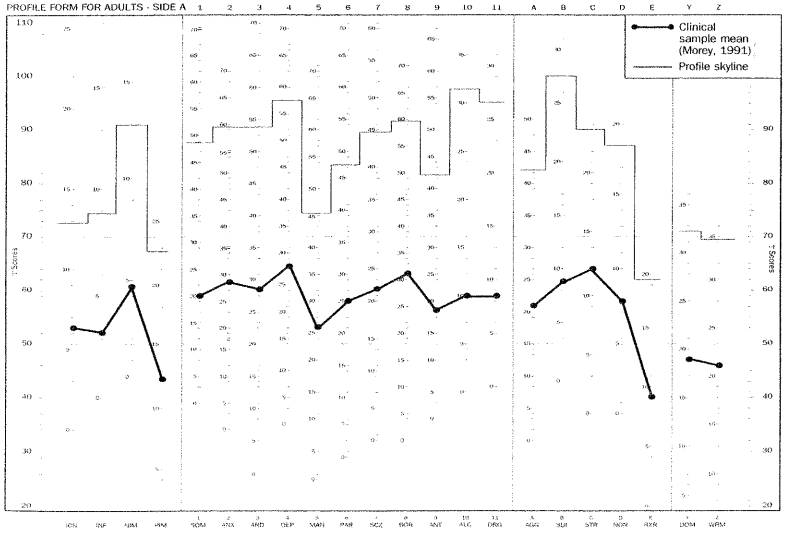

Figure 1-1. Mean PAI *T* scores for a clinical sample of adults (N = 1,246) and the skyline at 2 *SD* above the mean in that clinical sample.

skyline represent a marked elevation of scores relative to those of *patients in clinical settings*. Thus, interpretation of PAI profiles can be accomplished in comparison to both normal and clinical samples.

The *PAI Professional Manual* (Morey, 1991) provides normative transformations for a number of different comparisons. The appendices provide *T*-score transformations referenced against the clinical sample and a large sample of college students (N = 1,051), as well as for various demographic subgroups of the community standardization sample. Although the differences between different demographic groups were generally quite small, there are occasions where it may be useful to make comparisons with reference to particular groups. The raw score means and standard deviations needed to convert raw scores to *T* scores with reference to normative data provided by particular groups (men, women, Blacks, and respondents over age 60) are provided in the manual for this purpose. However, for most clinical and research applications, the use of the *T* scores derived from the full normative data is strongly recommended, because this sample was both large and representative of the general population.

Reliability of the PAI

The reliability of the PAI has been examined in a number of different studies that have examined the internal consistency, test-retest reliability, and configural stability of the instrument.

The internal consistency of the PAI has been examined in a number of different populations (Alterman et al., 1995; Boyle & Lennon, 1994; Morey, 1991; Rogers, Flores, Ustad, & Sewell, 1995; Schinka, 1995). This has involved the use of coefficient alpha (Cronbach, 1951), which can be interpreted as an estimate of the mean of all possible split-half combinations of items. The internal consistency alphas for the PAI full scales are satisfactory; in the *PAI Professional Manual*, Morey reports median alphas for the full scales of .81, .82, and .86 for normative, college, and clinical samples, respectively. As expected, the scales tend to appear more internally consistent in more heterogeneous samples. Alterman et al. found a median alpha of .78 in a sample of methadone maintenance patients; Schinka found a median alpha of .86 for full scales and .77 for the subscales in an alcoholic sample. Boyle and Lennon (1994) reported a median alpha of .84 in a mixed clinical–normal sample. Internal consistency estimates for the *ICN* and *INF* scales are consistently lower than those for other scales, because these scales do not measure theoretical constructs; instead, they measure the care with which the respondent completed the test. Lower alphas for such scales would be anticipated, as

carelessness might vary within a given sitting (e.g., a respondent might complete the first half of the test accurately, but complete the last half haphazardly).

The lowest internal consistency estimates for the PAI reported in the literature were obtained using the Spanish version of the instrument (Rogers et al., 1995), where an average alpha of .63 was obtained. Rogers and colleagues concluded that the internal consistency of the treatment consideration scales seemed to be most affected by the translation of the test. Examination of internal consistency estimates for the PAI full scales for groups defined by various demographic characteristics (Morey, 1991) does suggest that there is little variability in internal consistency (i.e., median scale alphas) as a function of race (i.e., Whites = .77, non-Whites = .78), gender (i.e., men = .79, women = .75), or age (i.e., under 40 years = .79, 40 years and over = .75).

The temporal stability of PAI scales has been examined by administering the test to respondents on two different occasions (Boyle & Lennon, 1994; Morey, 1991; Rogers et al., 1995). For the standardization studies, median test-retest reliability over a 4-week interval for the 11 full clinical scales was .86 (Morey, 1991), leading to standard error of measurement estimates for these scales on the order of 3 to 4 *T*-score points, with 95% confidence intervals of ±6 to 8 *T*-score points. Examination of the mean absolute *T*-score change values for scales also revealed that the absolute changes over time were quite small, on the order of 2 to 3 *T*-score points for most of the full scales (Morey, 1991). Boyle and Lennon (1994) reported a median test-retest reliability of .73 in their normal sample over 28 days. Rogers et al. (1995) found an average stability of .71 for the Spanish version of the PAI, administered over a 2-week interval.

Because multiple-scale inventories are often interpreted configurally, additional questions concerning the stability of configurations on the 11 PAI clinical scales are necessary. One such analysis (Morey, 1991) examined the inverse (or Q-type) correlation between each respondent's test and retest profiles. Correlations were obtained for each of the 155 respondents in the full retest sample, and a distribution of these within-subject profile correlations was obtained. Conducted in this manner, the median correlation over time of the clinical scale configuration was .83, indicating a substantial degree of stability in profile configurations over time.

Validity of the PAI

The validation of measures of clinical constructs is a process that requires accumulation of data concerning convergent and discriminant validity correlates. In Morey's (1991) examination of PAI validity, a number of the best available clinical

indicators were administered concurrently to various samples to determine their convergence with corresponding PAI scales. Furthermore, diagnostic and other clinical judgments concerning clinical behaviors (as rated by the treating clinician) were also examined to determine whether their PAI correlates were consistent with hypothesized relationships. Finally, a number of simulation studies were performed to determine the efficacy of the PAI validity scales in identifying response sets. To date, a number of studies have been conducted examining correlates of various PAI scales; in the *PAI Professional Manual*, Morey provides information about correlations of individual scales with over 50 concurrent indices of psychopathology. Noteworthy findings from these studies are described in the following paragraphs.

The PAI validity scales were developed to provide an assessment of the potential influence of certain response tendencies on PAI test performance. Two of these scales, Inconsistency (*ICN*) and Infrequency (*INF*), were developed to assess deviations from conscientious responding, whereas the other two validity scales, Negative Impression (*NIM*) and Positive Impression (*PIM*), were developed to provide an assessment of efforts at impression management by the respondent.

To model the performance of respondents completing the PAI in a random fashion, computer-generated profiles were created by generating random responses to individual PAI items and then scoring all scales according to their normal scoring algorithms. A total of 1,000 simulated protocols were generated for this analysis. Comparison of profiles derived from normal respondents, clinical respondents, and the random response simulations demonstrated a clear separation between scores of actual respondents and scores from the random simulations; 99.4% of these random profiles were identified as such by either *ICN* or *INF* (Morey, 1991).

To model the performance of respondents attempting to manage their impressions in either a positive or negative direction, studies have been performed (Morey, 1991) in which respondents were instructed to simulate such response styles. Comparisons of profiles for normal respondents, clinical respondents, and the corresponding response style simulation groups demonstrated a clear separation between scores for the actual respondents and scores for the simulated response groups. Respondents scoring above the critical level of *NIM* were 14.7 times more likely to be a member of the malingering group than of the clinical sample, whereas respondents scoring above threshold on *PIM* were 13.9 times more likely to be from the positive dissimulation sample than from a community sample. Subsequent studies generally support the ability of these scales to distinguish simulators from actual protocols under a variety of response set conditions

(e.g., Cashel, Rogers, Sewell, & Martin-Cannici, 1995; Rogers, Ornduff, & Sewell, 1993). Results of these studies are reviewed in greater detail in chapters 4 and 5.

In addition to such simulation studies, a number of correlational studies have been performed to determine the convergent and discriminant validity of the PAI validity scales as measured against other commonly used measures of similar constructs (Ban, Fjetland, Kutcher, & Morey, 1993; Costa & McCrae, 1992; Morey, 1991). For example, *NIM* correlated significantly (r = .54) with the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1967) *F* scale; *PIM* was associated with the Marlowe-Crowne (Crowne & Marlowe, 1957) Social Desirability scale (r = .56) as well as with the MMPI *K* (r = .47) and *L* (r = .41) scales (Morey, 1991). PAI scales *INF* and *ICN* displayed negligible correlations with any measures, an expected result as these scales were designed as relatively pure indicators of measurement error.

The clinical scales of the PAI were assembled to provide information about critical diagnostic features of 11 important clinical constructs. A number of different validity indicators have been used to provide information on the convergent and discriminant validity of the PAI clinical scales; these indicators can be divided into measures of "neurotic features," "psychotic features," and "behavior disorder features." Within the neurotic spectrum, correlations with the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985), the MMPI clinical and research scales (Hathaway & McKinley, 1967; Morey, Waugh, & Blashfield, 1985; Wiggins, 1966), and several specialized assessment instruments have been examined. These specialized instruments include the following: the Wahler Physical Symptoms Inventory (Wahler Inventory; Wahler, 1983), a broad measure of somatic complaints; the Beck Depression Inventory (BDI; Beck & Steer, 1987), the Beck Anxiety Inventory (BAI; Beck & Steer, 1990) and the Beck Hopelessness Scale (BHS; Beck & Steer, 1988), three widely used and well-validated measures of negative affect; the Hamilton Rating Scale for Depression (HAM-D: Hamilton, 1960), perhaps the most widely used measure of outcome in treatment studies of depression; the State-Trait Anxiety Inventory (STAI; Spielberger, 1983), a widely used measure that distinguishes between the situational and more enduring elements of anxiety; the Fear Survey Schedule (FSS; Wolpe & Lang, 1964), a comprehensive assessment of common fears; the Maudsley Obsessive-Compulsive Inventory (Maudsley Inventory; Rachman & Hodgson, 1980), a measure of severe obsessional ideation and contamination fears; and the Mississippi Scale for Combat-Related Posttraumatic Stress Disorder (Mississippi PTSD; Keane, Caddell, & Taylor, 1988).

Correlations between each of the full scale scores for the four PAI neurotic cluster scales and the validation measures described above follow hypothesized

patterns, demonstrating strong associations with other measures of neuroticism (Costa & McCrae, 1992; Montag & Levin, 1994; Morey, 1991). The strongest correlates for Somatic Complaints (*SOM*) were found with the Wiggins Health Concerns (r = .80) and Organic Problems (r = .82) content scales, the Wahler Inventory (r = .72), and the MMPI Hypochondriasis (r = .60) scale. Each of these measures is a fairly straightforward assessment of complaints regarding physical functioning, so this pattern of correlations is consistent with expectations. The *SOM* scale also displays small-to-moderate relationships with measures of distress, such as anxiety or depression. The *SOM* scale is generally the highest point of the PAI profile in a general medical population, although, even in such populations, the average score is typically below 70*T* (Osborne, 1994).

The Anxiety (*ANX*) scale demonstrated substantial correlations with a number of measures of negative affect, including the NEO-PI Neuroticism (r = .76) and Anxiety (r = .76), the STAI Trait Anxiety Inventory (r = .73), and the Wiggins Depression content (r = .76) scales. This finding is consistent with research results highlighting the prominent role of anxiety in many mental disorders; such a pattern should be anticipated, as *ANX* was intended to be a general measure of anxiety rather than a specific diagnostic indicator. In contrast, the Anxiety-Related Disorders (*ARD*) scale was designed to provide content relevant to more specific diagnostic differentiations; hence, the pattern of correlations tends to be more specific than that observed with *ANX*. The largest correlation for *ARD* was with the Mississippi PTSD scale (r = .81), and the second largest involved the FSS (r = .66); each of these scales directly parallels a disorder for which *ARD* was designed to provide coverage. The *ARD* scale has also been found to correlate with the probability of getting nightmares (r = .46), with *ARD-T* (r = .51), in particular, being associated with night terrors (Greenstein, 1993). The *ARD* scale (particularly *ARD-T*) also has been found to differentiate between women psychiatric patients who were victims of childhood abuse and women patients who did not experience such abuse (Cherepon & Prinzhorn, 1994).

The Depression (*DEP*) scale demonstrates its highest correlations with various well validated indicators of depression, such as the BDI (r = .81), the HAM-D (r = .78), and the Wiggins Depression content scale (r = .81). This is consistent with expectations, because these measures are widely used in the assessment of depression and related symptomatology. Other noteworthy correlates of the Depression scale include the MMPI *D* scale (r = .66), the Wiggins Poor Morale scale (r = .74), the NEO-PI Neuroticism (r = .69) and Depression (r = .70) scales, and the Beck Hopelessness scale (r = .67).

In addition, correlations with a number of other measures of related constructs can provide information relevant to the convergent and discriminant validity of the PAI "psychotic cluster" scales. For example, the MMPI, the NEO-PI, the Interpersonal Adjective Scale (IAS-R; Trapnell & Wiggins, 1990) and the clinician-rated Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) include scales that capture the cognitive and interpersonal abnormalities that characterize these disorders. Correlations between each the three PAI psychotic spectrum scales and these validation measures generally follow the expected pattern (Ban et al., 1993; Costa & McCrae, 1992; Morey, 1991). The Mania (*MAN*) scale has demonstrated its strongest correlations with Wiggins Hypomania (r = .63), Psychoticism (r = .58), and Hostility (r = .55) content scales; with the BPRS clinical ratings of Grandiosity (r = .48) and Conceptual Disorganization (r = .40); and with the MMPI *Ma* scale (r = .53). The Paranoia (*PAR*) scale demonstrated its largest correlations with the MMPI Paranoid personality disorder scale (r = .70), the Wiggins Psychoticism scale (r = .60), and various measures of hostility such as the Wiggins Hostility content scale (r = .54) and the NEO-PI Hostility facet scale (r = .55). A moderate correlation with the MMPI *Pa* scale was also observed (r = .45). The Schizophrenia (*SCZ*) scale has been found to correlate with the Wiggins Psychoticism content scale (r = .76) and the MMPI Schizotypal (r = .67) and Paranoid (r = .66) personality disorder scales. The *SCZ* scale was also positively correlated with the MMPI *Sc* scale (r = .55) and negatively associated with indices of sociability and social effectiveness such as the NEO-PI Agreeableness (r = –.49) and Gregariousness (r = –.57) scales. This pattern indicates that scores on the *SCZ* scale reflect disruptions in both the cognitive (e.g., delusions, hallucinations) and the interpersonal (e.g., limited social competence) realms of functioning. Finally, the *SCZ* scale has been found to distinguish schizophrenic patients from controls (Boyle & Lennon, 1994). In that study the schizophrenic sample did not differ significantly from a sample of alcoholics on *SCZ* scores, although the article suggested that many of the alcoholic patients completed the PAI during detoxification, which might complicate differential diagnosis based solely upon *SCZ* scores.

Information on the convergent and discriminant validity of the PAI scales in the behavior disorders cluster is also available. In addition to the NEO-PI, the IAS-R, and the MMPI, the PAI scales have been correlated with a number of specialized assessment instruments, including the Bell Object Relations Inventory (Bell Inventory; Bell, Billington, & Becker, 1985), a multifactorial questionnaire constructed to measure a variety of interpersonal attitudes and beliefs indicative of early pathological object relations thought to be at the core of the borderline syndrome

(Bell, Billington, Cicchetti, & Gibbons, 1988); the Michigan Alcoholism Screening Test (MAST; Selzer, 1971), a widely used and well validated measure of problem behaviors associated with drinking; the Drug Abuse Screening Test (DAST; Skinner, 1982), a measure, patterned after the MAST, that assesses the consequences of drug abuse; and the Self-Report Psychopathy test designed by Hare (1985) to assess his model of psychopathy.

Correlations between scores for the four PAI behavior disorder cluster scales and these validation measures follow expected patterns (Costa & McCrae, 1992; Kurtz, Morey, & Tomarken, 1993; Morey, 1991). The strongest correlates of the Borderline Features (*BOR*) scale are the MMPI Borderline personality disorder scale (r = .77), the NEO-PI Neuroticism scale (r = .67), and several different measures of hostility, such as the NEO-PI Hostility facet (r = .70). The *BOR* scale also displayed substantial correlations with the Bell Inventory Insecure Attachment scale (r = .63), the NEO-PI Impulsiveness facet (r = .52), and the Wiggins Family Problems (r = .63) and Psychoticism (r = .63) content scales. This pattern of anger, impulsiveness, and interpersonal clashes is consistent with the core features of the borderline syndrome. Other studies have supported the validity and utility of this scale in a variety of clinical contexts. The *BOR* scale in isolation has been found to distinguish borderline patients from unscreened controls with an 80% hit rate; it successfully identified 91% of these respondents as part of a discriminant function (Bell-Pringle, 1994). Classifications based on the *BOR* scale have been validated in a variety of domains related to borderline functioning, including depression, personality traits, coping, Axis I disorders, and interpersonal problems (Trull, 1995). These *BOR* scale classifications were also found to be predictive of 2-year outcome on academic indices in college students, even controlling for academic potential and diagnoses of substance abuse (Trull, Useda, Conforti, & Doan, 1995).

The PAI Antisocial Features (*ANT*) scale demonstrated its largest correlations with the Hare Psychopathy Scale (r = .82) and the MMPI Antisocial personality disorder scale (r = .77). Other correlates included the Wiggins Hostility (r = .57) and Family Problems (r = .52) content scales, the NEO-PI Excitement Seeking facet (r = .56), and the IAS-R cold interpersonal octant (r = .45). This pattern suggests that the *ANT* scale addresses the personality, interpersonal, and behavioral elements of psychopathy. The correlation with the MMPI *Pd* scale is positive, but not impressive (r = .34), suggesting that the two scales represent the core features of the disorder somewhat differently. The PAI Alcohol Problems (*ALC*) and Drug Problems (*DRG*) scales each demonstrate a similar pattern of correlates: strong correlations with corresponding measures of substance abuse and moderate associations with indicators of antisocial personality. *ALC* yields a correlation of .89 with

the MAST, whereas *DRG* correlates .69 with the DAST. The *ALC* scale has been found to differentiate patients in an alcohol rehabilitation clinic from both patients with schizophrenia and normal controls (Boyle & Lennon, 1994). The *DRG* scale has also been found to successfully discriminate drug abusers and methadone maintenance patients from general clinical and community samples (Alterman et al., 1995).

The treatment consideration scales of the PAI were assembled to provide indicators of potential complications in treatment that would not necessarily be apparent from diagnostic information. There are five of these scales: two indicators of potential for harm to self or others, two measures of the respondent's environmental circumstances, and one indicator of the respondent's motivation for treatment. These scales have been compared to a number of measures of related constructs. In addition to the NEO-PI, the IAS-R, and the MMPI, the scales have been correlated with a number of specialized assessment instruments. The BDI, BAI, and BHS provide convergent correlates for suicidal ideation. Also, the Suicide Probability Scale (SPS; Cull & Gill, 1982) serves as a concurrent indicator of suicide potential. The SPS has four subscales that assess hopelessness, suicidal ideation, negative self-evaluation, and hostility, in addition to yielding a total score for suicide probability. The State-Trait Anger Expression Inventory (STAXI; Spielberger, 1988) provides a marker for aggression that is broken down into six scales and two subscales. The Perceived Social Support scales (Procidano & Heller, 1983) provide an assessment of the subjective impact of supportive transactions between the respondent and his or her social support system; two separate scales assess support provided by the respondent's family and the respondent's friends. Finally, the Schedule of Recent Events (SRE; Holmes & Rahe, 1967) is a unit-scoring adaptation of the widely used Holmes and Rahe (1967) checklist of recent stressors, where respondents are asked to indicate major life changes that have taken place during the 12 months prior to evaluation.

Correlations between the PAI treatment consideration scales and such validation measures provide support for the construct validity of these PAI scales (Costa & McCrae, 1992; Morey, 1991). Substantial correlations have been identified between the Aggression (*AGG*) scale and the NEO-PI Hostility (r = .83) and STAXI Trait Anger (r = .75) scales. The *AGG* scale also was negatively correlated with the STAXI Anger Control scale (r = –.57). The Suicidal Ideation (*SUI*) scale was most positively correlated with the BHS (r = .64), the BDI (r = .61), the Suicidal Ideation (r = .56) and Total Score (r = .40) of the SPS; it also was found to be negatively correlated with the measures of perceived social support. As expected, the Nonsupport (*NON*) scale was found to be highly (and inversely) correlated with the social support measures: –.67 with PSS-Family and –.63 with PSS-Friends. *NON*

also was moderately associated with numerous measures of distress and tension. The Stress (*STR*) scale displayed its largest correlations with the SRE (r = .50) and also was associated with various indices of depression and poor morale. Finally, the Treatment Rejection (*RXR*) scale was found to be negatively associated with Wiggins Poor Morale (r = –.78) and the NEO-PI Vulnerability (r = –.54) scales, consistent with the idea that distress can serve as a motivator for treatment. The Treatment Rejection scale has been shown to be positively associated with indices of social support (r = .26 to .49), suggesting that people are less likely to be motivated for treatment if they have an intact and available support system as an alternative.

The interpersonal scales of the PAI were designed to provide an assessment of the interpersonal style of respondents along two dimensions: (a) a warmly affiliative versus a cold rejecting axis, and (b) a dominating and controlling versus a meekly submissive style. These axes provide a useful way of conceptualizing variation in normal personality as well as in many different mental disorders, and persons at the extremes of these dimensions may present with a variety of disorders. The *PAI Professional Manual* (Morey, 1991) describes a number of studies indicating that diagnostic groups differ on these dimensions; for example, spouse-abusers are relatively high on the Dominance (*DOM*) scale, whereas schizophrenics are low on the Warmth (*WRM*) scale. Correlations with related measures also provide support for the construct validity of these scales. For example, the correlations with the IAS-R vector scores are consistent with expectations, with PAI *DOM* associated with the IAS-R Dominance vector (r = .61) and PAI *WRM* associated with the IAS-R Love vector (r = .65). The NEO-PI Extroversion scale roughly bisects the high *DOM*/high *WRM* quadrant, as it is moderately positively correlated with both scales; this finding is consistent with previous research (Trapnell & Wiggins, 1990). The *WRM* scale was also correlated with the NEO-PI Gregariousness scale (r = .46), whereas *DOM* was associated with the NEO Assertiveness facet (r = .71).

In summary, the PAI scales have been found to associate in theoretically concordant ways with most major instruments for the assessment of diagnosis and treatment efficacy. Strategies for the interpretation of the PAI profile and its use in treatment planning and evaluation are presented in following sections.

Basic Interpretive Strategy

Because the development of the PAI emphasized the importance of both convergent and discriminant validity of the instrument, the interpretation of PAI protocols is relatively straightforward. For example, scales were designed to be generally

pure measures of the specific constructs; thus, an elevation on the *DEP* scale may be interpreted as indicating that the respondent reports a number of experiences consistent with the symptomatology of clinical depression. Interpretive hypotheses may be generated at four different levels: the *item* level, the *subscale* level, the *full scale* level, and the *configuration* level.

Interpretation of PAI responses at the item level are meaningful because the content of each item was assumed to be critical in determining its relevance for the assessment of the construct. For example, each item was reviewed by a panel of experts to ensure that its content was directly relevant to the specific clinical construct. As a result, a review of item content can provide specific information about the nature of the difficulties experienced by the respondent. In addition, 27 PAI items were identified as "critical items" based on two criteria: (a) importance of their content as an indicator of potential crisis situations, and (b) very low endorsement rates in normal individuals. Endorsement of any of these items should be followed by more detailed questioning that can clarify the nature and severity of these concerns.

The PAI subscales were constructed as an aid in isolating the core elements of the different clinical constructs measured by the instrument. These subscales can serve to clarify the meaning of full scale elevations, and may be used configurally in diagnostic decision-making. For example, many patients typically come to clinical settings with marked distress and dysphoria; this often leads to elevations on most unidimensional depression scales. However, unless other manifestations of the syndrome are present, this does not necessarily indicate that Major Depressive Disorder is the likely diagnosis. In the absence of features such as vegetative signs, lowered self-esteem, and negative expectancies, the diagnosis may not be warranted even with a prominent elevation on a unidimensional depression scale. On the PAI, such a pattern would lead to an elevation on *DEP-A*, representing the dysphoria and distress, but no elevations on *DEP-P* (the vegetative signs) and *DEP-C* (the cognitive signs). As a result, an overall elevation on *DEP* in this instance would not be interpreted as diagnostic of major depression because of the lack of supporting data from the subscale configuration.

Interpretation of PAI full scale scores is aided by comparison to two referents: expected scores in the community and expected scores in clinical patients. As described earlier, the PAI profile form (Figure 1-1) provides a skyline marking an elevation of 2 standard deviations with respect to the clinical sample. The similarity of expected scores for these two populations varies a great deal across scales. For example, the interpersonal scales *DOM* and *WRM* have distributions that are

quite similar in both community and clinical samples; thus, marked elevations (or very low scores) are noteworthy regardless of the nature of the client. On the other hand, the *RXR* scale (which was designed to identify risk for early treatment termination) has a markedly different distribution in clinical and community samples. A majority of clinical respondents who are currently in treatment obtain scores that are considerably lower than those of community respondents, who are typically not in psychological treatment and have little interest in it. Thus, a *T* score of 50 on *RXR* in a client presenting for psychotherapy, although "average" for a community sample, is actually considerably above the expected score for respondents in clinical settings. In this instance, the *RXR* score should be interpreted as indicating potentially significant resistance to change for this client. In contrast, an *RXR* score of 50*T* in an individual who was administered the PAI for personnel selection purposes would be unremarkable. In these two examples, the differences in the assessment question leads to differences in the interpretation of the information yielded by a normative transformation.

The broadest level of PAI interpretation involves the analysis of scale configuration. Traditionally, the premise behind multidimensional inventories such as the PAI has been that the *combination* of information provided by the multiple scales is greater than any of its parts; hence, most previous research focused on the profile yielded by such an inventory, rather than any single scale elevations. There are a variety of ways to examine profile configuration; to date, there have been five research approaches to studying the configural use of PAI profile data. These approaches include the use of mean profiles, profile codetypes, cluster profiles, actuarial functions, and conceptually driven configural decision rules. These differing approaches can be applied to different issues in decision-making, including diagnostic (e.g., Is this a schizophrenic patient or a depressed patient?), intervention (e.g., Does this patient require inpatient treatment?), or protocol-related (e.g., Is this a valid PAI protocol?) issues. Each of these approaches will be discussed throughout this guide in the context of these different types of decisions.

The following chapters will focus on the four different interpretive levels in an effort to resolve certain dilemmas the test user may face in interpreting the PAI. The initial focus is on understanding the composition and interpretation of the individual scales; this is followed by a discussion of the meaning of different two-point combinations of scales (codetypes). The remainder of this interpretive guide explores specific issues commonly encountered in PAI interpretation: Is this patient malingering or defensive? What diagnoses should be considered? What is the person's characteristic view of self and of others? What initial steps should be considered in planning treatment? In all cases, the available data are used to address these questions, but, as is the case with any assessment instrument, many

questions require further study. It is hoped that current and future PAI users can help to fill the gaps in this literature, so that subsequent editions of this guide can incorporate the advances made possible by such work.