

Assessment of Perinatal Substance Abusers: Experiences of One Perinatal-20 Project

Deborah L. Haller and Karen S. Ingersoll

INTRODUCTION

This chapter identifies and discusses issues relevant to the psychosocial evaluation of perinatal substance abusers. The conclusions are based on 5 years of experience with 133 women treated at Virginia Commonwealth University's National Institute on Drug Abuse (NIDA)-funded Treatment Research Unit, the Center for Perinatal Addiction (CPA). The chapter begins with a brief overview of measurement methodology relevant to clinical research and then describes the initial rationale for test selection and presents some measurement problems that have been encountered. After delineating weaknesses in the test protocol, the authors recommend adjustments and suggest areas of exploration for future clinical research projects.

GENERAL ISSUES OF MEASUREMENT

The first step in selecting any assessment instrument is to clarify the purpose of testing (Anastasi 1982; Cronbach 1984). In a research project, the purpose will emanate from the research questions. For example, an investigation of the base rates for mental illness in a given population might necessitate use of an instrument that makes categorical psychiatric diagnoses (e.g., the Structured Clinical Interview for *DSM-III-R* [SCID] [Spitzer et al. 1988]) using the *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised) (DSM-III-R)* (American Psychiatric Association 1987). If a researcher wishes to address change in functioning as a consequence of treatment, a different kind of measurement tool is indicated, one that yields scalable data (e.g., the Minnesota Multiphasic Personality Inventory-2 [MMPI-2] [Hathaway and McKinley 1989] and the Millon Clinical Multiaxial Inventory-II [MCMI-II] [Millon 1990]). Thus, for instrument selection to be most effective, research objectives must be specific. The investigator also must consider the psychometric properties of available tests. The two most important properties are reliability (consistency of the test) and validity (usefulness of the test). Anastasi (1982) gives a thorough explication of these issues; a brief review follows.

Reliability

Three types of reliability are relevant: internal consistency, temporal stability, and interrater agreement. If a test purports to measure a single domain, its internal consistency is important. This concept is less relevant to longer measures with multiple scales; however, good internal consistency within scales is expected. Tests also should possess moderate-to-high test-retest reliability, when readministered within a short timeframe, if their aim is to measure an abiding trait. In contrast, test-retest reliability is irrelevant when a test is used to measure a state rather than a trait. If observational or rater-scored measures are used, interrater reliability becomes important. For example, scoring of the SCID relies on the judgment of trained raters who complete a lengthy interview with the subject. If interrater reliability is poor, the investigator would have little confidence in the diagnoses generated because they would be more reflective of the raters than of the subject.

Validity

Once a test is found to possess adequate reliability, validity must be considered. Validity is the degree to which a test measures what it purports to measure. Four types of validity are relevant: construct-related, criterion-related, content, and face validity. *Construct-related* validity is the degree to which a test measures a hypothesized construct, such as “addiction.” Construct-related validity is inferred from examining the relationships among tests purporting to measure the same construct (providing convergent validity) and those purporting to measure different constructs (providing discriminant validity). Stated simply, if a given measure of addiction correlates well with established tests of addiction and does not correlate well with tests of less related constructs (e.g., intelligence), it probably possesses adequate construct validity. *Criterion-related* validity generally refers to a test’s predictive ability. Tests found to have good predictive validity may be useful as screening tools to match patients to treatments. *Content* validity is usually established by a panel of expert raters who “judge” that a given test adequately measures a domain. For example, to be useful, an achievement test should accurately reflect an individual’s performance. *Face* validity is the test’s acceptability to the user and the subject as a relevant-appearing instrument. Face validity is clinically important because research subjects quickly tire of tests that seem irrelevant to their lives. Subjects are more likely to enact a negative response set (a test-taking stance or attitude that influences responding, such as exaggerated responding in an attempt to convince others of severity of symptoms) or refuse to continue testing if the instrument lacks face validity. On the other hand, face-valid tests can lead to defensive or exaggerated responding by subjects who wish to create a

particular impression. In some instances, a non-face-valid (i.e., empirically derived) instrument, such as the MMPI-2, may be of greater utility for these reasons.

Standardization

Identifying appropriately normed, well-standardized tests that are suitable for studying perinatal substance abusers is difficult. Few appropriately normed instruments are available to study women. There are even fewer such tests to target women who are pregnant, belong to an ethnic minority, or are addicted to cocaine. Most standardized tests were developed using older males as subjects. Early studies often failed to include women in the standardization group or to provide descriptive statistics for the standardization sample by sex or race.

In general, researchers should build on previous research findings rather than design new instruments; however, this is not always possible in the perinatal substance abuse field. Accordingly, some investigators have developed new tests that tap domains of particular relevance to this population, whereas others have modified existing instruments to meet their needs. Findings for pregnant substance abusers from instruments established with men (such as the Addiction Severity Index [ASI] [McLellan et al. 1980]) also are being collected and published; this means that population-specific norms should be available in the future. It is crucial that limitations in measurement of perinatal substance abusers be discussed in the literature and in public forums so that investigators can make informed decisions regarding psychosocial assessment of this population. Along these lines, Haller and colleagues (1993a) recently noted possible sex-related differences for perinatal substance abusers on the MMPI-2 and MCMI-II, two widely used objective personality inventories. However, until large samples are collected from multiple sites, interpretations of many tests must remain tentative.

Test Administration Issues

Other issues of importance in selecting measurement instruments include test version, ease of administration, cost, and ease of scoring and interpretation. Some tests, such as the MMPI (Hathaway and McKinley 1967), have large bodies of established norms and interpretive data available for older test versions. An investigator may be more comfortable with interpretation of an older version of a test, already have a large data set using the older version, or own a computer-scoring program for an older version, making continued use of this version highly attractive. In comparison, another researcher may be unencumbered by such restraints and may therefore elect to use the newest version of the test. When the

Perinatal-20 projects were initiated, it was unclear whether recent test versions for several prominent instruments (including the MMPI and MCMI-II) would assess the same domains or be interpretable using the same assumptions as the older versions. As a result, investigators generally selected test versions for idiosyncratic reasons. It is reassuring that recent studies have found the consistency of MMPI and MMPI-2 profile shapes, elevation, and scatter to be good (Harrell et al. 1992). In addition, most clinicians have rated their interpretations of MMPI and MMPI-2 to be similar (Clavelle 1992), although the issue of codetype¹ equivalence remains unresolved.

Types of Assessment Instruments

Structured interviews, self-report measures, and behavior ratings are the most common types of measurement. All have costs and benefits to be considered, assuming relative equivalence of relevance to the purpose of testing and psychometric standing. *Structured clinical interviews* are among the most costly of techniques, requiring several trained raters who first must achieve solid interrater agreement. Typically, they will spend several hours interviewing a subject and coding the results. Test interpretation relies on the accuracy of coding. Structured interviews often yield truncated, limited information; however, as the result of strict standardization, they generally provide high-quality categorical information. Structured clinical interviews used in the CPA project included the SCID, ASI, and Structured Interview for DSM-III-R Personality-Revised (SIDP-R) (Pfohl et al. 1989).

In contrast, *self-report scales* are low cost, readily available, generally well standardized, and easily administered. Many can be interpreted by a single psychometrician, resulting in lower personnel costs. Computer-scoring and -interpreting programs also are readily available for many of these instruments. However, self-report scales can be vulnerable to subject biases and may not be appropriate for use with subjects with limited verbal skills. A crucial but sometimes neglected issue is the likelihood of subject acceptance or resistance. For example, many subjects balk at completing the MMPI-2 because of its length.

Behavior ratings include observational measures administered by trained raters; informant reports from raters, therapists, and peers; and performance tests in which the subject performs a sample behavior. *Observational measures* are costly in terms of personnel and the need

¹In the MMPI-2, the codetype is determined on the basis of high points (i.e., the scales that are most elevated). These scales define the codetype. For example, someone with the two highest scale elevations on scales 2 and 4 has a 2-4 codetype, with the highest scale named first.

for constant supervision to ensure that standardized procedures are consistently followed. Although they may yield information with better external validity than self-report measures, observational measures can be difficult to interpret. *Informant reports*, although inexpensive, are vulnerable to response sets, inconsistent reporting, and extreme tendencies. *Performance tests* may be more or less costly depending on the complexity of the behavior to be assessed; they are also vulnerable to response sets, coding errors, and lack of standardization.

Summary

Researchers should consider the relative costs and benefits associated with available instruments before making test selections. The first consideration should be the instrument's ability to address the research questions. Next, psychometric properties, specifically reliability and validity, must be considered. Although test standardization that includes the target population is important, norms for perinatal substance abusers are generally unavailable. Similarly, norms for women (particularly ethnic women) are rare. Investigators thus are left with the options of developing new instruments or using existing ones. The advantage of the latter approach is the opportunity for norms to be established for the particular population of interest and for direct comparison of the study population with related populations, such as male substance abusers. The final decision in the test selection process also should take into consideration test version, ease of administration, cost, and ease of scoring and interpretation.

The advantages of investigators working in a given field using the same instruments and the same test versions are obvious. When each project makes independent decisions regarding test protocol, this inadvertently creates a barrier to multisite comparisons and the development of new norms for substance-abusing women based on a large available sample.

DEVELOPMENT OF THE INITIAL CENTER FOR PERINATAL ADDICTION TESTING PROTOCOL

When the Perinatal-20 research projects were initiated, little was known about the target population of pregnant and recently postpartum substance abusers. Accordingly, CPA researchers decided to take a "shotgun" approach to assessment. This strategy was adopted in hopes that sufficient descriptive data would be obtained to characterize this population. A broad-based testing program also was expected to facilitate identification of specific areas where assessment efforts should be focused.

The testing program was designed to collect data relevant to studying treatment outcomes. To accomplish this task, the authors included instruments that previously had been used in outcome studies with related populations. There was special interest in comparing findings for female substance abusers with those for male substance abusers and female psychiatric patients. Accordingly, measures were selected that had been widely used in mental health and substance abuse treatment research (e.g, the ASI, MMPI, SCID, and Beck Depression Inventory [BDI] [Beck 1978]). As a general strategy, the authors focused on tests with good psychometric properties and large bodies of literature to support their efficacy.

Another objective was to compare the psychometric performance of several commonly used instruments purported to measure the same domains. Through this comparison process, the authors hoped to identify specific tests that were both acceptable to subjects and capable of providing consistent, useful information in a cost-effective manner. To accomplish this task, multiple measures of personality disorder, depression, and substance abuse were administered. Unfortunately, at that time, none of the instruments had norms available for perinatal substance abusers. As a general strategy, the authors supported the development of perinatal norms for use with existing tests as a more effective and cost-effective approach than new test development.

The initial domains of interest included cognitive functioning, psychiatric diagnosis (both *DSM-III-R* Axis I [acute disorders] and Axis II [personality disorders] [American Psychiatric Association 1987]), personality functioning, addiction and psychiatric severity, global psychiatric functioning, and family functioning.

Cognitive Functioning

With regard to cognitive functioning, the authors used a screening tool, the Shipley Institute of Living Scale (SILS) (Shipley 1967), which yields an intelligence quotient (IQ) estimate correlating with the Wechsler Adult Intelligence Scale Full-Scale Intelligence Quotient (WAIS IQ) (Zachary 1986). We did not anticipate many subjects with subnormal intelligence in this population. Unfortunately, approximately one-third of the sample evidenced some cognitive impairment (Haller et al. 1993*b*). For these cases, it proved necessary to supplement the screening instrument with a psychologist-administered Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler 1981) to better define the impairment and to potentially qualify the subjects for community-based mental retardation services. This two-tiered assessment strategy has worked reasonably well; it provides general information on those who are functioning within

the normal range of intellectual ability and more specific information about the type and extent of impairment for those who are not. However, administration of the WAIS-R requires a psychologist or skilled psychometrician because test interpretation and scoring are complex, and the entire procedure is both time consuming and expensive.

Psychiatric Diagnosis

From the outset, the authors distinguished between personality functioning or “style” and psychiatric diagnosis (*DSM-III-R*) in the assessment process. One significant drawback to the diagnosis approach is that no information is available for those who fail to meet diagnostic criteria (i.e., who fall below the threshold for detection). Unfortunately, the yes-no categorization typical of diagnosis-oriented instruments such as the SCID or the SIDP-R provides little useful information beyond base rates for the various mental disorders. Also, instruments such as these are not particularly helpful in conducting treatment outcome research because that change is difficult to detect unless someone recovers completely or, conversely, develops a mental disorder during treatment.

The authors employed multiple measures designed to yield psychiatric diagnoses. These included the SCID, SIDP-R, MCMI-II, and MMPI-2. The SCID is a structured clinical interview that identifies Axis I psychiatric disorders. Similarly, the SIDP-R is a structured clinical interview that identifies Axis II psychiatric disorders. In contrast, the MCMI-II comprises numerous scales paralleling *DSM-III-R* diagnoses, both Axis I and Axis II. Finally, scores on several clinical and research scales derived from the MMPI item pool (including the MMPI Personality Disorder Scales [Morey 1988]) were compared with findings for the other tests that measure similar domains (e.g., SIDP-R and MCMI-II).

Data derived from these multiple sources have allowed the authors to begin to characterize the population in terms of psychopathology. Relatively high rates of emotional disturbance were found across the board (Haller et al. 1993b). For the substance abuse disorders (and most other Axis I disorders), there has been relatively good agreement across tests, suggesting little need to employ multiple instruments. One exception is the MCMI-II, which appears to underestimate the prevalence of drug disorders (Haller and Dawson 1994). On the other hand, clinical diagnosis of substance abuse disorders was as effective as SCID diagnosis (93 vs. 94 percent). In comparison, non-substance-abuse Axis I disorders were grossly underdiagnosed by clinicians compared with all the assessment tools used (e.g., 12 percent for depressive disorders compared with 29 percent by the SCID and 37 percent by the MCMI-II). The bias against diagnosing non-substance-abuse Axis I disorders is

curious because staff members were trained mental health professionals with at least a master's degree. One explanation is that people who work in the addiction field are conservative about making Axis I diagnoses until they have had ample opportunity to observe the subject drug-free.

In comparison, intertest agreement for Axis II disorders was poor. Base rates for the various disorders fluctuated widely, depending on the instruments being compared (Haller 1995). Interestingly, some personality disorders were measured more reliably than others. For instance, base rates for histrionic personality disorder varied from a low of 2 percent on the MMPI-2 to a high of 63 percent on the MCMI-II; in contrast, the rates for schizoid personality disorder varied between 7 percent on the SIDP-R and 14 percent on the MCMI-II. Overall, clinicians did an exceedingly poor job of identifying Axis II disorders. For example, for antisocial personality disorder (ASP), the base rate for detection was 2 percent for clinicians compared with 33 to 73 percent by test; indeed, the rate of "confirmed" diagnosis (two or more tests agreeing) for ASP was 57 percent. These discrepant findings highlight the importance of using supplemental tests to diagnose Axis II disorders in a clinical setting.

The lack of consistency in measurement of personality disorders is a serious problem. Because investigators continue to publish studies that use different instruments, the literature may be biased. That is, findings may reflect test-sensitivity differences rather than true population differences. Clearly, studies that use different test instruments cannot be compared with confidence. More important, the lack of agreement between test instruments designed to measure the same construct may indicate fundamental difficulties with the personality disorder constructs. Indeed, the current classification system has been criticized for having excessive overlap among personality disorders (Widiger et al. 1986) and for failing to consider the personality disorders as continuous, dimensional constructs (Cloninger 1987).

Personality Functioning

Both the MMPI-2 and the MCMI-II have been widely used to study treatment outcomes in male substance abusers and in both male and female psychiatric patients. Using these instruments allowed the authors to compare findings for substance-abusing women with these reference groups. Both instruments were recently revised; we elected to use the most recent versions (MMPI-2 and MCMI-II) so that changes in procedure would not need to be made once the study had commenced. (Note: The MCMI-III [Millon 1994] is now available.)

Both the MMPI-2 and the MCMI-II have yielded interesting data about personality style for the sample of mostly African-American, cocaine-dependent women. Overall, findings are similar to those reported in the literature for men, with a few important distinctions. Specifically, at intake these women are more psychologically disturbed than their male counterparts, as evidenced by higher test elevations and more pathological profile configurations. However, sex differences appear to diminish with treatment; that is, after these women undergo 5 months of treatment, findings on the MMPI-2 approximate those reported in the literature for male substance abusers (Haller and Dawson 1994). Test profiles for perinatal substance abusers who accept and reject treatment are also different, with refusers looking similar to accepters after 5 months of treatment (Haller et al. 1995). Both tests are useful in assessing personality change as a function of treatment. Significant decreases were evidenced for most scales on both the MMPI-2 and the MCMI-II, and the frequency of certain personality configurations or patterns generated changed considerably from intake to month 5 (Haller and Dawson 1994).

The authors were surprised to experience relatively few difficulties in the administration of lengthy personality inventories. Although many subjects had low general intelligence, all but one subject were able to read sufficiently to complete self-administered instruments (n=155 admissions). Although the test content is probing, the vast majority of subjects were tolerant of this. Approximately 85 percent of self-administered objective personality test protocols were valid at first administration; the remainder were deemed invalid because of subjects' adopting a deviant response set (15 percent of MMPI-2s and 17 percent of MCMI-IIs). It was standard procedure, whenever possible, to retest subjects who generated invalid protocols. At the time of retest, valid MMPI-2 protocols were generated by 42 percent of subjects and valid MCMI-II protocols by 50 percent.

Addiction and Psychiatric Severity

The ASI was used to assess severity of dysfunction in multiple domains. Of particular importance to the psychopathology project were ratings of alcohol, other drugs, and psychiatric severity. Although few administration difficulties have been noted for the ASI, the test may not be as sensitive to change in this population as in others. Nevertheless, in preliminary studies, changes were found from discharge (for treatment completers) to first followup (6 months postdischarge). More specifically, there is a tendency for subjects who have completed treatment to improve in terms of medical, psychiatric, and drug severity. However, the authors are unsure about the test's ability to assess change in other domains. For example, composite scores in the employment domain tend to become higher (meaning more dysfunctional) as a function of treatment. This

suggests that unemployment may not be perceived as a problem until the treatment program identifies it as such. ASI findings to determine the overall appropriateness of this instrument with the perinatal population currently are being evaluated. On the positive side, perinatal norms now have been developed by the instrument's author (A.T. McLellan, personal communication, 1993). Normative information will greatly increase the utility of the test with the perinatal population.

Although the authors do not have personal experience with the Addiction Problem Survey (APS) (DePhilippis et al. 1994, p. 56), we are interested in evaluating this instrument. The APS is a brief, self-administered instrument that assesses severity in the same domains as the ASI. The test's authors (DePhilippis et al. 1994) compared findings for the ASI and APS; initial results have been promising. Therefore, there may be shortcuts to obtaining information similar to that provided by the ASI.

Global Psychiatric Functioning

Global psychiatric functioning was assessed by means of multiple measures, including the Symptom Checklist-90-R (SCL-90-R) (Derogatis 1993), BDI (Beck 1978), and Hamilton Depression Rating Scale (Hamilton 1967). At intake, the SCL-90-R findings reflected high levels of emotional distress on every scale; all elevations subsided with treatment. However, relatively little additional information was obtained through use of the SCL-90-R that was not available from the objective personality tests. Information obtained from the two depression scales was also redundant (identification by BDI, 64 percent; by Hamilton, 54 percent; by both, 43 percent). Clearly, only one depression scale is needed if this domain is to be independently assessed.

Family Functioning

When this work was begun, the authors had little personal experience with the assessment of families and family functioning. Many subjects continued to be involved with their families of origin and also had families of their own; clearly, family relations and interactions were salient. Therefore, instruments were sought that would characterize a woman's family of origin, current adult relationship with a significant other, and current family constellation. Because few measures were available that could address these needs, we selected what appeared to be relevant instruments, despite limited documentation as to their validity and reliability.

The Family of Origin Scale (FOS) (Hovestadt 1985) is a 40-item self-report inventory designed to assess adults' perceptions of the intimacy and

autonomy of the families in which they were reared. Subsequently, doubt was cast on the validity and multidimensionality assumptions of the FOS (Gavin and Wamboldt 1992; Fedor 1992; Manley et al. 1990; Mazer et al. 1990). The authors also have found the FOS to provide only limited information on the population. For example, there is no evidence of change in perinatal substance abusers' views of their families of origin despite treatment. This may be an accurate reflection; on the other hand, it may indicate that the instrument is insensitive to changes that do occur. Also measured were subjects' current family status in areas of competence, cohesion, leadership, and expressiveness using the Self-Report Family Inventory (SFI) (Beavers et al. 1985). This is a relatively well-researched instrument that appears to provide meaningful data to the project. For instance, on the SFI, leadership (the degree to which the parents are "in charge") decreased from intake to month 5 and then increased at followup. This may indicate that perinatal families become more chaotic initially under the stress of substance abuse treatment. However, parental leadership reemerges, and the family appears to reorganize by followup. An alternative interpretation is that of defensive posturing at intake, followed by the subject's admission of problems, and subsequent real improvements in leadership. Because no psychometrically sound instruments were available to measure subjects' primary relationships, a questionnaire was adapted from the Acquaintance Description Form (Wright and Wright 1990) to create the Relationship Assessment Form (RAF) (McCall 1991). The RAF provides information on subjects' tendencies to minimize difficulties, take responsibility, place blame, and display other behaviors relevant to their relationships with spouses or lovers. Data analyses on all family measures are under way; initial results seem promising for the SFI and RAF. However, family measurement is an area that deserves further attention in perinatal substance abuse research.

FUTURE DIRECTIONS FOR ASSESSMENT OF PERINATAL SUBSTANCE ABUSERS

During the past 5 years, the authors identified several significant gaps in our assessment program. Unfortunately, some important domains were completely neglected. These included neuropsychological functioning (including learning disabilities), experience with violence (both as victim and perpetrator), social support, and parenting experience. We also failed to include stages and processes of change data in our initial test protocol. Clearly, the concept of readiness to change behavior should be viewed as an important covariate in psychotherapy outcome research. Hindsight also suggests that information related to the substance use experience should have been collected. More specifically, data on craving and alcohol and other drug expectancies would be extremely useful when trying to

understand who enters, stays, and succeeds in treatment. The influence of other cognitive factors, such as control orientation, probably needs to be considered in a well-balanced test protocol as well. In essence, by assessing primarily psychiatric domains, the authors ignored normal cognitive processes and learning experiences.

We are attempting to fill these identified gaps by adding to the protocol carefully selected instruments that address population-specific needs. The following instruments are being used in the Center for Substance Abuse Treatment's new Center for Perinatal Addiction-Residential Program (CPA-RP) project protocol or are under consideration for use in other clinical research projects involving this population.

Neuropsychological Functioning

Subjects now complete a brief battery of tests sensitive to brain dysfunction, including the Trailmaking Test, Parts A and B (Trails: The Halstead-Reitan . . . 1985), and the Logical Memory and Visual Reproduction subtests of the Wechsler Memory Scale-Revised (WMS-R) (Wechsler 1987). Also, we have added the reading subtest of the Wide Range Achievement Test III (Jastak and Wilkinson 1993) to obtain a literacy rating. All these are well-standardized tests that have been widely used in both research studies and clinical practice and have norms available for comparison purposes.

Experience With Violence

Few instruments are available that address participation in violent acts, either as victim or perpetrator. We were interested in pursuing the four domains of violation highlighted by the Jefferson Medical College of Thomas Jefferson University in its structured clinical interview: verbal, physical, sexual, and personal freedom. However, a self-administered instrument that focused on recent participation in aggressive activities was needed. Thus, with the assistance of the perinatal group at Jefferson Medical College, we are modifying the Self-Report of Abuse (SRA) (Family Center 1991). The result is the Aggressive Acts Questionnaire (AAQ) (Haller and Woodard 1994), a 40-item instrument that addresses involvement in aggressive acts within the past 30 days only. Scales that assess the four domains of violation, from the perspective of both victim and perpetrator, are included. Norms for males and females and blacks and whites are being obtained for comparison purposes, and a profile sheet depicting extent, frequency, and response to violence is currently under development.

Social Support

Although many measures of social support exist, the majority of these have not been adequately researched. Few have norms, and even fewer have been investigated in a rigorous manner. After reviewing numerous instruments, the authors chose two that were developed by a single research group, the Arizona Social Support Interview Schedule (ASSIS) (Barrera 1981; Heitzmann and Kaplan 1988) and the Inventory of Social Supportive Behaviors (ISSB) (Barrera 1981; Heitzmann and Kaplan 1988). The ASSIS is a structured interview that assesses an individual's perceived social support, actual social support, and satisfaction with received support along with perceived and actual conflict about support. The test measures the domains of material aid, physical assistance, intimate interaction, guidance, feedback, and positive social interactions. The ISSB is a self-administered paper-and-pencil measure of social support. It consists of 40 items rated on a 5-point scale to show how often each event occurred during the preceding month (1=not at all; 5=every day). Factor studies have revealed clusters labeled Guidance, Emotional Support, Tangible Support, Cognitive, Information, and Social Support; the first three account for most of the variance.

Parenting Experience

The Parenting Stress Index (PSI) (Abidin 1992) is designed to identify parent-child systems under stress, the risk for dysfunctional parenting, and the development of emotional pathology in a child. The PSI consists of 120 items answered on a 5-point-type scale. In addition to yielding a total stress score, the PSI generates separate scores reflecting stress in both child and parent domains along with a score for stressful life events. The child domain score represents stress the parent experiences on the basis of the child's temperament and is made up of six subscales: adaptability, acceptability, demandingness, mood, distractibility/hyperactivity, and reinforces parent. The parent domain score assesses stress caused by the parent's own characteristics and social support and is made up of the following seven subscales: depression, attachment, restriction of role, sense of competence, social isolation, relationship with spouse, and parent's health. The score for stressful life events provides a context for the other scores.

Stages and Processes of Change

Subjects' readiness for behavioral change has been studied extensively in nonpatient self-changers. These studies have resulted in the development of a transtheoretical model of change that implies that universal stages and processes of change across problem behaviors may be more important

in predicting change than either problem severity or psychiatric status (Prochaska and DiClemente 1992; Prochaska et al. 1994). Researchers are only beginning to apply the transtheoretical model to substance abusers, and to cocaine abusers in particular (J.S. Rossi, personal communication, October 1993).

Drug Use Experience

Toward the end of the NIDA-funded CPA project, the authors introduced several measures of alcohol or other drug experience, and we have continued to use these in the CPA-RP. Of special interest were subjects' self-efficacy, locus of control in relation to substance use, and craving experiences. Pilot studies are being conducted using a self-efficacy scale adapted from a smokers' scale (DiClemente 1981); the Drinking-Related Internal-External scale (Donovan and O'Leary 1978), adapted for alcohol and other drugs; and several cocaine craving scales, including the Minnesota Cocaine Craving Scale (Halikas et al. 1991) and the Cocaine Craving Questionnaire (Jaffe et al. 1989). It is planned to add instruments measuring the related area of drug expectancies in the near future. The Cocaine Effects Expectancy Questionnaire (Schafer and Brown 1991) is a promising instrument for this purpose.

SUMMARY AND RECOMMENDATIONS

A significant amount of information about perinatal substance abusers was gained through use of CPA's initial test protocol. In particular, the authors identified the need for routine cognitive assessment. Unless a woman's intellectual capacity is known, it is difficult to design programming that will adequately meet her needs. Although the initial program curriculum was too advanced for many of the program's subjects, knowledge about the level of intellectual functioning allowed for appropriate adjustments to be made. At the same time, the majority of subjects tested in the normal range of intellectual ability. For this reason, indepth intellectual assessment would have been an unnecessary expense. It seems most appropriate to adopt a screening strategy and use specialized tests as a supplement whenever findings are abnormal. In addition, the initial protocol failed to include tests for literacy and learning ability. Accordingly, the initial protocol was expanded to include measures of these domains.

Although structured clinical interviews appear to be a preferred form of instrumentation for many grant agencies, the authors have found instruments such as the SCID and SIDP-R to be of relatively little value when cost-benefit comparisons are made. Although these instruments provide high-quality information about base rates for the various psychiatric

disorders, they are not particularly useful in treatment outcome studies. Perhaps the yes-no categorization that is typical of such interviews is simply too gross a measure, especially when many women do not meet the criteria for psychiatric diagnosis. Also, diagnoses alone do not provide the wealth of clinical information that can be obtained from other types of tests.

The objective, self-report personality inventories seem more useful. Their assets include scores on every subject that can be compared at multiple test administration points, ease of administration, and cost-effectiveness. Computer-scoring and -interpreting programs are available for programmatic support as well. Although change on some psychiatric dimensions, such as depression or anxiety, can be measured easily and reliably using disorder-specific instruments such as the BDI, score declines on these psychiatric dimensions also are reflected on the depression subscales of the multidimensional inventories. Thus, depending on the breadth of a testing program, instruments like the MMPI-2 and the MCMI-II can serve multiple purposes. The authors were pleasantly surprised to learn that most subjects tolerated these tests, despite their length and relative lack of face validity relevant to substance abuse. Reading ability also was not a problem (most personality inventories are written on the seventh-grade reading level), and the ability of most subjects to produce valid test results was generally high.

The authors now believe that the almost exclusive focus on measurement of psychiatric domains was a mistake. This emphasis accounted for neglect of such important constructs as learning disability, social support, stage of change, violence, and the parenting experience in the initial test protocol. We now are taking a more multidimensional approach to the evaluation of perinatal substance abusers. It is hoped that, at some point in the future, we will be able to determine the relative importance of these domains in predicting treatment outcome.

An area that was completely overlooked in the initial test protocol pertained to alcohol or other drug response and cognitions about alcohol or other drug use. For instance, we failed to include measures of craving, alcohol or other drug expectancies, and control orientation. All these may be important mediating variables that should be considered, and the purely psychiatric approach is now seen as too simplistic. Perhaps individuals with certain personality characteristics, who are at particular stages of change and who have a given control orientation and specific responses to alcohol or other drug cues, are most likely to be successful in treatment. A multidimensional approach to evaluation is strongly advocated.

The validity of instruments that assess Axis II disorders needs to be considered further. Unfortunately, there is no accepted standard for

detection of personality disorders. *DSM-III-R* (American Psychiatric Association 1987) criteria are highly behavioral in nature, making personality disorders somewhat difficult to assess via objective inventories that do not assess behavior directly. Similarly, face-valid structured clinical interviews allow opportunity for self-report bias and may underestimate the frequency of personality problems. For these reasons, the authors arbitrarily decided that agreement between at least two instruments (as to presence or absence of a given disorder) would be required for diagnosis of personality disorder to provide a minimum standard of convergent validity. Although this procedure increased the confidence level in diagnosing Axis II disorders, it is likely overly conservative because disorders that are diagnosed on a single test are routinely discounted. This procedure is also costly and time consuming. Either researchers need to use more rigorous assessment standards, such as use of multiple tests (possibly generating false negatives), or there needs to be agreement in the field to use specific instruments so that findings can be compared across future studies.

REFERENCES

- Abidin, R.R. *Parenting Stress Index, Form 6*. Charlottesville, VA: Pediatric Psychology Press, 1992.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (Third Edition-Revised)*. Washington, DC: American Psychiatric Press, 1987. 567 pp.
- Anastasi, A. *Psychological Testing, 5th Edition*. New York: Macmillan, 1982. 784 pp.
- Barrera, M. Preliminary development of a scale of social support. *Am J Community Psychol* 9:435-447, 1981.
- Beavers, W.R.; Hampson, R.B.; and Hulgus, Y.F. Commentary: The Beavers system approach to family assessment. *Fam Process* 24:398-405, 1985.
- Beck, A.T. *Beck Depression Inventory (BDI)*. The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1978.
- Clavelle, P.R. Clinicians' perceptions of the comparability of the MMPI and MMPI-2. *Psychol Assessment* 4:466-472, 1992.
- Cloninger, C.R. A systematic method for clinical description and classification of personality variants. *Arch Gen Psychiatry* 44:573-588, 1987.
- Cronbach, L.J. *Essentials of Psychological Testing, 4th Edition*. New York: Harper & Row, 1984. 630 pp.

- DePhilippis, D.; Metzger, D.S.; and Navaline, H.A. The Addiction Problem Survey (APS): A valid, self-report measure of addiction problem severity. In: Harris, L.S., ed. *Problems of Drug Dependence, 1993: Proceedings of the 55th Annual Scientific Meeting, The College on Problems of Drug Dependence, Inc., Volume II*. National Institute on Drug Abuse Research Monograph 141. NIH Pub. No. 94-3749. Washington, DC: Supt. of Docs., U.S. Govt. Print. Off., 1994.
- Derogatis, L.R. *Symptom Checklist-90-R (SCL-90-R)*. Minneapolis, MN: National Computer Systems, 1993.
- DiClemente, C.C. Self-efficacy and smoking cessation maintenance: A preliminary report. *Cognit Ther Res* 5:175-187, 1981.
- Donovan, D.M., and O'Leary, M.R. The drinking-related locus of control scale: Reliability, factor structure, and validity. *J Stud Alcohol* 39:759-784, 1978.
- Family Center. "The Self-Report of Abuse (SRA)." Unpublished structured clinical interview, Jefferson Medical College, 1991.
- Fedor, A.J. A validation of the Family of Origin Scale using alcoholic and mental health patients. *Dissertation Abstracts Int* 52:2411, 1992.
- Gavin, L.A., and Wamboldt, F.S. A reconsideration of the Family of Origin Scale. *J Marital Fam Ther* 18:179-188, 1992.
- Halikas, J.A.; Kuhn, K.L.; Crosby, R.; Carlson, G.; and Crea, F. The measurement of craving in cocaine patients using the Minnesota Cocaine Craving Scale. *Compr Psychiatry* 32:22-27, 1991.
- Haller, D.L. Unpublished data, Virginia Commonwealth University, 1995.
- Haller, D.L., and Dawson, K.S. "Changes in Personality as a Function of Intensive Outpatient Therapy: A Study of Perinatal Substance Abusers." Unpublished manuscript, Virginia Commonwealth University, 1994.
- Haller, D.L.; Dawson, K.S.; and Ingersoll, K.S. "Treatment Acceptors vs. Rejectors: A Study of Perinatal Substance Abusers." Unpublished document, 1995.
- Haller, D.L.; Knisely, J.S.; Dawson, K.S.; and Schnoll, S.H. Perinatal substance abusers: Psychological and psychosocial characteristics. *J Nerv Ment Dis* 181(8):509-513, 1993b.
- Haller, D.L.; Motley, C.E.W.; and Schnoll, S.H. Personality and addiction, focus on women. In: Watson, R.R., ed. *Addictive Behaviors in Women*. Totowa, NJ: Humana Press, 1993a.
- Haller, D.L., and Woodard, C. "The Aggressive Acts Questionnaire (AAQ)." Unpublished instrument, Virginia Commonwealth University, 1994.
- Hamilton, M. Development of a rating scale for primary depressive illness. *Brit J Soc Psychol* 6:278-296, 1967.
- Harrell, T.H.; Honaker, L.M.; and Parnell, T. Equivalence of the MMPI-2 with the MMPI in psychiatric patients. *Psychol Assessment* 4:460-465, 1992.

- Hathaway, S.R., and McKinley, J.C. *The Minnesota Multiphasic Personality Inventory Manual*. Minneapolis, MN: National Computer Systems, 1967.
- Hathaway, S.R., and McKinley, J.C. *Minnesota Multiphasic Personality Inventory-2 (MMPI-2)*. The Regents of the University of Minnesota. Minneapolis, MN: National Computer Systems, 1989.
- Heitzmann, C.A., and Kaplan, R.M. Assessment methods for measuring social support. *Health Psychol* 7(1):75-109, 1988.
- Hovestadt, A.J. A Family of Origin Scale. *J Marital Fam Ther* 11:287-297, 1985.
- Jaffe, J.H.; Cascella, N.G.; Kumor, K.M.; and Sherer, M.A. Cocaine-induced cocaine craving. *Psychopharmacology* 97(1):59-64, 1989.
- Jastak, S., and Wilkinson, G.S. *Wide Range Achievement Test-Revision 3*. Wilmington, DE: Jastak Associates, 1993.
- Manley, C.M.; Searight, H.R.; Binder, A.F.; and Russo, R.J. The Family-of-Origin Scale: Factorial validity for adolescents. *Fam Ther* 17(1):75-82, 1990.
- Mazer, G.E.; Mangrum, O.L.; Hovestadt, A.J.; and Brashear, R.L. Further validation of the Family of Origin Scale: A factor analysis. *J Marital Fam Ther* 16:423-426, 1990.
- McCall, D. "The Relationship Assessment Form." Unpublished instrument, Virginia Commonwealth University, 1991.
- McLellan, A.T.; Luborsky, L.; O'Brien, C.P.; and Woody, G.E. An improved evaluation instrument for substance abuse patients: The Addiction Severity Index. *J Nerv Ment Dis* 168:26-33, 1980.
- Millon, T. *Millon Clinical Multiaxial Inventory-II*. Minneapolis, MN: National Computer Systems, 1987. 249 pp.
- Millon, T. *MCMI-III Manual*. Minneapolis, MN: National Computer Systems, 1994.
- Morey, L.C. *The MMPI Personality Disorder Scales: A Manual and Guide to Interpretation*. 1988.
- Pfohl, B.; Blum, N.; Zimmerman, M.; and Stangl, D. *Structured Interview for DSM-III-R Personality-Revised (SIDP-R)*. Ames, IA: Department of Psychiatry, University of Iowa, 1989.
- Prochaska, J.O., and DiClemente, C.C. Stages of change in the modification of problem behaviors. In: Hersen, M.; Eisler, R.M.; and Miller, P.M., eds. *Progress in Behavioral Medicine*. Sycamore, IL: Sycamore Publishing, 1992.
- Prochaska, J.O.; Velicer, W.F.; Rossi, J.S.; Goldstein, M.S.; Marcus, B.H.; Rakowski, W.; Fiore, C.; Harlow, L.; Redding, C.A.; Rosenbloom, D.; and Rossi, S.R. Stages of change and decisional balance for 12 problem behaviors. *Health Psychol* 13:39-46, 1994.
- Schafer, J., and Brown, S.A. Marijuana and cocaine effect expectancies and drug use patterns. *J Consult Clin Psychol* 59:558-565, 1991.

- Shibley, W.C. *Shibley Institute of Living Scale (SILS)*. Los Angeles: Western Psychological Services, 1967.
- Spitzer, R.L.; Williams, J.; and Gibbon, M. *Structured Clinical Interview for DSM-III-R: Non Patient Version (SCID)*. New York: Biometrics Research Department, New York State Psychiatric Institute, 1988.
- Trails: The Halstead-Reitan Neuropsychological Test Battery*. Trail Making Test. Tucson, AZ: Neuropsychology Press, 1985.
- Wechsler, D. *Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1981.
- Wechsler, D. *Wechsler Memory Scale-Revised (WMS-R)*. The Psychological Corporation. San Antonio, TX: Harcourt Brace Jovanovich, 1987.
- Widiger, T.A.; Frances, A.; Warner, L.; and Bluhm, C. Diagnostic criteria for the borderline and schizotypal personality disorders. *J Abnorm Psychol* 95:43-51, 1986.
- Wright, P.H., and Wright, K.D. Measuring codependents' close relationships: A preliminary study. *J Subst Abuse* 2:335-344, 1990.
- Zachary, R.A. *Shibley Institute of Living Scale: Revised Manual*. Los Angeles: Western Psychological Services, 1986.

ACKNOWLEDGMENT

This research was supported by NIDA grant R18-DA-06094.

AUTHORS

Deborah L. Haller, Ph.D.
Associate Professor of Psychiatry, Internal Medicine, and Anesthesiology
Departments of Psychiatry, Internal Medicine, and Anesthesiology
(804) 828-9925 (Tel)
(804) 828-9906 (Fax)
dhaller@gems.vcu.edu (Internet)

Karen S. Ingersoll, Ph.D.
Assistant Professor of Psychiatry
Department of Psychiatry
(804) 828-7456 (Tel)
(804) 828-9906 (Fax)
kingersoll@gems.vcu.edu (Internet)

Medical College of Virginia Hospitals
Virginia Commonwealth University
Box 980109
Richmond, VA 23298-0109

[Click here to go to page 163](#)