OUTPATIENT SCREEN: An Inventory of Scientific Findings

Professional Online Testing Solutions, Inc.
www.online-testing.com

TABLE OF CONTENTS

PREFACE

This document is a cumulative research record of Outpatient Screen's development. Research studies are presented chronologically from earliest to most recent. Research presented herein serves as an approximation of the Outpatient Screen's reliability, validity and accuracy. Outpatient Screen's development focused on the unique needs of YKHC clientele. Alaskan Natives, especially in rural areas, face special challenges, some of which may be due to generational and culture-related factors. Assessments designed with specific populations in mind are often more suited to effectively gather information about these clients despite cultural and other challenges. To facilitate culturally-sensitive assessment and enhance client understanding, Outpatient Screen will be translated into the Yupik, Cupik and other Native Alaskan dialects as deemed appropriate by YKHC.

Outpatient Screen questions and scales were obtained from established Behavior Data Systems (BDS) tests that have been standardized on different male and female adult populations: municipal court defendants, substance (alcohol and drugs) abusers, probationers, domestic violence perpetrators, DUI/DWI offenders, etc. Representative YKHC Alaskan Native clients will be incorporated into the Outpatient Screen database for standardization and research purposes.

Outpatient Screen is a computer-scored assessment instrument designed specifically for the assessment of Alaskan Native clients. When treatment staff members, judges, probation officers and other professionals refer clients for evaluation, Outpatient Screen results are used to help determine when counseling or treatment is warranted. If warranted, programs and services would be provided by YKHC or referred to out-of-region agencies. Outpatient Screen incorporates an Intervention Checklist, allowing clients to indicate their desire to participate in (or continue to participate in) selected services and programs. This information is helpful when determining individual treatment needs.

Outpatient Screen is a brief and easily administered test. The 103 test items are in true/false and multiple choice format and can be completed in less than 30 minutes. Once test data is inputted online, scoring and report printing are completed in less than 3 minutes. Outpatient Screen is composed of five pre-existing, empirically-based scales: Truthfulness, Anger, Alcohol, Drugs and Stress Management. The scales included in Outpatient Screen have been developed and fine-tuned over a 20-year period. Items included in each scale were selected based on their reliability, validity and other statistical properties.

The Outpatient Screen report explains what a client's attained scores mean and makes specific score-related intervention and treatment recommendations. Its five scales are a comprehensive profile of client risk and needs. As noted, the Intervention Checklist provides additional important information about a client's motivation and willingness to work through problems. BDS test scales used in the Outpatient Screen have demonstrated reliability, validity and accuracy. The five scales attain strong correlations with both experienced staff judgment and other established tests. No decision or diagnosis should be based solely on Outpatient Screen results. Outpatient Screen research is ongoing in nature, so that Outpatient Screen reports can provide test users, assessors and evaluators with the most accurate information possible.

INTRODUCTION

OUTPATIENT SCREEN

Outpatient Screen is designed to evaluate violence-prone clients, clients with substance (alcohol and other drugs) misuse disorders and the emotionally distressed. It is used to identify and measure the severity of client problems. In Outpatient Screen reports, quantitative information is obtained by evidence based measures (scales) which independently generate risk (percentile) scores. Scale development is based upon nearly 20 years of research. In addition, explanatory paragraphs describe attained scores and contain specific score-related recommendations. Each scale score is also presented graphically in the comprehensive Outpatient Screen report (profile).

OUTPATIENT SCREEN
MEASURES OR SCALES

1. **Truthfulness Scale**
2. **Anger Scale**
3. **Alcohol Scale**
4. **Drugs Scale**
5. **Stress Management Scale**

The Outpatient Screen is a brief, easily administered and interpreted risk screening or assessment instrument that represents the latest developments in psychometric techniques and computer technology. The Outpatient Screen is scored and interpreted with a computer which generates Outpatient Screen reports.

The Outpatient Screen can be administered individually or in groups. Automated scoring procedures help insure objectivity and accuracy. The Outpatient Screen is to be used in conjunction with a review of available records, a focused interview and experienced court staff judgment.

The Outpatient Screen was designed to provide carefully developed measures (called scales) of several behavioral patterns and traits of interest to those working with perpetrators of domestic violence. The measures (scales) incorporated in the Outpatient Screen further the understanding of the client. In addition, they provide important information on the client's test-taking attitude, emotional/behavioral adjustment, cooperativeness and much more.

UNIQUE FEATURES

Truth Correction: A sophisticated psychometric technique permitted by computerized scoring involves "truth-corrected" scores which are calculated individually for Outpatient Screen scales. Since it would be naive to assume everybody responds truthfully while completing any self-report test, the Truthfulness Scale was developed. **The Truthfulness Scale establishes how truthful an individual is in terms of Outpatient Screen responses.** Correlations between the Truthfulness Scale and all other scales permit identification of error variance associated with untruthfulness. This error variance can then be added back into scale scores, resulting in more accurate "Truth-Corrected" scores. Unidentified denial

wants you to know. Truth-Corrected scores reveal what the client is trying to hide. Truth-Corrected scores are more accurate than raw scores.

Risk Range Percentile Scores: Each Outpatient Screen scale is scored independently of the other scales. The Truthfulness Scale applies a truth-correction factor so that each scale score is referred to as a Truth-Corrected scale score. These Truth-Corrected scale scores are converted to the percentile scores that are reported in the client Outpatient Screen report.

Outpatient Screen scale percentile scores represent *degree of severity*. Degree of severity is defined for all scales as follows: **Low Risk** (zero to 39th percentile), **Medium Risk** (40th to 69th percentile), **Problem Risk** (70th to 89th percentile), and **Severe Problem** or **Maximum Risk** (90th to 100th percentile).

Standardization data is statistically analyzed where percentile scale scores are derived from obtained scale scores from offender populations. The cumulative distributions of truth-corrected scale scores determine the cut-off scores for each of the four risk range and severity categories. Individual scale score calculations are automatically performed and results are presented in the Outpatient Screen report numerically (percentile), by attained risk category (narrative) and graphically (Outpatient Screen profile).

Outpatient Screen Database: Every time an Outpatient Screen is scored, the test data is automatically stored in the Outpatient Screen online database. This database will be statistically analyzed annually, at which time future Outpatient Screen test updates would reflect demographic changes or trends that might have occurred. This unique and proprietary database also enables the formulation of annual summary reports that are descriptive of the populations tested. Summary reports provide important testing information, for budgeting, planning, management and program description.

Confidentiality (Delete Client Names): Many agencies and programs are rightfully concerned about protecting their clients' confidentiality. The proprietary Delete Client Names option is provided to allow deletion of client names from the online Outpatient Screen database. Deleting client names does not delete demographic information or test data. It only deletes the client names when the option is used. Once the client names are deleted there can be no further editing of the client names.

DESCRIPTION OF EVIDENCE BASED MEASURES OR SCALES

Scales used in the Outpatient Screen were developed from large item pools. Initial item selection was a rational process based upon clearly understood definitions of each scale. Each scale is evidence-based. Subsequently, items and scales were analyzed for final test selection. The original pool of potential test items was analyzed and the items with the best statistical properties were retained. **Final test and item selection was based on each item's statistical properties**. It is important that users of the Outpatient Screen familiarize themselves with the definition of each scale. For that purpose a description of each Outpatient Screen scale follows.

Truthfulness Scale: This scale is a measure of the truthfulness of the client while completing the Outpatient Screen. Obtained scores are categorized in terms of percentiles and risk levels, i.e., Low Risk, Medium Risk, Problem Risk, and Severe Problem (Maximum Risk).

All interview and self-report information is subject to the dangers of untrue answers due to defensiveness, guardedness or deliberate falsification. The straightforward nature of any self-report questionnaire may appear to some people as intrusive -- giving rise to denial, faking and even distortion. The Truthfulness Scale identifies these self-protective, recalcitrant and guarded people who minimize or even conceal information. It is equally important to establish that the client understood the test items he or she was responding to, and the Truthfulness Scale also helps identify client comprehension or lack thereof. A high Truthfulness Scale score (at or above the 90th percentile) invalidates all scale scores.

Alcohol Scale: This scale identifies alcohol use and quantifies the severity of abuse (if present). Attained scale scores are categorized in terms of percentiles and severity intervention levels. An elevated score at or above the 90th percentile identifies dependency and severe problems. *Alcohol* refers to beer, wine and other liquor and is a legal (licit) substance.

Drug Scale: This scale identifies drug involvement and measures the severity of abuse, when warranted. As with all other Outpatient Screen scale scores, attained scores are categorized in terms of percentiles and severity intervention levels.

A drug may be broadly defined as any chemical substance that affects living processes. This definition includes marijuana, crack, cocaine, heroin, ecstasy, barbiturates, etc. The Drugs Scale incorporates both illicit drug involvement and prescription drug abuse.

Independent Alcohol and Drug scales provide specific alcohol and/or drug problem identification and problem measurement, and facilitate effective matching of problem type and severity to treatment modality and intensity.

Anger Scale: Measures the client's self-assertiveness, angry or aggressive behavior. Anger usually refers to social dominance with a hostile tendency. Angry conduct disorders are characterized by persistent, domineering, punitive and even assaultive verbal and physical conduct.

An elevated Anger Scale score is indicative of increased acting out behavior and impulsiveness. Severe Problem (90 to 100th percentile) scorers represent the extreme and can represent a problematic lack of social concern. Severe problem scorers sometimes cross the line between aggressive behavior and inappropriate acting out.

An angry person who also manifests substance (alcohol and other drugs) abuse exacerbates impaired judgment and associated acting out -- a malignant combination. At the least these people can be bothersome and distracting.

Elevated Anger Scale and Stress Management Scale scores can codetermine defiant, rebellious, confrontational and protesting behaviors. The Anger Scale can be interpreted independently or in combination with other Outpatient Screen scales.

Stress Management Scale: The Outpatient Screen Stress Management Scale includes stress items and measures stress handling abilities. This scale goes beyond just measuring stress. It measures how well the individual handles, manages or copes with stress. Clients that score in the severe (90^{th} percentile and higher) range consistently have other serious (diagnosable) emotional or mental health problems.

Stress exacerbates other symptoms of emotional, attitudinal, interpersonal and substance-related problems. Frequency and magnitude of impaired Stress Management are important factors in understanding the substance abuser. A Stress Management Scale score at or above the 90th percentile is typically indicative of a diagnosable mental health problem. It is important to assess or measure the degree of severity of stress coping ability problems. This is done with the Stress Management Scale.

OUTPATIENT SCREEN DEVELOPMENT

Scales used in the Outpatient Screen were developed from large item pools. Initial item selection was a rational process based upon clearly understood definitions of each scale. Subsequently, test items and scales were analyzed for scale item inclusion. Final item selection (and inclusion of scale items) was based upon each items statistical properties.

In the beginning, three Ph.D. level psychologists invited experienced staff at several treatment agencies, shelters and batterer programs to share their ideas as to relevant areas of inquiry. This input helped conceptualize the scales used today in the Outpatient Screen. Then, large item pools were developed for each scale. In a series of preliminary studies these item pools were given to domestic violence offenders, probationers, inpatients, outpatients and other groups.

Evidence-based Outpatient Screen scales (or measures) were finally developed by statistically relating scale item configurations to known client groups. The Outpatient Screen was then normed against the identified client population. Outpatient Screen will be standardized upon the YKHC client population once test data is available.

RESEARCH STUDIES

Reliability refers to consistency of results regardless of who uses the instrument. Preexisting scales now utilized in Outpatient Screen are objective, verifiable and reproducible. Validity refers to a test measuring what it is purported to measure. The Outpatient Screen scales were validated in a series of studies that are summarized in this document. However, it should be re-emphasized that Outpatient Screen research is ongoing in nature.

STRESS QUOTIENT

The Stress Quotient (SQ) or Stress Management Scale is based upon the following mathematical equation:

$$SQ = CS/S \times k$$

The Stress Quotient (SQ) scale is a numerical value representing a person's ability to handle or cope with stress relative to their amount of experienced stress. CS (Coping Skill) refers to a person's ability to cope with stress. S (Stress) refers to experienced stress. *k* (Constant) represents a constant value in the SQ equation to establish SQ score ranges. The SQ includes measures of both stress and coping skills in the derivation of the Stress Quotient (SQ) score. The better an individual's coping skills, compared to the amount of experienced stress, the higher the SQ score.

The Stress Quotient (SQ) scale equation represents empirically verifiable relationships. The SQ scale (and its individual components) lends itself to research. Nine studies were conducted to investigate the validity and reliability of the Stress Quotient or Stress Management Scale.

Validation Study 1: This study was conducted (1980) to compare SQ between High Stress and Low Stress groups. The High Stress group (N=10) was comprised of 5 males and 5 females. Their average age was 39. Subjects for the High Stress group were randomly selected from outpatients seeking treatment for stress. The Low Stress group (N=10) was comprised of 5 males and 5 females (average age 38.7) randomly selected from persons not involved in treatment for stress. High Stress group SQ scores ranged from 32 to 97, with a mean of 64.2. Low Stress group SQ scores ranged from 82 to 156, with a mean of 115.7. The t-test statistical analysis of the difference between the means of the two groups indicated that the High Stress group had significantly higher SQ scores than the Low Stress group ($t = 4.9$, $p < .001$). This study shows that the SQ or Stress Management Scale is a valid measure of stress coping. The Stress Management Scale significantly discriminates between high stress individuals and low stress individuals.

Validation Study 2: This study (1980) evaluated the relationship between the SQ scale and two criterion measures: Taylor Manifest Anxiety Scale and Cornell Index. These two measures have been shown to be valid measures of anxiety and neuroticism, respectively. If the SQ or Stress Management Scale is correlated with these measures it would indicate that the SQ or Stress Management Scale is a valid measure. In the Taylor Manifest Anxiety Scale, high scores indicate a high level of anxiety. Similarly, in the Cornell Index high scores indicate neuroticism. Negative correlation coefficients between the two measures and the SQ were expected because high SQ scores indicate good Stress Management. The three tests were administered to forty-three (43) subjects selected from the general population. There were 21 males and 22 females ranging in age from 15 to 64 years. Utilizing a product-moment correlation, SQ scores attained a -.70 correlation with the Taylor Manifest Anxiety Scale and -.75 with the Cornell Index. Both correlations were significant, in the predicted direction, at the $p < .01$ level. These results support the finding that the Stress Management Scale is a valid measure of Stress Management. The reliability of the SQ was investigated in ten subjects (5 male and 5 female) randomly chosen from this study. A split-half correlation analysis was conducted on the SQ items. The product-moment correlation coefficient (r) was .85, significant at the $p < .01$ level. This correlation indicates that the SQ or Stress Management Scale is a reliable measure. These results support the Stress Management Scale as a reliable and valid measure.

Validation Study 3: In this study (1981) the relationship between the SQ Scale and the Holmes Rahe Social Readjustment Rating Scale (SRRS) was investigated. The SRRS, which is comprised of a self-rating of stressful life events, has been shown to be a valid measure of stress. Three correlation analyses were done. SRRS scores were correlated with SQ scores and separately with two components of the SQ scale: Coping Skill (CS) scores and Stress (S) scores. It was hypothesized that the SQ and SRRS

correlation would be negative, since subjects with lower SQ scores would be more likely to either encounter less stressful life events or experience less stress in their lives. It was also predicted that subjects with a higher CS would be less likely to encounter stressful life events, hence a negative correlation was hypothesized. A positive correlation was predicted between S and SRRS, since subjects experiencing more frequent stressful life events would reflect more experienced stress. The participants in this study consisted of 30 outpatient psychotherapy patients. There were 14 males and 16 females. The average age was 35. The SQ and the SRRS were administered in counterbalanced order. The results showed there was a significant positive correlation (product-moment correlation coefficient) between SQ and SRRS ($r = .4006$, $p<.01$). The correlation results between CS and SRRS was not significant ($r = .1355$, n.s.). There was a significant positive correlation between S and SRRS ($r = .6183$, $p<.001$). The correlations were in predicted directions. The significant correlations between SQ and SRRS as well as S and SRRS support the construct validity of the SQ or Stress Management Scale.

Validation Study 4: This validation study (1982) evaluated the relationship between factor C (Ego Strength) in the 16 PF Test as a criterion measure and the SQ in a sample of juveniles. High scores on factor C indicate high ego strength and emotional stability, whereas high SQ scores reflect good coping skills. A positive correlation was predicted because emotional stability and coping skills reflect similar attributes. The participants were 34 adjudicated delinquent adolescents. They ranged in age from 15 to 18 years with an average age of 16.2. There were 30 males and 4 females. The Cattell 16 PF Test and the SQ scale were administered in counterbalanced order. All subjects had at least a 6.0 grade equivalent reading level. The correlation (product-moment correlation coefficient) results indicated that Factor C scores were significantly correlated with SQ scores ($r = .695$, $p<.01$). Results were significant and in the predicted direction. These results support the SQ or Stress Management Scale as a valid measure of Stress Management in juvenile offenders.

In a subsequent study the relationship between factor Q4 (Free Floating Anxiety) on the 16 PF Test and S (Stress) on the SQ scale was investigated. High Q4 scores reflect free floating anxiety and tension, whereas high S scores measure experienced stress. A high positive correlation between Q4 and S was predicted. There were 22 of the original 34 subjects included in this analysis since the remainder of the original files were unavailable. All 22 subjects were male. The results indicated that Factor Q4 scores were significantly correlated (product-moment correlation coefficient) with S scores ($r = .584$, $p<.05$). Results were significant and in predicted directions. The significant correlations between factor C and SQ scores as well as factor Q4 and S scores support the construct validity of the SQ scale.

Validation Study 5: Psychotherapy outpatient clients were used in this validation study (1982) that evaluated the relationship between selected Wiggins MMPI (Minnesota Multiphasic Personality Inventory) supplementary content scales (ES & MAS) as criterion measures and the SQ scale. ES measures ego strength and MAS measures manifest anxiety. It was predicted that the ES and SC correlation would be positive, since people with high ego strength would be more likely to possess good coping skills. Similarly, it was predicted that MAS and S correlations would be positive, since people experiencing high levels of manifest anxiety would also likely experience high levels of stress. The subjects were 51 psychotherapy outpatients ranging in age from 22 to 56 years with an average age of 34. There were 23 males and 28 females. The MMPI and the SQ were administered in counterbalanced order. The correlation (product-moment correlation coefficient) results indicated that ES and CS were positively significantly correlated ($r = .29$, $p<.001$). MAS and S comparisons resulted in an r of .54, significant at the $p < .001$ level. All results were significant and in predicted directions.

In a related study (1982) utilizing the same population data (N=51) the relationship between the Psychasthenia (Pt) scale in the MMPI and the S component of the SQ scale was evaluated. The Pt scale in the MMPI reflects neurotic anxiety, whereas the S component of the SQ scale measures stress. Positive Pt and S correlations were predicted. The correlation (product-moment correlation coefficient) results indicated that the Pt scale and the S component of the SQ scale were significantly correlated ($r = .58$, $p<.001$). Results were significant and in the predicted direction. The significant correlations between MMPI scales (ES, MAS, Pt) and the SQ scale components (CS, S) support the construct validity of the SQ or Stress Management Scale.

Reliability Study 6: The reliability of the Stress Quotient (SQ) or Stress Management Scale was investigated (1984) in a population of outpatient psychotherapy patients. There were 100 participants, 41 males and 59 females. The average age was 37. The SQ was administered soon after intake. The most common procedure for reporting inter-item (within test) reliability is with Coefficient Alpha. The reliability analysis indicated that the Coefficient Alpha of 0.81 was highly significant ($F = 46.74$, $p<.001$). Highly significant inter-item scale consistency was demonstrated.

Reliability Study 7: (1985) The reliability of the Stress Quotient (SQ) or Stress Management Scale was investigated in a sample of 189 job applicants. There were 120 males and 69 females with an average age of 31. The SQ was administered at the time of pre-employment screening. The reliability analysis indicated that the Coefficient Alpha of 0.73 was highly significant ($F = 195.86$, $p<.001$). Highly significant Cronbach Coefficient Alpha reveals that all SQ scale items are significantly ($p<.001$) related and measure one factor or trait.

Validation Study 8: Chemical dependency inpatients were used in a validation study (1985) to determine the relation between MMPI scales as criterion measures and the Stress Quotient (SQ) Scale or Stress Management Scale. The SQ is inversely related to other MMPI scales, consequently, negative correlations were predicted. The participants were 100 chemical dependency inpatients. There were 62 males and 38 females with an average age of 41. The SQ and the MMPI were administered in counterbalanced order. The reliability analysis results indicated that the Coefficient Alpha of 0.84 was highly significant ($F = 16.20$, $p<001$). Highly significant inter-item scale consistency was demonstrated.

The correlation (product-moment correlation coefficient) results between the Stress Quotient (SQ) and selected MMPI scales were significant at the $p < .001$ level and in predicted directions. The SQ correlation results were as follows: Psychopathic Deviate (-0.59), Psychasthenia (-.068), Social Maladjustment (-0.54), Authority Conflict (-0.46), Taylor Manifest Anxiety Scale (-0.78), Authority Problems (-0.22), and Social Alienation (-0.67). The most significant SQ correlation was with the Taylor Manifest Anxiety Scale. As discussed earlier, stress exacerbates symptoms of impaired adjustment as well as emotional and attitudinal problems. These results support the Stress Quotient or Stress Management Scale as a valid measure of Stress Management.

Validation Study 9: In a replication of earlier research, a study (1986) was conducted to further evaluate the reliability and validity of the Stress Quotient (SQ). The participants were 212 inpatients in chemical dependency programs. There were 122 males and 90 females with an average age of 44. The SQ and MMPI were administered in counterbalanced order. Reliability analysis of the SQ scale resulted in a Coefficient Alpha of 0.986 ($F = 27.77$, $p<.001$). Highly significant inter-item scale consistency was again demonstrated. Rounded off, the **Coefficient Alpha for the SQ was 0.99**.

In the same study (1986, inpatients), product-moment correlations were calculated between the Stress Quotient (SQ) and selected MMPI scales. The SQ correlated significantly (.001 level) with the following MMPI scales: Psychopathic Deviate (Pd), Psychasthenia (Pt), Anxiety (A), Manifest Anxiety (MAS), Ego Strength (ES), Social Responsibility (RE), Social Alienation (PD4A), Social Alienation (SC1A), Social Maladjustment (SOC), Authority Conflict (AUT), Manifest Hostility (HOS), Suspiciousness/Mistrust (TSC-II), Resentment/Aggression (TSC-V) and Tension/Worry (TSC-VII). **All SQ correlations with selected MMPI scales were significant (at the .001 level of significance) and in predicted directions.** These results support the SQ scale or Stress Management Scale as a valid measure of Stress Management.

The studies cited above demonstrate empirical relationships between the SQ scale (Stress Management Scale) and other established measures of stress, anxiety and coping skills. This research demonstrates that the Stress Quotient (SQ) or Stress Management Scale is a reliable and valid measure of Stress Management. The SQ has high inter-item scale reliability. The SQ also has high concurrent (criterion-related) validity with other recognized and accepted tests. The SQ scale permits objective (rather than subjective) analysis of the interaction of these important variables. In the research that follows, the **Stress Quotient** or **SQ** is also referred to as the **Stress Management Scale**.

OUTPATIENT SCREEN RESEARCH

Early in development, the scales now used in the Outpatient Screen were administered to normals (by definition not offenders, probationers, defendants, etc.), college students, substance abuse patients, inmates and defendants. The Outpatient Screen does differentiate between "normals" and clients with known problems. And, scale scores correlate well with other tests measuring similar behaviors.

10. Validation of the Truthfulness Scale

The Truthfulness Scale is an important psychometric scale as these scores establish how truthful the respondent was while completing the Outpatient Screen. Truthfulness Scale scores determine whether or not profiles are accurate and are integral to the calculation of Truth-Corrected scale scores.

The Truthfulness Scale identifies respondents who were self-protective, recalcitrant and guarded, as well as those who minimized or even concealed information while completing the test. Truthfulness Scale items are designed to detect respondents who try to fake good or put themselves into a favorable light. These scale items are statements about oneself that most people would agree to. The following statement is an example of a Truthfulness Scale item, "Sometimes I worry about what others think or say about me."

This preliminary study used the Truthfulness Scale items that are now included in the Outpatient Screen to determine if these Truthfulness Scale items could differentiate between respondents who were honest from those trying to fake good. It was hypothesized that the group trying to fake good would score higher on the Truthfulness Scale than the group instructed to be honest.

Method

Seventy-eight Arizona State University college students (1985) enrolled in an introductory psychology class were randomly assigned to one of two groups. Group 1 comprised the “Honest” group and Group 2 comprised the “Fakers” group. Group 1 was instructed to be honest and truthful while completing the test. Group 2 was instructed to "fake good" while completing the test, but to respond "in such a manner that their faking good would not be detected." The test, which included the Outpatient Screen Truthfulness Scale, was administered to the subjects and the Truthfulness Scale was embedded in the test as one of the six scales. Truthfulness Scale scores were made up of the number of deviant answers given to the 21 Truthfulness Scale items.

Results

The mean Truthfulness Scale score for the Honest group was 2.71 and the mean Truthfulness Scale score for Fakers was 15.77. The results of the correlation (product-moment correlation coefficient) between the Honest group and the Fakers showed that the Fakers scored significantly higher on the Truthfulness Scale than the Honest group ($r = 0.27$, $p < .05$). The Truthfulness Scale successfully measured how truthful the respondents were while completing the test. The results of this study reveals that the Truthfulness Scale accurately detects "Fakers" from those students that took the test honestly.

11. Validation of Four Selected Scales using Criterion Measures

In general terms, a test is valid if it measures what it is supposed to measure. The process of confirming this statement is called validating a test. A common practice when validating a test is to compute a correlation between it and another (criterion) test that purports to measure the same thing and that has been previously validated. For the purpose of this study, four scales (Truthfulness, Alcohol, Drug, Stress Management) were validated with comparable scales on the Minnesota Multiphasic Personality Inventory (MMPI). The MMPI was selected for this validity study because it is the most researched, validated and widely used objective personality test in the United States. The selected scales were validated with MMPI scales as follows. The Truthfulness Scale was validated with the L Scale. The Alcohol Scale was validated with the MacAndrew Scale and Psychopathic Deviant Scale. The Drug Scale was validated with the MacAndrew and Psychopathic Deviant Scale. The Stress Management Scale was validated with the Taylor Manifest Anxiety, Psychasthenia, Social Maladjustment and Social Alienation scales or measures.

Method

One hundred (100) chemical dependency inpatients (1985) were administered both the selected scales and the MMPI. Tests were counterbalanced for order effects -- half were given the scales first and half the MMPI first.

Results and Discussion

Product-moment correlation coefficients were calculated between selected scales and MMPI scales. These results are summarized in Table 1. The correlation results presented in Table 1 show that all Outpatient Screen scales significantly correlated (.001 level of significance) with all represented MMPI scales. In addition, all correlations were in predicted directions.

Table 1. (1985) Product-moment correlations between MMPI scales and selected scales

MMPI SCALES (MEASURES)	SCALES (MEASURES)			
	Truthfulness	Alcohol	Drug	Stress
L (Lie) Scale	0.72	-0.38	-0.41	0.53
Psychopathic Deviant	-0.37	0.52	0.54	-0.59
Psychasthenia	-0.34	0.38	0.41	-0.68
Social Maladjustment	-0.25	0.34	0.26	-0.54
Authority Conflict	-0.43	0.31	0.47	-0.46
Manifest Hostility	-0.45	0.34	0.47	-0.58
Taylor Manifest Anxiety	-0.58	0.47	0.46	-0.78
MacAndrew	-0.40	0.58	0.62	-0.33
Social Alienation	-0.47	0.35	0.45	-0.67

The **Truthfulness Scale** correlates significantly with all of the represented MMPI scales in Table 1. Of particular interest is this scale's highly significant positive correlation with the MMPI Lie (L) Scale. A high L Scale score on the MMPI invalidates other MMPI scale scores due to untruthfulness. This helps in understanding why the Truthfulness Scale is significantly, but negatively, correlated with the other represented MMPI scales. Similarly, the MMPI L Scale correlates significantly, but negatively, with the other Outpatient Screen scales.

The **Alcohol Scale** correlates significantly with all represented MMPI scales. This is consistent with the conceptual definition of the Alcohol Scale and previous research that has found that alcohol abuse is associated with mental, emotional and physical problems. Of particular interest are the highly significant correlations with the MacAndrew (r = 0.58) Scale and the Psychopathic Deviant (r = 0.52) Scale. High MacAndrew and Psychopathic Deviant scorers on the MMPI are often found to be associated with substance abuse. Similarly, the **Drug Scale** correlates significantly with the MacAndrew (r = 0.62) Scale and the Psychopathic Deviant (r = 0.54) Scale.

The **Stress Management Scale** is inversely related to MMPI scales, which accounts for the negative correlations shown in Table 1. The positive correlation with the L scale on the MMPI was discussed earlier, i.e., Truthfulness Scale. It should be noted that stress exacerbates symptoms of impaired adjustment and even psychopathology. The Stress Management Scale correlates most significantly with the Taylor Manifest Anxiety (r = -0.78) Scale, the Psychasthenia (r= -0.68) Scale and the Social Alienation (r= -0.67) Scale.

These findings strongly support the validity of Outpatient Screen scales. All of the Outpatient Screen scales were highly correlated with the MMPI criterion scale they were tested against. The large correlation coefficients support the validity of the Outpatient Screen. All product-moment correlation coefficients testing the relation between Outpatient Screen scales and MMPI scales were significant at the $p < .001$ level.

12. Relationships Between Selected Scales and Polygraph Examination

A measure that has often been used in business or industry for employee selection is the Polygraph examination. The polygraph exam is most often used to determine the truthfulness or honesty of an individual while being tested. The Polygraph examination is more accurate as the area of inquiry is more "situation" specific. Conversely, the less specific the area of inquiry, the less reliable the Polygraph examination becomes. Three scales were chosen for this study; Truthfulness Scale, Alcohol Scale and Drug Scale. The Truthfulness Scale was chosen because it is used to measure the

truthfulness of the respondent at the time of assessment. The Alcohol and Drug scales are well suited for comparison with the polygraph exam because of the situation specific nature of the scales. Alcohol and Drug scale items are direct and relate specifically to alcohol and drug use. The comparison with the Truthfulness Scale is less direct because of the subtle nature of Truthfulness Scale items. The Truthfulness Scale is affected by the respondent's attitude, emotional stability and tendencies to fake good. The Alcohol and Drug scales were expected to be highly correlated with the polygraph results and the Truthfulness Scale would show a somewhat less but nonetheless significant correlation.

Method

One hundred and eighty-nine (189) job applicants (1985) were administered both the selected scales and the Polygraph examination. Tests were given in a counterbalanced order, half of the applicants were given the scale items first and the other half of the applicants were administered the polygraph first. The subjects were administered the scales and polygraph exam in the same room in the same session with the examiner present for both tests.

Results

The product-moment correlation results between the Polygraph exam and selected scales indicated there was a significant positive correlation between the Truthfulness Scale and Polygraph exam ($r = 0.23$, $p<.001$). Similarly, significant positive relationships were observed between the Polygraph exam and the Alcohol Scale ($r = 0.54$, $p<.001$) and the Drug Scale ($r = 0.56$, $p<.001$).

In summary, this study supports the validity of the Truthfulness, Alcohol and Drug scales. There were strong positive relationships between the selected scales and the Polygraph examination. The highly significant product-moment correlations between the selected scales and Polygraph examinations demonstrate the validity of the Truthfulness, Alcohol and Drug measures.

These results are important because the Polygraph exam is a direct measure obtained from the individual being tested rather than a rating by someone else. This is similar to self-report such as utilized in the Outpatient Screen. The fact that there was a very strong relationship between Polygraph results and scales shows that this type of information can be obtained accurately in self-report instruments.

These results indicate that the Truthfulness Scale is an accurate measure of the respondent's truthfulness at the time of assessment. The Truthfulness Scale is an essential measure in self-report instruments. There must be a means to determine the honesty or "correctness" of the respondents answers and there must be a means to adjust scores when the respondent is less than honest. The Truthfulness Scale addresses both of these issues. The Truthfulness Scale measures truthfulness and then applies a correction to other scales based on the Truthfulness Scale score. The Truthfulness Scale ensures accurate assessment.

13. Validation of Scales in a Sample of Substance Abuse Inpatients

The present study (1987) was conducted to validate the selected scales in a sample of substance abuse inpatients in a chemical dependency facility.

Selected scales in the Minnesota Multiphasic Personality Inventory (MMPI) were used as criterion measures for the different scales. The Truthfulness Scale was validated with MMPI L Scale, F Scale and K Scale. The Alcohol Scale was validated with MMPI MacAndrew Scale (MAC) and Psychopathic Deviate-Obvious (PD-O). The Drug Scale was validated with MMPI MacAndrew Scale and Psychopathic Deviate-Obvious. The Stress Management Scale was validated with MMPI Psychasthenia (PT), Anxiety (A), Taylor Manifest Anxiety (MAS) and Tension/Worry (TSC-VII). The MMPI scales were chosen to compare to the scales because they measure similar attributes.

Method

The subjects used in the study were 212 substance (alcohol and other drugs) abuse inpatients in chemical dependency facilities. The scales currently used in Outpatient Screen and MMPI scales were administered in counterbalanced order.

Results and Discussion

The product-moment correlation results are summarized in Table 2.

The **Truthfulness Scale** correlates significantly in predicted directions with selected MMPI criterion scales, L Scale (lie, $p<.001$), F Scale (validity, $p<.001$) and K Scale (validity correction, $p<.001$). Other significant correlations with traditional MMPI scales include: PD (Psychopathic deviate, $p<.001$), ES (Ego Strength, $p<.001$), and RE (Social responsibility, $p< .001$); Harris MMPI subscales: PD2 (Authority Problems, $p<.001$), PD4 (Social Alienation, $p<.001$), SCIA (Social Alienation, $p<.001$); Wiggins MMPI content scales: SOC (Social Maladjustment, $p<.001$), HOS (Manifest Hostility, $p<.001$); Wiener-Harmon MMPI subscales: PDO (Psychopathic Deviant-Obvious, $p<.001$); Tryon, Stein & Chu MMPI cluster scales: TSC-V (Resentment/Aggressive, $p<.001$).

The **Alcohol Scale** correlates significantly in predicted directions with selected MMPI criterion scales: MAC (MacAndrew scale, $p<.001$), and PD-O (Psychopathic Deviate Obvious, $p<.021$). The **Drug Scale** correlates significantly in predicted directions with selected MMPI criterion scales: MAC (MacAndrew scale, $p<.001$), and PD-O (Psychopathic Deviate Obvious, $p<.001$).

The **Stress Management Scale** correlates significantly in predicted directions with selected MMPI criterion scales: PT (Psychasthenia, $p<.001$), A (Anxiety, $p<.001$), MAS (Taylor Manifest Anxiety, $p<.001$), PD4 (Social Alienation, $p<.001$) and TSC-VII (Tension/Worry, $p<.001$).

Table 2. Selected Scale-MMPI Product-moment Correlations (1987)
Inpatients, Chemical Dependency Facilities (N = 212)

MMPI SCALES (MEASURES)	SCALES (MEASURES)			
	Truthfulness	Alcohol	Drug	Stress
L	0.60	-0.24	-0.15	-0.30
F	-0.34	0.32	0.32	0.49
K	0.39	-0.28	-0.29	-0.51
MAC	-0.30	0.35	0.37	0.28
PD-O	-0.35	0.22	0.33	0.53
PD2	-0.26	0.18	0.17	0.07
PD	-0.33	0.21	0.33	0.39
HOS	-0.45	0.25	0.33	0.46
TSC-V	-0.46	0.34	0.28	0.58
ES	0.25	-0.27	-0.25	-0.51
RE	0.41	-0.27	-0.34	-0.45
SOC	-0.19	0.17	0.08	0.39
PD4	-0.41	0.20	0.28	0.55
SCIA	-0.36	0.27	0.32	0.39
PT	-0.39	0.27	0.24	0.58
A	-0.41	0.31	0.31	0.68
MAS	-0.44	0.25	0.18	0.65
TSC-VII	-0.41	0.33	0.29	0.66

These findings strongly support the validity of the selected scales in this sample of chemical dependency inpatients. All of the selected scales were highly correlated with the MMPI criterion scales they were tested against. Inpatients in chemical dependency facilities are known to have substance abuse problems and these correlation results confirm the validity of the instruments. These findings, then support the validity of the selected scales.

The Alcohol and Drug scales are direct measures of alcohol and drug use and abuse, respectively, whereas the MacAndrew Scale was developed from discriminant analysis and does not include a truthfulness scale. The MacAndrew Scale items do not relate specifically to alcohol and drugs. Hence, the correlations between the MacAndrew Scale and the Alcohol and Drug scales could be affected by the lack of a truthfulness measure which is a deficiency of the MacAndrew Scale. However, the correlation coefficients were significant.

Where MMPI scales are closely related (by definition) to the selected scales, the correlation coefficients were significant. For example, the Truthfulness Scale and the MMPI L Scale both measure tendencies to fake good, and the correlation was very highly significant at r = .60. The correlation between Stress Management Scale and MMPI Tension/Worry Scale was r = -.66. This study supports the validity of the selected scales.

14. Reliability Study of Selected Scales in Two Samples of Domestic Violence Defendants

Any approach to detection, assessment, or measurement must meet the criteria of reliability and validity. Reliability refers to an instrument's consistency of results regardless of who uses it. This means that the

outcome must be objective, verifiable, and reproducible. Ideally, the instrument or test must also be practical, economical, and accessible. Psychometric principles and computer technology insures accuracy, objectivity, practicality, cost-effectiveness and accessibility.

This study (1991) was conducted to test the scale reliability in two different samples of domestic violence defendants. Within-test reliability measures to what extent a test with multiple scales measuring different factors, measures each factor independent of the other factors (scales) in the test. It also measures to what extent items in each scale consistently measure the particular trait (or factor) that scale was designed to measure. Within-test reliability measures are referred to as inter-item reliability. The most common method of reporting within-test (scale) inter-item reliability is with coefficient alpha.

Method

There were two samples of domestic violence defendants included in this study (1991). **The subjects in Group 1 consisted of 168 domestic violence defendants.** There were 158 (94%) males, and 10 (6.0%) females. The demographic composition of this sample is summarized as follows: Age: 16 to 20 years (7.1%); 21 to 25 years (16.1%); 26 to 30 years (25.6%); 31 to 35 years (22.6%); 36 to 40 years (14.3%); 41 to 45 years (6.5%); 46 to 50 years (3.6%); 51 to 55 years (2.4%); and 56 to 60 years (1.8%). Ethnicity: Caucasian (97%) and Black (3.1%). Education: 8th Grade or less (12.5%); Some High School (38.7%); G.E.D. (5.4%); High School Graduate (36.3%); Some College (4.8%); Technical/Business School (1.2%); College Graduate (0.6%); and Professional/Graduate School (0.6%). Marital Status: Single (17.3%); Married (42.9%); Divorced (25.6%); and Separated (14.3%).

Group 2 consisted of 525 domestic violence defendants. There were 416 (79.2%) males and 109 (20.8%) females. Age: 18 to 20 (16.9%); 21 to 25 (17.1%); 26 to 30 (21.1%); 31 to 35 (17.1%); 36 to 40 (15.2%); 41 to 45 (7.4%); 46 to 50 (2.3%); 51 to 55 (1.3%); 56 to 60 (1.0%); 60 to 65 (0.4%). Ethnicity: Caucasian (65.3%); Black (23.2%); Hispanic (9.3%); Asian (0.4%); American Indian (1.3%) and Other (0.4%). Education: 8th Grade or less (7.1%); Some High School (29.2%); G.E.D. (5.9%); High School Graduate (37.4%); Some College (14.7%); Technical/Business School (0.2%); College Graduate (4.4%); Professional/Graduate School (1.1%); and Missing (0.2%). Marital Status: Single (50.6%); Married (35.7%); Divorced (6.5%); Separated (7.3%); and Missing (0.2%).

Reliability coefficients are presented in Table 3.

Table 3. Reliability coefficient alphas. Domestic violence defendants. (1991, N = 693)
All coefficient alphas are significant at p<.001.

Scales	1 D.V. Defendants N = 168	2 D.V. Defendants N = 525
Truthfulness Scale	.85	.85
Alcohol Scale	.92	.89
Drug Scale	.89	88
Anger Scale	.83	84
Stress Management	.91	.91

The results of this study support the reliability (internal consistency) of the selected scales. All coefficient alphas are significant at p<.001.

15. Reliability of Selected Scales in Two Samples of Convicted Domestic Violence Offenders

This study (1992) was conducted to test the reliability (internal consistency) of selected scales in two samples of clients. All respondents (N = 729) were convicted domestic violence offenders.

Method

There were two samples of clients who participated in the study (1992).

Group 1 consisted of 153 clients. There were 141 males (92%) and 12 females (8%). This sample is described as follows: Age: Under 18 (45.1%); 18 to 25 (17.6%); 26 to 35 (25.5%); 36 to 45 (6.5%); 46 to 55 (3.3%); and over 55 (2.0%). Ethnicity: Caucasian (42.5%); Black (8.5%); Hispanic (32.0%); Asian (5.2%); American Indian (0.7%), and Other (11.1%). Education: 8th Grade or less (2.0%); Some High School (5.9%); G.E.D. (4.6%); High School Graduate (73.2%); Some College (7.8%); Technical/Business School (1.3%); and College Graduate (5.2%). Marital Status: Single (45.1%); Married (43.8%); Divorced (4.6%); Separated (5.9%); and Widowed (0.7%).

Group 2 consisted of 576 adjudicated clients. Of these 576 offenders, 489 were male (84.9%) and 87 were female (15.1%). This sample is described as follows: Age: Under 18 (17.7%); 18 to 25 (28.6%); 26 to 35 (33.0%); 36 to 45 (14.9%); 46 to 55 (4.2%); over 55 (1.6%). Ethnicity: Caucasian (62.3%), Black (15.6%); Hispanic (15.8%); Asian (1.9%); American Indian (0.7%); and Other (3.6%). Education: 8th Grade or less (8.3%); Some High School (24.5%); G.E.D. (3.6%); High School Graduate (46.7%); Some College (11.6%); Technical/Business School (0.5%); College Graduate (3.8%); and Graduate/Professional Degree (0.9%). Marital Status: Single: (46.0%); Married (38.0%); Divorced (5.9%); Separated (9.0%); Widowed (1.0%).

Coefficient alpha is considered an important indicator of internal consistency or reliability. These coefficients are reported in Table 4. The total number of clients included this study was 729.

Table 4. Reliability coefficient alphas. (1992, N = 729)
All coefficient alphas are significant at p<.001.

Outpatient Screen Scales	1 D.V. Offenders N = 153	2. D.V. Offenders N = 576
Truthfulness Scale	.85	.86
Alcohol Scale	.93	.92
Drug Scale	.92	.89
Anger Scale	.81	.86
Stress Management	.90	.92

The results of this study demonstrate the reliability (internal consistency) of the selected scales.

16. Reliability of Selected Scales and Sex Differences in Violent Offenders

Because sex differences were found in other assessment instruments, sex differences in selected scales were examined. For the most part, the clients that have been studied have been primarily male with only a very small percentage being female. The purpose of the present study was to test for sex differences in the different scales among clients.

There were two samples of clients included in the present study (1993), but because the two samples were from different regions of the United States, the databases were kept separate. There were a total of 269 clients included in the study.

Method

There were two samples of clients included in this study (1993). The group contained 152 clients and the second group contained 117 offenders. The demographic composition of group 1 was as follows: There were 152 offenders, 137 were male and 15 were female. Age: 16 to 25 (35.3%); 26 to 35 (35.3%); 36 to 45 (23.5%); 46 to 55 (3.8%); over 55 (2.0%). Ethnicity: Caucasian (56.2%); Black (42.5%); Hispanic (1.3%). Education: 8th Grade or less (6.5%); Some High School (26.8%); G.E.D. (4.6%); High School Graduate (35.9%); Some College (17.6%); College Graduate (5.9%); Graduate/Professional Degree (2.6%). Marital Status: Single (51.6%); Married (30.1%); Divorced (3.9%); Separated (13.7%); and Widowed (0.7%).

Group 2 consisted of 117 domestic violence diversion program participants, 87 (74%) were male and 30 (26%) were female. The demographic composition of this sample was as follows: Age: 18 to 25 (25.6%); 26 to 35 (48.7%); 36 to 45 (17.1%); 46 to 55 (6.8%); and over 55 (1.7%). Ethnicity: Caucasian (63.2%); Black (4.3%); Hispanic (27.4%); Asian (1.7%); American Indian (0.9%); and Other (2.6%). Education: 8th Grade or less (8.5%); Some High School (19.7%); G.E.D. (0.9%); High School Graduate (50.4%); Some College (16.2%); and College Graduate (4.3%). Marital Status: Single (29.1%); Married (46.2%); Divorced (11.1%); Separated (10.3%); and Widowed (3.4%).

Reliability coefficient alphas (internal consistency) for the two clients are reported in Table 5. There were a total of 269 offenders included in this study (1993).The results in Table 5 support the internal consistency or reliability of the selected scales. These findings are in close agreement with previous reliability research.

Table 5. Reliability coefficient alphas in two samples of clients.
All coefficient alpha are significant at p<.001. (1993, N = 269)

Scales	1 Offenders N = 152	2 Offenders N = 117
Truthfulness Scale	.86	.85
Alcohol Scale	.91	.91
Drug Scale	.90	.88
Anger Scale	.85	.85
Stress Management	.92	.91

The results of Group 2 were used to identify sex differences. There was an insufficient number of females in Group 1. Even in Group 2 (N = 117) the distributions for males and females were not equivalent, meaning that the variances of distributions were unequal (and not normally distributed). Because of this t-test comparisons could not be done and the sex differences were tested using the Wilcoxon signed-rank test. The Wilcoxon sign-rank test results indicated no statistically significant gender difference at the .05 level.

17. Validation of the Selected Scales Using Evaluator Ratings

The present study (1993, N = 559) was conducted to determine the relationship between experienced staff ratings and selected scales. Domestic violence diversion program staff screened program applicants for admission as part of their normal routine. While evaluator rating studies tend to be adversely affected by inter-rater reliability, these studies can provide sound validation when the measures to be rated are well defined.

Evaluators were instructed to interview each client, administer and score the selected scales and review client's police reports. After completing their screening procedure, staff were to rate each client. The evaluators were instructed to rate the applicants on scale correlate measures, i.e., truthfulness in interview, substance (alcohol and other drugs) abuse, aggressiveness, violence proneness, and Stress Management. These ratings were to be completed before scales were scored.

Results

The results of this study (1993) are presented in Table 8. The correlations between evaluator ratings and selected scales are significant with the exception of the Stress Management Scale.

Table 6. Product-moment correlations between staff ratings and selected scales. (1993, N = 559)

Scales	Agreement Coefficients	Significance Level
Truthfulness Scale	.10	P< .02
Alcohol Scale	.54	P< .01
Drug Scale	.50	P< .01
Anger Scale	.44	P< .01
Stress Management	.03	P< .57

Scale scores were available after approximately 30 to 35 minutes of testing time. The agreement between staff ratings and scale scores were highly significant. The less significant correlation between Truthfulness Scale scores and staff ratings of client truthfulness ratings is to be expected. The literature consistently notes that individuals tend to under-report their problems when evaluated for referral.

The non-significant correlation between the Stress Management Scale and evaluator ratings is in marked contrast to the Stress Management scale's high concurrent validity with clinical and chemical dependency client populations. In post-study interviews, most staff reported that they did not score the scales until later that day or the next day. In other words, scale score results were, in most cases, unavailable at the time of staff ratings. When the Stress Management Scale is compared to other objective instruments designed to measure stress or anxiety, highly significant correlations are demonstrated.

These results support the validity of the selected scales. Domestic violence evaluator ratings of clients were significantly correlated with scale scores. All scales but the Stress Management Scale were highly correlated with evaluator ratings. The highest correlation coefficients were found with the Alcohol, Drug and Anger Scales. These measures are well defined and evaluator ratings of these scales were in close agreement with the selected scales. These results provide validation evidence for the Outpatient Screen as an accurate instrument for client assessment.

18. Replication of a Scale Validation Study Using Evaluator Ratings

A study (1993) was conducted to replicate an earlier study (cited above) that investigated the relationship between selected scales and staff ratings. It was decided that the earlier study may have been affected by differences in evaluator procedures during the study. Not all evaluators consistently followed the study procedures as instructed.

For clarity, staff were instructed to rate each defendant after they interviewed the applicant, reviewed their TII results and read the police report. Defendants were being screened for admission into a violence-related diversion program. The scales was administered as part of the intake procedure, but scored after all staff ratings were completed.

The results of this study are presented in Table 9. All product-moment correlations demonstrated significant relationships between experienced staff ratings and selected scale scores.

Table 7. Product-moment correlations between staff ratings and selected scales. (1993, N = 1350)
All correlations are significant at p<.01.

Outpatient Screen Scales	Agreement Coefficients
Truthfulness Scale	.34
Alcohol Scale	.53
Drug Scale	.47
Anger Scale	.43
Stress Management	.38

The correlations between staff ratings and the Alcohol, Drug and Anger Scales were in close agreement to those found in the earlier study. However, the correlations with the Truthfulness and Stress Management scales are much higher than previously found. Apparently, evaluators were more consistent in following study procedures and the ratings were based on all available information.

Staff completed this normal assessment procedure, including reviewing scale score results, prior to completing their defendant ratings. These agreement coefficients are all significant, in predicted directions and impressive.

19. A Study of Reliability of Selected Scales in Domestic Violence Defendants

This study (1994) tested the reliability of the Outpatient Screen in a sample of domestic violence defendants.

Method

There were 255 domestic violence defendants included in the present study. There were 237 males (92.9%) and 18 females (7.1%). The demographic composition of the defendants was as follows: Age: 16 to 25 (29.4%); 26 to 35 (43.9%); 36 to 45 (19.2%); 46 to 55 (5.5%); 56 to 65 (2.0%). Ethnicity: Caucasian (51.4%); Black (47.8%), Hispanic (0.4%); American Indian (0.4%). Education: 8th Grade or less (3.9%); Some High School (33.7%); G.E.D. (6.7%); High School Graduate (38.0%); Some College (14.1%); Technical/Business School (0.4%); College Graduate (14.1%); and Professional/ Graduate School (0.8%). Marital Status: Single (47.1%); Married (39.2%); Divorced (7.1%); and Separated (6.7%).

Reliability coefficient alphas are presented in Table 8.

Table 8. Reliability coefficient alphas. Domestic violence defendants. (1994, N = 255)
All coefficient alphas are significant at p<.001.

Scales	Coefficient Alpha
Truthfulness Scale	.87
Alcohol Scale	.92
Drug Scale	.88
Anger Scale	.87
Stress Management	.90

This study supports the reliability (internal consistency) of the selected scales. All scales have highly significant reliability coefficient alphas.

23. Reliability of Selected Scales in Two Samples of Offenders

Two samples of offenders were included in the present study (1995) to further investigate reliability in different samples and assessment milieu. The groups of offenders represented two different geographical areas of the country in similar domestic violence evaluation programs. The purpose of the present study (1995) was to investigate reliability of the selected scales across different offender samples. People often develop firm masculine or feminine identifications that contribute to consistent "sex differences" or gender differences on psychometric tests. Outpatient Screen is a risk assessment instrument that measures risk from a variety of perspectives, notably, risk of alcohol and drug abuse, violence, control and mental health. If sex differences exist in these areas then male and female respondents are likely to score differently on these scales. This study also investigated sex differences in selected scales in one of the samples included in the study.

Method

The selected scales were administered to two different samples of offenders. The total number of offenders involved in the study (1995) was 1,821. Group 1 consisted of 611 offenders. There were 530 (86.7%) males and 81 (13.3%) females. The demographic composition of this sample is described as follows: Age: 16 to 20 (10.4%); 21 to 30 (44.9%); 31 to 40 (31.6%); 41 to 50 (10.5%) and 51 to 65 (2.7%). Ethnicity: Caucasian (41.6%); Black (57.0%); Hispanic (1.0%); Asian (0.3%); American Indian (0%) and Other (0.2%). Education: 8th Grade or less (3.3%); Some High School (29.2%); G.E.D. (4.8%); High School Graduate (51.8%); Some College (6.7%); Technical/ Business School (0.7%); College Graduate (3.1%) and Professional/Graduate School (0.5%). Marital Status: Single (48.6%); Married (38.1%); Divorced (8.0%); and Separated (5.2%).

Group 2 consisted of 1,210 domestic violence defendants (1,074 males and 136 females). The demographic composition of this sample is as follows: Age 19 and under (6.2%); 20 to 29 (42.6%); 30 to 39 (34.8%); 40 to 49 (12.8%); 50 to 59 (2.7%); 60 and older (0.8%). Ethnicity: Caucasian (50.2%); Black (42.7); Hispanic (5.2%); Asian (0.3%); American Indian (0.6%); and Other (0.7%). Education: 8th Grade or less (5%); Some High School (30%); G.E.D. (62%); High School Graduate (44.4%); Some College (9.3%); Technical/Business School (0.9%); College Graduate (3.3%); Professional/ Graduate School (0.5%). Marital Status: Single (44.5%); Married (38.9%); Divorced (8.8%) and Separated (1.8%).

Reliability coefficient alphas are presented in Table 9.

Table 9. Reliability coefficient alphas. Two sample of offenders.
All coefficient alphas are significant at p<.001. (1995, Total N = 1,821)

Outpatient Screen Scale	1 Offenders N = 611	2 Offenders N = 1,210
Truthfulness Scale	.87	.87
Alcohol Scale	.91	.90
Drug Scale	.89	.88
Anger Scale	.87	.88
Stress Management	.92	.93

This study strongly supports the reliability (internal consistency) of the selected scales. All coefficient alphas were statistically significant at p<.001.

Sex differences were investigated using offenders in Group 2. T-tests were calculated for the selected scales to evaluate possible gender differences. These results are presented in Table 10.

Significant sex differences were found on three scales, i.e., Alcohol Scale, Drug Scale and the Anger Scale. Significant sex differences were not observed in terms of the Truthfulness Scale or Stress Management Scale.

Table 10. Sex differences in group 2 offenders. (1995, N = 1,210)

Outpatient Screen Scales	Males Mean (N=1,074)	Females Mean (N=136)	T value	Significance Level
Alcohol Scale	8.27	6.20	3.23	p<.001
Drug Scale	5.62	4.17	2.74	p=.006
Anger Scale	8.22	7.27	1.99	p=.047

Based on this research, gender specific norms (or separate male and female scoring procedures) have been established in the scoring procedure for men and women on the Alcohol Scale, Drug Scale and Anger Scale. In general, males tend to make more straightforward admissions on these items then females. Gender and other demographic-related differences will continue to be explored in subsequent research.

20. Outpatient Screen Reliability Study in Four Samples of Clients

Four client samples were included in the present study (1999) to further investigate scale reliability in different offender samples. The groups represented domestic violence defendants from different geographical areas of the country, but the offender assessment programs were similar. The purpose of the present study (1999) was to investigate reliability of the Outpatient Screen in different client samples.

Method

The Outpatient Screen was administered to four groups of clients. There were a total of 841 offenders included in this study (1999). **Group 1** consisted of 306 clients. This sample included 267 (87.3%) males and 39 (12.7%) females. The demographic composition of Group 1 is as follows: Age: 16 to 20 years (3.9%), 21 to 25 (19.6%), 26 to 30 (24.5%), 31 to 35 (20.6%), 36 to 40 (18.3%), 41 to 45 (7.8%); 46 to 50 (3.9%), 51 to 55 (1.0%), and over 60 (0.3%). Ethnicity: Caucasian (70.9%), Black (22.2%), Hispanic (3.6%), Asian (1.3%), and Native American (2.0%). Education: 8th grade or less (2.0%), Some High School (22.9%), G.E.D. (13.4%), High School Graduate (37.6%), Some College (19.3%), College Graduate (1.6%), Technical/Business School (2.9%), and Professional/Graduate School (0.3%). Marital Status: Single (39.9%), Married (30.4%), Divorced (17.6%), Separated (11.8%), and Widowed (0.3%).

Group 2 consisted of 287 clients. There were 255 males (88.9%) and 32 females (11.1%). The demographic composition of this sample is as follows: Age: 16 to 20 years (6.3%), 21 to 25 (19.7%), 26 to 30 (28.5%), 31 to 35 (64; 22.5%), 36 to 40 (38; 13.4%), 41 to 45 (16; 5.6%), 46 to 50 (8; 2.8%) and 51 to 55 (1.1%). Ethnicity: Caucasian (72.8%), Black (20.9%), Hispanic (3.8), Asian (0%), American Indian (1.7%) and Other (0.7%). Education: 8th grade or less (1.8%), Some High School (19.4%), G.E.D. (11.7%), High School Graduate (40.6%), Some College (16.6%), College Graduates (5.3%), Technical/Business School (2.5%) and Professional/Graduate School (2.1%). Marital Status: Single (65.1%), Married (17.2%), Divorced (12.6%), Separated (4.5%) and Widowed (0.7%).

Group 3 consisted of 95 clients. There were 78 males (82.1%) and 17 females (17.9%). The demographic composition of this sample is as follows: Ethnicity: Caucasian (84.2%), Black (1.1%), Hispanic (11.6%), Asian (1.1%), and Other (2.1%). Education: 8th grade or less (4.2%), Some High School (20.0%), G.E.D. (11.6%), High School Graduate (36.8%), Some College (11.6%), Technical/Business School (6.3%), College Graduates (5.3%), and Professional/Graduate School (4.2%). Marital Status: Single (16.8%), Married (42.1%), Divorced (17.9%), Separated (22.1%), and Widowed (1.1%).

Group 4 consisted of 153 adjudicated clients. This sample contained 147 (96.1%) males and 6 (3.9%) females. The demographic composition of this sample is as follows: Age: 19 years and younger (2.6%), 20 to 29 years (45.1%), 30 to 39 years (36.5%), 40 to 49 years (11.8%) and 50 to 59 years (3.9%). Ethnicity: Caucasian (47.7%), Black (47.1%), Hispanic (1.3%), Native American (0.7%) and Other (3.3%). Education: 8th grade or less (1.3%), Some High School (35.3%), G.E.D. (3.9%), High School Graduates (45.8%), Some College (9.8%), College Graduate (3.3%) and Professional/Graduate School (0.7%). Marital Status: Single (52.9%), Married (33.3%), Divorced (9.2%), Separated (3.9%) and Widowed (0.7%).

Reliability coefficient alphas are presented in Table 11. The total number of clients included in the study was 841.

Table 11. Reliability coefficient alphas. 841 clients (1999)
All coefficient alphas are significant at p<.001.

Scale	1 DV Offenders N = 306	2 DV Offenders N = 287	3 DV Offenders N = 95	4 DV Offenders N = 153
Truthfulness	.85	.87	.86	.89
Alcohol Scale	.93	.93	.94	.89
Drug Scale	.88	.87	.92	.91
Anger Scale	.85	.87	.90	.85
Stress Management	.92	.90	.92	.91

These results support the internal consistency (reliability) of the selected scales. All coefficient alphas are significant at p<.001. The selected scales are demonstrated to be a reliable domestic violence offender assessment instrument in different offender samples.

21. Reliability of the Selected Scales in Large Samples of Clients

In 2000, two large client assessment programs were added to the database. A study (2000) was conducted to determine the reliability of the selected scales in these two new probationer samples. **The first group contained 1,209 clients.** There were 1,074 males (88.8%) and 135 females (11.2%). The demographic composition of this group is as follows: Age: Under 19 (6.2%), 20 to 29 (42.6%), 30 to 39 (34.9%), 40 to 49 (12.8%), 50 to 59 (2.7%), 60 and over (0.7%). Education: 8th grade or less (5.1%), Some High School (30.1%), G.E.D. (6.2%), High School Graduate (44.6%), Some College (9.3%), Technical/Business School (0.9%), College Graduate (3.3%), Professional/Graduate School (0.4%). Ethnicity: Caucasian (50.4%), Black (42.9%), Hispanic (5.2%), Asian (0.3%), Native American (0.6%), and Other (0.6%). Marital Status: Single (44.6%), Married (39.0%), Divorced (8.9%), and Separated (7.5%).

Group 2 consisted of 1,478 clients. The demographic composition of this group is as follows: Males (1,283; 86.8%); Females (195; 13.2%). Age: 19 years and younger (7.6%), 20 to 29 years (40.0%), 30 to 39 years (36.0%), 40 to 49 years (12.8%), 50 to 59 years (2.8%), 60 and over (0.8%). Ethnicity: Caucasian (35.1%), Black (62.7%), Hispanic (1.4%), Asian (0.3%), Native American (0.3%), and Other (0.2%). Education: 8th grade or less (5.8%), Some High School (36.0%), G.E.D. (4.1%), High School Graduates (39.0%), Some College (12.1%), Technical/Business School (0.7%), College Graduates (2.2%), Professional/Graduate School (0.2%). Marital Status: Single (46.8%), Married (35.1%), Divorced (10.2%), Separated (7.6%), and Widowed (0.2%).

Reliability coefficient alphas are presented in Table 12. There were a total of 2,687 clients included in the study.

Table 12. Reliability coefficient alphas. N = 2,687 clients (2000).
All coefficient alphas are significant at p<.001.

Scale	1 D.V. Offenders N = 1,209	2 D.V. Offenders N = 1,478
Truthfulness Scale	.85	.86
Alcohol Scale	.89	.88
Drug Scale	.86	.85
Anger Scale	.86	.85
Stress Management	.92	.90

These results support the internal consistency of the selected scales. Reliability refers to consistency of results regardless of who uses the instrument.

This study (2000, N=1,478) supports the reliability (internal consistency) of the selected scales.

22. Validity, Reliability and Accuracy of the Selected Scales

This study (2001) was conducted to test the validity, reliability and accuracy of the selected scales. Two statistical procedures were used in the present study to test the validity of scales. The first procedure involved t-test comparisons between first offenders and multiple offenders (discriminant validity) and the second procedure involved statistical decision-making (predictive validity). For the *t*-test comparisons, a first offender was defined as an offender who did not have a prior arrest and a multiple offender was defined as an offender who had one or more prior arrests. Several discriminant validity tests were conducted. Number of alcohol arrests was used to define first offenders and multiple offenders to test the Alcohol Scale. Similarly, number of drug arrests was used for the Drug Scale. Self-reported 'total number of arrests' was used to categorize offenders for other scale analyses. Because risk is often defined in terms of severity of problem behavior it is expected that multiple offenders would score significantly higher on the different scales than first offenders. This was an empirical question that was tested in the present study.

In assessment, a measurement can be considered a prediction. For example, the Alcohol Scale is a measure of alcohol abuse or severity of abuse. Alcohol Scale scores would predict if an individual has an alcohol problem. A benchmark that can be used for the existence of an alcohol problem is treatment. If an individual has been in alcohol treatment then the individual is known to have had an alcohol problem. Therefore, the Alcohol Scale should predict if an individual has been in treatment.

Statistical decision-making is closely related to predictive validity of a test. The quality of statistical decision-making and test validity are both assessed by the accuracy with which the test (Alcohol Scale) classifies *known* cases (treatment). In the present study predictive validity was evaluated in the selected scales by using contingency tables defined by scale scores and either treatment or desire for treatment.

Risk range percentile scores are calculated for each selected scale. These risk range percentile scores are derived from scoring equations based on responses to scale items, Truth-Corrections and prior criminal history information. These scores are then converted to percentile scores. There are four risk range categories: **Low Risk** (zero to 39th percentile), **Medium Risk** (40 to 69th percentile), **Problem Risk** (70 to 89th percentile) and **Severe Problem or Maximum Risk** (90 to 100th percentile). Risk range percentile scores represent degree of severity.

Analysis of the accuracy of risk range percentile scores involves comparing the risk range percentile scores obtained from client test results to the predicted risk range percentages as defined above. The percentages of clients expected to fall into each risk range is the following: Low Risk **(39%)**, Medium Risk **(30%)**, Problem Risk **(20%)** and Severe Problem or Maximum Risk **(11%)**. The actual percentage of probationers falling in each of the four risk ranges, based on their risk range percentile scores, was compared to these predicted percentages.

Method

There were three client samples used in the study. The total number of participants was 5,122. **Group 1 consisted of 604 offenders.** There were 521 males (86.3%) and 83 females (13.7%). The demographic composition of this sample is as follows: Age: 19 and under (4%), 20 - 29 (34.3%), 30 - 39 (38.2%), 40 - 49 (16.6%), 50 - 59 (5%) and 60 and over (2%). Ethnicity: Caucasian (60.9%), Black (21.7%), Hispanic (15.9%), Asian (0.2%), Native American (0.5%) and Other (0.8%). Education: 8th grade or less (11.1%), Some High School (3.3%), G.E.D. (5%), High School graduate (37.6%), Some college (10.9%), Technical/Business school (0.5%), College graduate (2.3%) and Professional/Graduate school (0.2%). Marital Status: Single (47%), Married (33.8%), Divorced (12.3%), Separated (6.6%) and Widowed (0.2%).

Group 2 consisted of 1,239 offenders. There were 1,068 males (86.2%) and 171 females (13.8%). The demographic composition of this sample is as follows: Age: 19 and under (5%), 20 - 29 (36.7%), 30 - 39 (39.9%), 40 - 49 (14.9%), 50 - 59 (2.9%) and 60 and over (0.6%). Ethnicity: Caucasian (48.8%), Black (47.2%), Hispanic (2.7%), Asian (0.4%), Native American (0.3%) and Other (0.7%). Education: 8th grade or less (8%), Some High School (29.1%), G.E.D. (4%), High School graduate (42.3%), Some college (13%), Technical/Business school (0.2%), College graduate (3.3%) and Professional/Graduate school (0.1%). Marital Status: Single (43.1%), Married (39.9%), Divorced (10.8%), Separated (6%) and Widowed (0.3%).

Group 3 consisted of 3,279 offenders. There were 2,786 males (85%) and 493 females (15%). The demographic composition of this sample is as follows: Age: 19 and under (5.4%), 20 - 29 (36.2%), 30 - 39 (37.8%), 40 - 49 (16.2%), 50 - 59 (3.5%) and 60 and over (0.9%). Ethnicity: Caucasian (70.7%), Black (7.9%), Hispanic (13.6%), Asian (1%), Native American (2.6%) and Other (4.1%). Education: 8th grade or less (4.8%), Some High School (27.7%), G.E.D. (7.1%), High School graduate (39.8%), Some college (14.4%), Technical/Business school (2%), College graduate (3.5%) and Professional/Graduate school (0.6%). Marital Status: Single (39%), Married (40.2%), Divorced (13.2%), Separated (7.2%) and Widowed (0.4%).

Reliability coefficient alphas for the three groups are presented in Table 13. There were a total of 5,122 offenders included in the study (2001).

Table 13. Reliability coefficient alphas in three adult offender samples. (2001, Total N=5,122)
All coefficient alphas are significant at p<.001.

Scale	1 Offenders N = 604	2 Offenders N = 1,239	3 Offenders N = 3,279
Truthfulness Scale	.87	.88	.88
Alcohol Scale	.94	.93	.94
Drug Scale	.91	.90	.92
Anger Scale	.89	.89	.90
Stress Management	.93	.92	.93

These results support the reliability of the selected scales. All coefficient alphas were significant at p<.001. All coefficient alphas for selected scales are very highly significant.

T-tests were calculated for the five scales to assess possible sex differences in the three offender groups. These results are presented in Table 14.

Table 14. T-test comparisons of sex differences. (2001)
Offender Sex Differences (Total N = 5,122)

SCALE	Group 1 604 Offenders	Group 2 1,239 Offenders	Group 3 3,279 Offenders
Truthfulness Scale	n.s.	n.s.	n.s.
Alcohol Scale	t=3.83, p<.001	t=3.13, p=.002	t=6.11, p<.001
Drug Scale	n.s.	n.s.	t=3.12, p=.002
Anger Scale	t=4.80, p<.001	t=2.83, p=.005	t=10.96, p<.001
Stress Management	n.s.	t=3.16, p=.002	n.s.

Significant sex differences were demonstrated on two of the scales, i.e., Alcohol and Violence, in all groups. Significant sex differences were found on the Stress Management Scale in Group 2. Significant sex differences were found on the Drug Scale in Group 3. Based on this (2001) study, gender-specific norms have been established in the computerized scoring software for men and women on the Alcohol, Violence, Drug and Stress Management scales. This is an example of the value of ongoing Outpatient Screen research. With more accurate and fair measures, assessment personnel can be more confident in their assessment-related decisions.

The risk range percentile scores for Group 3 are presented in Table 15.

Table 15. Risk Range Percentile Scores for Group 3, N = 3,279 offenders.

Risk Range	Truthful-ness	Alcohol	Drug	Violence	Stress Coping	**Predicted**
Low	39.1	41.8	37.6	38.3	39.2	**39%**
Medium	31.4	27.9	30.9	30.3	29.7	**30%**
Problem	18.2	19.4	20.8	20.2	20.0	**20%**
Maximum	11.3	10.9	10.7	11.2	11.1	**11%**

These results show that obtained risk range percentile scores closely approximated the predicted risk range percentile scores for each of the selected scales presented in Table 20. The results of the comparisons between attained risk percentages and predicted percentages for Group 3 shows that all obtained scale risk range percentile scores were within 2.8 percent of predicted. The largest difference between obtained and predicted risk range percentages occurred on the Low Risk category. For the Problem Risk and Severe Problem Risk categories, all but two comparisons showed that the obtained percentages were within one percentage point of predicted. **This embodies accurate risk assessment.**

The t-test comparisons of first offenders' and multiple offenders' scale scores are presented in Tables 16 through 18. There were 3,279 clients used in this analysis.

Table 16. T-test comparisons between first offenders and multiple offenders.
Offender status defined by total number of arrests. (N = 3,279)

Scale	First Offenders Mean (N=1,251)	Multiple Offenders Mean (N=2,028)	T-value	Level of significance
Truthfulness Scale	9.08	7.27	t = 9.08	p<.001
Anger Scale	16.93	30.95	t = 28.76	p<.001
Stress Management	109.54	104.36	t = 3.46	p=.001

Table 17. T-test comparison of Alcohol Scale between first offenders and multiple offenders.
Offender status defined by number of alcohol arrests.

Scale	First Offenders Mean (N=2,454)	Multiple Offenders Mean (N=825)	T-value	Level of significance
Alcohol Scale	6.92	22.62	t = 31.85	p<.001

Table 18. T-test comparison of Drug Scale between first offenders and multiple offenders.
Offender status defined by number of drug arrests.

Scale	First Offenders Mean (N=3,110)	Multiple Offenders Mean (N=169)	T-value	Level of significance
Drug Scale	4.97	21.37	t = 16.39	p<.001

These t-test results support the discriminant validity of the selected scales. All t-test comparisons of first offenders and multiple offenders were significant on the Alcohol, Drug, Violence and Stress Management scales. The Truthfulness Scale showed that first offenders scored significantly higher than multiple offenders. The mean scale score on the Stress Management Scale indicated that first offenders had higher scores on average (better stress handling abilities) than multiple offenders.

T-test results of the Anger Scale indicated that multiple offenders scored much higher than first offenders. The very large significant difference between first and multiple offenders strongly support the discriminant validity of the Anger Scale. T-test results of the Alcohol Scale and Drug Scale, where offender status was defined by alcohol arrests and drug arrests, respectively, also showed very large significant differences between first and multiple offenders. These results strongly support the discriminant validity of the Alcohol Scale, Drug Scale and Anger Scale.

The test of predictive validity for the Alcohol Scale is presented in Table 19. Defendants who scored between the 40th and 69th percentile are not included in the table because the table distinguishes between problem and no problem behavior. No problem is defined as an Alcohol Scale score at or below the 39th percentile, whereas alcohol-related problematic behavior is defined as an Alcohol Scale score in the 70th or above percentile range. Alcohol treatment information was obtained from offenders' answers to test items concerning treatment or desire for treatment.

Table 19. Predictive validity for the Alcohol Scale using scale scores and alcohol treatment.

	Alcohol Treatment		
Alcohol Scale	No AA and No Desire	Attends AA or Desires Treatment	Number in each category
Low Risk (zero to 39th percentile)	1,362 (.83)	8 (.01)	1,370
Problem or Severe Problem Risk (70 to 100th percentile)	284 (.17)	709 (.99)	993
	1,646	717	N = 2,363

These results show that for the 717 offenders who reported having attended AA or who desired alcohol treatment, 709 offenders, or 99 percent, had Alcohol Scale scores at or above the 70th percentile. Similarly, of the 1,646 offenders who reported no AA or no desire for alcohol treatment, 1,362 offenders or 76 percent had Alcohol Scale scores in the Low Risk or no problem range. This percentage is reasonable because offenders could have a drinking problem without having been in treatment. Combining these results gives an overall accuracy of the Alcohol Scale of **88 percent**. This is very accurate considering that a highly accepted diagnostic procedure, the mammogram, is about 70 percent accurate. These results show there is a very strong positive correlation between Alcohol Scale scores and alcohol treatment.

The predictive validity test of the Drug Scale was done in the same way using drug treatment as the criterion. **Of the 207 offenders who desired drug treatment 207 or 100 percent had Drug Scale scores in the 70th percentile or higher (Problem Risk and above). Of the 2,429 offenders who did not have treatment 1,604 (66%) had Drug Scale scores in the Low Risk (no problem) range.** The overall accuracy of the Drug Scale in predicting drug treatment was **69 percent**. These results show there is a very strong positive correlation between the Drug Scale and drug treatment.

For the Anger Scale, 86 percent of the offenders who desired domestic violence treatment, had Anger Scale scores at or above the 70th percentile and the overall accuracy was 78 percent. This means that there is a very strong positive correlation between Anger Scale scores and desire for violence treatment.

23. A Study of Validity, Reliability and Accuracy of Selected Scale in Five Samples of Offenders

This study (2002) was conducted to further test the validity, reliability and accuracy of the selected scales in different samples. The study replicates the statistical procedures of reliability, validity and accuracy that was presented in earlier research studies.

Method

There were five client samples used in this study (2002). The total number of participants was 7,905. **Group 1 consisted of 903 clients.** There were 757 males (83.8%) and 146 females (16.2%). The demographic composition of this sample is as follows: Age: 19 and under (4.9%), 20 - 29 (38.5%), 30 - 39 (35.4%), 40 - 49 (15.4%), 50 - 59 (4%) and 60 and over (1.4%). Ethnicity: Caucasian (57.4%), Black (34.4%), Hispanic (4.7%), Asian (0.8%), Native American (0.4%) and Other (0.8%). Education: 8th grade or less (11%), Some High School (27.4%), G.E.D. (5.1%), High School graduate (38.2%), Some college (12.5%), Technical/Business school (0.1%), College graduate (3.7%) and Professional/Graduate school (1.1%). Marital Status: Single (45.7%), Married (35.7%), Divorced (11.2%), Separated (5.3%) and Widowed (0.2%).

Group 2 consisted of 1,157 clients. There were 989 males (85.5%) and 168 females (14.5%). The demographic composition of this sample is as follows: Age: 19 and under (4.1%), 20 - 29 (41.1%), 30 - 39 (34.8%), 40 - 49 (15.5%), 50 - 59 (3.1%) and 60 and over (0.8%). Ethnicity: Caucasian (50.4%), Black (3.2%), Hispanic (17%), Asian (1%), Native American (1.9%) and Other (15%). Education: 8th grade or less (3.8%), Some High School (20.8%), G.E.D. (10.6%), High School graduate (40.4%), Some college (15.5%), Technical/Business school (1.5%), College graduate (3.2%) and Professional/Graduate school (0.4%). Marital Status: Single (44.4%), Married (34.7%), Divorced (10.1%), Separated (3.5%) and Widowed (0.3%).

Group 3 consisted of 1,626 clients. There were 1,396 males (85.9%) and 230 females (14.1%). The demographic composition of this sample is as follows: Age: 19 and under (5.7%), 20 - 29 (34.1%), 30 - 39 (37.5%), 40 - 49 (16.5%), 50 - 59 (4.8%) and 60 and over (1.3%). Ethnicity: Caucasian (69.8%), Black (9.2%), Hispanic (10.8%), Asian (1%), Native American (2.6%) and Other (2.5%). Education: 8th grade or less (5.6%), Some High School (25.3%), G.E.D. (7.4%), High School graduate (40.2%), Some college (13.7%), Technical/Business school (1.7%), College graduate (3.1%) and Professional/Graduate school (0.9%). Marital Status: Single (41.7%), Married (35.9%), Divorced (13.5%), Separated (6.3%) and Widowed (0.2%).

Group 4 consisted of 3,190 clients. There were 2,690 males (84.3%) and 500 females (15.7%). The demographic composition of this sample is as follows: Age: 19 and under (5.2%), 20 - 29 (36.7%), 30 - 39 (35.3%), 40 - 49 (17.5%), 50 - 59 (4.1%) and 60 and over (1.1%). Ethnicity: Caucasian (69%), Black (14.3%), Hispanic (11.2%), Asian (0.9%), Native American (2.4%) and Other (2.4%). Education: 8th grade or less (5.8%), Some High School (25.4%), G.E.D. (8.0%), High School graduate (39.4%), Some college (14.1%), Technical/Business school (1.8%), College graduate (4.7%) and Professional/Graduate school (0.7%). Marital Status: Single (41%), Married (39.6%), Divorced (12.5%), Separated (6.5%) and Widowed (0.5%).

Group 5 consisted of 1,029 clients. There were 919 males (89.3%) and 110 females (10.7%). The demographic composition of this sample is as follows: Age: 19 and under (4.4%), 20 - 29 (40.2%), 30 - 39 (37.2%), 40 - 49 (15%), 50 - 59 (2.6%) and 60 and over (0.6%). Ethnicity: Caucasian (69%), Black (15.3%), Hispanic (9.7%), Asian (0.4%), Native American (3.4%) and Other (2.1%). Education: 8th grade or less (4%), Some High School (24.4%), G.E.D. (9.6%), High School graduate (40.7%), Some college (14.9%), Technical/Business school (3.2%), College graduate (2.8%) and Professional/Graduate school (0.4%). Marital Status: Single (44.3%), Married (31.5%), Divorced (15.8%), Separated (8.1%) and Widowed (0.3%).

Accuracy of the Selected Scales

Scale scores are classified according to the following four risk range categories: **Low Risk** (zero to 39th percentile or 39% of the clients), **Medium Risk** (40 to 69th percentile or 30%), **Problem Risk** (70 to 89th percentile or 20%) and **Severe Problem or Maximum Risk** (90 to 100th percentile or 11%).

Risk range percentile scores represent degree of severity. Accuracy of the Outpatient Screen scale scores is determined by the closeness of obtained scores to the predicted risk range percentages as defined above. The actual percentages of offenders falling in each of the four risk ranges are presented in the graph and table below.

Table 20. Risk Range Percentile Scores for Group 5, N = 1,029 offenders.

	Truthfulness	Alcohol	Drug	Violence	Stress Mgmnt	**Predicted**
Risk Range	%	%	%	%	%	**%**
Low	39.7	37.1	39.0	38.0	38.7	**39%**
Medium	29.3	32.2	28.0	31.9	30.2	**30%**
Problem	21.1	19.0	22.5	19.5	19.8	**20%**
Severe Problem	9.9	11.7	10.5	10.6	11.3	**11%**

As shown in the above, the obtained risk range percentages for all risk categories and all scales were within **2.5** percentage points of the predicted risk ranges.

Reliability of Selected Scales

Reliability coefficient alphas for the five groups are presented in Table 21. There were a total of 7,905 clients included in the study (2002).

Table 21. Reliability coefficient alphas in five adult offender samples. (Total N=7,905)
All coefficient alphas are significant at p<.001.

Scale	**1 Offenders N = 903**	**2 Offenders N = 1,157**	**3 Offenders N = 1,626**	**4 Offenders N = 3,190**	**5 Offenders N = 1,029**
Truthfulness Scale	0.89	0.87	0.88	0.88	0.87
Alcohol Scale	0.93	0.93	0.94	0.94	0.95
Drug Scale	0.87	0.90	0.91	0.91	0.92
Anger Scale	0.90	0.90	0.91	0.90	0.90
Stress Management	0.90	0.94	0.94	0.93	0.93

These results support the reliability of the selected scales. All coefficient alphas were significant at p<.001. All coefficient alphas for the selected scales are well above the generally accepted level of 0.80 for assessment tests.

Validity of the Selected Scales

T-test comparisons between first offenders and multiple offenders are presented in Tables 22 to 24 for offenders in Group 5. A first offender was defined as an offender who did not have a prior arrest and a multiple offender one or more prior arrests. Several discriminant validity tests were conducted. There were 1,029 clients used in this analysis.

Table 22. T-test comparisons between first offenders and multiple offenders in Group 5. Offender status defined by number of domestic violence arrests. (2002, N=1,029)

Scale	First Offenders Mean (N=790)	Multiple Offenders Mean (N=239)	T-value	Level of significance
Truthfulness Scale	7.80	6.55	t = 3.55	p<.001
Anger Scale	24.64	39.74	t = 13.09	p<.001
Stress Management	110.36	99.53	t = 3.63	p<.001

Table 23. T-test comparison of Alcohol Scale between first offenders and multiple offenders. Offender status defined by number of alcohol arrests.

Scale	First Offenders Mean (N=741)	Multiple Offenders Mean (N=288)	T-value	Level of significance
Alcohol Scale	8.04	22.48	t = 15.49	p<.001

Table 24. T-test comparison of Drug Scale between first offenders and multiple offenders. Offender status defined by number of drug arrests.

Scale	First Offenders Mean (N=946)	Multiple Offenders Mean (N=83)	T-value	Level of significance
Drug Scale	11.25	19.46	t = 6.84	p<.001

The Violence and Stress Management Scales accurately differentiated between first offenders and multiple offenders. These results show that having domestic violence arrests is associated with having higher severity levels for control, violence and stress problems. **These t-test results strongly support the discriminant validity of the Violence and Stress Management Scales.**

The Truthfulness Scale shows that first offenders score higher than multiple offenders. There appears to be a trend in offender assessment where first time offenders try to fake good more often than multiple offenders. This finding has been found in the other tests as well. The Alcohol Scale and Drug Scale accurately differentiated between multiple offenders and first offenders. **These results strongly support the discriminant validity of the Alcohol Scale and Drug Scale.**

The **predictive validity** analysis shows that the Alcohol Scale accurately identified offenders who have alcohol problems. Those offenders who have been in alcohol treatment or desire treatment were identified as having alcohol problems.

As shown in Table 25, offenders who reported having been in alcohol treatment or desired treatment, 226 offenders, or 100 percent, had Alcohol Scale scores at or above the 70th percentile. Nearly 100 percent of the clients who had alcohol treatment scored in the Problem or Severe Problem risk range on the Alcohol Scale. The Alcohol Scale was accurate in identifying clients with known alcohol problems.

Table 25. Predictive validity of the Alcohol Scale using scale scores and alcohol treatment.

	Alcohol Treatment		
Alcohol Scale	No Treatment or desire	Treatment or desire treatment	Number in each category
Low Risk (zero to 39th percentile)	363 (83%)	1 (0%)	364 (55%)
Problem or Severe Problem Risk (70 to 100th percentile)	75 (17%)	226 (100%)	301 (45%)
	438 (66%)	227 (34%)	N = 665

363 of the 438 offenders (83%) who reported no alcohol treatment had Alcohol Scale scores in the Low Risk or no problem range. 589 (226 + 363) of the 665 offenders gives an overall accuracy of the Alcohol Scale of **89 percent**. This is very accurate assessment. These results show that the Alcohol Scale accurately identified alcohol problems.

The Drug Scale accurately identified offenders who have drug problems. Using drug treatment responses, it was determined that **158 of the 186 offenders (85%) who reported having been in drug treatment or desired treatment had Drug Scale scores in the Problem Risk range and above.**

24. Scale Validity and Accuracy in a Large Sample of Clients

This study (2003) investigated validity and accuracy of the selected scales in a large sample of offenders. There were 7,941 offenders included in this study. These clients were tested in a variety of testing milieus throughout the US and Canada. These include counseling agencies, treatment centers, community corrections, probation and judicial centers.

Method and Results

There were 7,941 clients included in this study (2000). There were 6,565 males (82.7%) and 1,376 females (17.3%). Demographic composition of these participants is as follows: Age: 19 & under (6%); 20-29 (36%); 30-39 (35%); 40-49 (19%); 50-59 (4%) and 60 & Over (1%). Ethnicity: Caucasian (63%); Black (19%), Hispanic (13%) and Other (5%). Education: Eighth grade or less (7%); Some H.S. (27%); H.S. graduate/GED (47%); Some college (14%) and College graduate (4%). Marital Status: Single (44%); Married (36%); Divorced (12%); Separate (7%) and Widowed (1%).

The court-history information for these participants is as follows: Age of first conviction: 15 & under (12%); 16-20 (34%); 21-25 (19%); 26-30 (12%); 31-35 (9%); 36-40 (7%); 41-45 (4%); 46-50 (2%); 51 & over (2%). Misdemeanor convictions: None (26%); One (25%); Two (17%); Three (11%); Four (6%); Five or more (14%). Felony convictions: None (73%); One (16%); Two (6%); Three (2%); Four (1%);

Five or more (2%). Times on probation: None (33%); One (34%); Two (19%); Three (7%); Four (3%); Five or more (3%). Probation revocations: None (83%); One (11%); Two (3%); Three (1%); Four (1%); Five or more (1%). Times on parole: None (91%); One (7%); Two (1%); Three or more (1%). Parole revocations: None (95%); One (3%); Two (1%); Three or more (2%). Total number of times arrested: None (11%); One (24%); Two (19%); Three (13%); Four (9%); Five or more (25%). Times sentenced to jail: None (50%); One (22%); Two (12%); Three (6%); Four (4%); Five or more (7%). Times sentenced to prison: None (89%); One (8%); Two (2%); Three (1%); Four (1%); Five or more (1%). Years incarcerated: None (84%); One (7%); Two (3%); Three (2%); Four (1%); Five or more (4%). Domestic violence arrests: None (26%); One (52%); Two (14%); Three (4%); Four (1%); Five or more (2%). Alcohol arrests: None (55%); One (22%); Two (10%); Three (5%); Four (2%); Five or more (6%). Drug arrests: None (83%); One (11%); Two (3%); Three (1%); Four (1%); Five or more (2%). Assault arrests: None (73%); One (16%); Two (4%); Three (1%); Four (1%); Five or more (5%).

Accuracy of the Outpatient Screen

Participant scale scores are classified according to the risk (degree of severity) they represent. Four categories of risk are assigned: **Low risk** (zero to 39th percentile), **Medium risk** (40 to 69th percentile), **Problem risk** (70 to 89th percentile), and **Severe Problem** (90 to 100th percentile). By definition the expected percentage of participants assigned to each risk category is, 39% in Low risk, 30% in Medium risk, 20% in Problem risk and 11% in Severe Problem. The actual percentages of participants placed in the four risk categories based on their scale scores are compared to these expected percentages. Table 26 presents these comparisons. The differences between obtained and expected are shown in parentheses.

Table 26. Risk Range Percentile Scores, N = 7,941 offenders.

Scale	Low Risk (39%)		Medium Risk (30%)		Problem Risk (20%)		Severe Problem (11%)	
Truthfulness Scale	39.5	(0.5)	30.9	(0.9)	19.4	(0.6)	10.2	(0.8)
Alcohol Scale	38.9	(0.1)	30.5	(0.5)	20.0	(0.0)	10.6	(0.4)
Drugs Scale	40.6	(1.6)	30.5	(0.5)	18.6	(1.4)	10.3	(0.7)
Anger Scale	38.0	(1.0)	30.1	(0.1)	20.7	(0.7)	11.1	(0.1)
Stress Management	39.1	(0.1)	30.0	(0.0)	20.0	(0.0)	10.9	(0.1)

As shown in the graph and table above, scale scores are accurate. The objectively obtained percentages of participants falling in each risk range are very close to the expected percentages for each risk category.

For those participants who are identified as having problems (Problem and Severe Problem risk ranges or 31% of the participants), the obtained percentages were also accurate. The problem risk ranges for all scales are in close agreement to the expected percentage. These results demonstrate that scale scores accurately identify domestic violence risk.

Reliability of the Selected Scales

Table 27. Reliability coefficient alphas (2003, Total N = 7,941).
All coefficient alphas are significant at p<.001.

Scale	Alpha
Truthfulness Scale	.88
Alcohol Scale	.93
Drugs Scale	.91
Anger Scale	.90
Stress Management	.93

These results support the statistical reliability of the selected scales. Reliability coefficients were well above the generally accepted level (0.75) for tests.

Validity of the Selected Scales

Two different statistical procedures are presented that demonstrate the validity of the Outpatient Screen. The first validation procedure compares first offenders and multiple offenders (discriminant validity). Multiple offenders are defined as offenders who have two or more domestic violence arrests. Because risk of domestic violence is defined in terms of severity of risk it is expected that multiple offenders would attain significantly higher scale scores than first-time offenders.

T-test comparisons were used to study the statistical significance between first and multiple offenders. There were 6,255 first offenders and 1,686 multiple offenders (2 or more arrests). These results are presented in Table 28.

Table 28. T-test comparisons between first offenders and multiple offenders. (2003, N=7,941)

Scale	First Offenders Mean	Multiple Offenders Mean	T-value	Significance
Truthfulness Scale	9.15	8.01	t =7.65	p<.001
Alcohol Scale	8.17	13.72	t = 16.11	p<.001
Drugs Scale	4.60	6.64	t = 8.70	p<.001
Anger Scale	21.42	35.77	t = 35.26	p<.001
Stress Management	110.67	99.58	t = 10.22	p<.001

Note: The Stress Management Scale is reversed in that the higher the score the better one copes with stress.

These results show that multiple offenders score significantly higher on the Alcohol, Drugs, Anger Scales and Stress Management Scales than first offenders. These results support the discriminant validity of the Alcohol, Drugs, Violence and Stress Management Scales. The Truthfulness Scale shows that first offenders score significantly higher than multiple offenders. Results on the Truthfulness Scale suggest that first offenders may try to fake good, whereas multiple offenders see no reason to further deny their problems. These results strongly support the discriminant validity of the Outpatient Screen.

The second validity procedure studied the accuracy at which the Alcohol and Drugs scales identified problem drinkers and drug abusers. To be considered accurate a client test must accurately identify problem clients (drinkers or drug abusers). The criterion in this analysis for identifying offenders as problem drinkers or drug abusers is having been in treatment (alcohol or drug). Having been in treatment identifies offenders as having had an alcohol or drug problem. If a person has never had an alcohol or drug problem it is very likely they have not been treated for an alcohol or drug problem. Thus, offenders are

separated into two groups, those who had treatment and those who have not had treatment. Then, offender scores on the Alcohol and Drug Scales are compared. It is predicted that offenders with an alcohol and/or drug treatment history will score in the problem risk range (70th percentile and above) on the Alcohol and/or Drug Scales. Non-problem is defined in terms of low risk scores (39th percentile and below) on the Alcohol and/or Drug Scales. Substance abuse treatment information is obtained from offender answers to scale items regarding alcohol and drug treatment.

Predictive validity analysis shows that Alcohol and Drug Scales accurately identify offenders who have had alcohol and/or drug treatment. The Alcohol Scale is very accurate in identifying clients who have alcohol problems. There were 1,382 offenders who reported having been in alcohol treatment and these offenders are classified as problem drinkers. Of these 1,382 offenders, 1,370 individuals, or 99.1 percent, had Alcohol Scale scores at or above the 70th percentile. **The Alcohol Scale correctly identified nearly all of the offenders categorized as problem drinkers.** This is very accurate assessment. These results validate the Alcohol Scale.

Similar results were found for the Drugs Scale. There were 1,337 offenders who reported having been in drug treatment. All 1,337 individuals, or 100 percent, had Drug Scale scores at or above the 70th percentile. **These results strongly support the validity of the Drugs Scale.**

Gender Differences

Possible male/female scale score differences were investigated in this study and these results are shown in Table 29.

Table 29. T-test comparisons between males and females. (2003, N=7,941)

Scale	Males Mean	Females Mean	T-value	Significance
Truthfulness Scale	9.04	8.29	t =4.50	p<.001
Alcohol Scale	9.88	6.82	t = 10.37	p<.001
Control Scale	8.57	9.18	t = 2.93	p=.003
Drugs Scale	5.19	4.26	t = 4.10	p<.001
Anger Scale	25.42	19.98	t = 13.61	p<.001
Stress Management	109.43	102.99	t = 5.23	p<.001

Note: The Stress Management Scale is reversed in that the higher the score the better one copes with stress.

These results demonstrate significant male/female differences on all scales. The Truthfulness, Alcohol, Drugs and Anger Scales show that males score significantly higher than females, whereas the Stress Management Scale scores show that females attain significantly higher scores than males. These results indicate that separate scoring procedures are needed for males and females for fair and accurate risk assessment. Accurate assessment must take into account differences between males and females patterns of responding to scale items.

Discussion

The participants in this study were clients taken from a variety of testing milieus. There were 7,941 offenders included in this study from different areas around the US and Canada. With such a diverse sample of clients these results have wide applicability. The majority of the offenders (82.7%) were male and most (78.3%) were first time clients.

These results strongly support the reliability, validity and accuracy of the tested scales. Scale scores were within 2.1 percent of predicted percentages for all four risk range classification categories.

Reliability coefficients for all tested scales were at or above 0.90. All coefficients were significant at $p<0.001$. The results of two validity studies validate the scales. Scale score comparisons of first and multiple offenders show the scores significantly differentiate between first and multiple offenders. Multiple offenders score significantly higher than first offenders on the Alcohol, Drugs and Anger Scales. Furthermore, the Alcohol Scale accurately identified 99.1 percent of problem drinkers and the Drugs Scale accurately identified 100 percent of problem drug users.

25. A Study of Selected Scales in a Sample of Probationers

This study (2003) included clients being tested in a statewide probation department offender assessment program. Statistical reliability, validity and accuracy of the selected scales were examined. There were 833 offenders included in this study.

Method and Results

There were 833 clients included in this study (2003). There were 737 males (88.5%) and 96 females (11.5%). Demographic composition of these participants is as follows: Age: 19 & under (5.8%); 20-29 (37%); 30-39 (38.5%); 40-49 (16%); 50-59 (2.4%) and 60 & Over (0.4%). Ethnicity: Caucasian (67.4%); Black (17.6%), Hispanic (9.3%), Native American (2.9%) and Other (2.7%). Education: Eighth grade or less (2.7%); Some H.S. (23.8%); H.S. graduate/GED (49.5%); Some college (20.2%) and College graduate (3.7%). Marital Status: Single (48.3%); Married (28.2%); Divorced (15.1%); Separated (8.1%) and Widowed (0.2%).

Reliability coefficient alphas are presented in Table 30.

Table 30. Reliability coefficient alphas (2003, Total N = 833).
All coefficient alphas are significant at p<.001.

Scale	Alpha
Truthfulness Scale	.87
Alcohol Scale	.95
Drugs Scale	.91
Anger Scale	.89
Stress Management	.94

These results are in close agreement to those found in other studies reported above. The Outpatient Screen achieved high statistical reliability. All scale reliability coefficients were at or above .90.

Accuracy of the Selected Scales

The percentages of offenders classified in the four risk ranges based on their scale scores are presented in Table 31. The differences between obtained and expected percentages are shown in parentheses.

Table 31. Risk Range Percentile Scores (2000, N = 833)

Scale	Low Risk (39%)		Medium Risk (30%)		Problem Risk (20%)		Severe Problem (11%)	
Truthfulness	41.4	**(2.4)**	29.8	**(0.2)**	18.0	**(2.0)**	10.8	**(0.2)**
Alcohol	37.1	**(1.9)**	31.7	**(1.7)**	20.7	**(0.7)**	10.5	**(0.5)**
Drug	37.6	**(1.4)**	30.8	**(0.8)**	20.6	**(0.6)**	11.0	**(0.0)**
Violence	38.8	**(0.2)**	30.3	**(0.3)**	19.4	**(0.6)**	11.5	**(0.5)**
Stress Management	38.9	**(0.1)**	29.9	**(0.1)**	20.5	**(0.5)**	10.7	**(0.3)**

As shown in the above graph and table, the obtained risk range percentages for all risk categories and all scales were within 2.0 percentage points of the predicted risk ranges.

Validity of the Selected Scales

In the discriminant validity analyses "Number of domestic violence arrests," "Number of alcohol arrests" and "Number of drug arrests" were used to define first offenders and multiple offenders. There were 226 domestic violence multiple offenders, 256 alcohol multiple offenders and 67 drug multiple offenders.

Table 32. T-test comparisons between first offenders and multiple offenders. (2000, N = 833)

Outpatient Screen Scale	**First Offenders Mean Score**	**Multiple Offenders Mean Score**	**T-value**	**Level of significance**
Truthfulness Scale	8.81	7.63	t = 3.10	p=.002
Alcohol Scale *	7.43	21.67	t = 15.98	p<.001
Drugs Scale *	4.95	15.63	t = 7.66	p<.001
Anger Scale	25.27	39.87	t = 13.86	p<.001
Stress Management	108.82	98.96	t = 3.08	p=.002

*Offender status defined by alcohol arrests and drug arrests. Stress Management scores are reversed in that higher scores mean better Stress Management.

These results demonstrate that multiple offenders scored significantly higher on the Alcohol, Drugs, Violence and Stress Management Scales than did first offenders. These scales accurately differentiated between first offenders and multiple offenders. These results support the discriminant validity of the Alcohol, Drugs, Violence and Stress Management scales. There are very large scale score differences on the Alcohol, Drugs and Anger Scales between first and multiple offenders. These scales clearly indicate that multiple offenders are at risk in comparison to first offenders.

The Truthfulness Scale shows that first offenders attained significantly higher scores than multiple offenders. This result has been found in previous studies. First offenders seemingly try to minimize their problems more often than multiple offenders. Multiple offenders appear to be more experienced and know their histories are well documented by the probation department. The Truthfulness Scale has been validated in previous research studies.

The second validity procedure studied the accuracy at which the Alcohol Scale and Drugs Scale identified problem drinkers and drug abusers. See the previous study for a discussion on this analysis. Offenders who have been in alcohol or drug treatment are predicted to score in the problem risk ranges (70^{th} percentile and above) on the Alcohol and Drugs Scales. The predictive validity analysis shows that the Alcohol and Drugs Scales accurately identified offenders who have alcohol or drug problems.

Treatment information was obtained from offenders' answers to specific items concerning alcohol and drug treatment. These analyses compared offenders who scored in the problem risk ranges (70th percentile and above) with offenders who scored in the low risk range (39th percentile and below).

Of the 191 offenders who reported having been in alcohol treatment or desired treatment, 190 offenders, or 99.5 percent, had Alcohol Scale scores at or above the 70th percentile. Nearly 100 percent of the offenders who had alcohol treatment scored in the problem risk ranges on the Alcohol Scale. These results validate the Alcohol Scale. Of the 168 offenders who reported having been in drug treatment or desired treatment 163 or 97 percent had Drugs Scale scores in the problem risk ranges. These results validate the Drugs Scale.

Discussion

Results of these statewide probation department offenders were consistent with the general population clients. These results strongly support the reliability, validity and accuracy of the examined scales. All scale scores were within 2.0 percent of predicted percentages for all four risk range classification categories. Reliability coefficients for all scales were at or above 0.90. All coefficients were significant at $p<0.001$. Scales significantly differentiate between first and multiple offenders. Multiple offenders score significantly higher than first offenders on the Alcohol, Drugs, Violence and Stress Management Scales. Furthermore, the Alcohol Scale accurately identified 99.5 percent of problem drinkers and the Drugs Scale accurately identified 97 percent of problem drug abusers. The scales studies are accurate, reliable and valid.

26. Study of Selected Scales with a Large Sample of Domestic Violence Offenders

This study (2010) summarizes results for **10,676** adult domestic violence offenders. Offenders were tested during the time-period beginning in January 2003 and ending in March 2010. Outpatient Screen test data was gathered online. The purpose of this study was to evaluate the validity of the Outpatient Screen by analyzing results from the 10,676 Outpatient Screen test administrations.

Method

There were 10,676 clients included in this study (2010). There were 8,187 males (76.7%) and 2,489 females (23.3%). The demographic composition of this sample is as follows: Age: 19 and under (11.8%), 20-29 (38.3%), 30-39 (30.5%), 40-49 (15.7%), 51-60 (3.4%) and 61 and over (0.3%). Ethnicity: Caucasian (71.7%); African American (9.7%); Hispanic (14.4%); Asian (2.0%); Native American (1.4%); Other (0.8%). Education: 8th grade or less (5.5%); some High School (26.6%); GED (9.5%) High School Graduate (34.9%); Some College (14.4%); Trade/Technical School (2.9%); College Graduate (5.1%); Advanced Degree (1.3%) Marital Status: Single or Never Married (46.1%); Married (31.9%); Divorced (12.8%) Separated (8.8%); Widowed (0.5%).

Validity

For the following validity analyses, the prediction criterion was offender status. By comparing the scale scores of First and Multiple Offenders, the analyses examined whether test scales could distinguish between offenders with known different levels of problem severity. In the first analysis, it was predicted that *Violent* Multiple Offenders (two or more domestic violence or two or more assault arrests) would obtain significantly higher Anger Scale and Stress Management Scale scores than *Violent* First Offenders (one or no domestic violence or general assault arrests). *T*-test results (presented in Table 59) demonstrated that Violent Multiple Offenders did indeed score significantly higher than Violent First Offenders on all three scales.

Table 33. *T*-test Comparisons between *Violent* First and Multiple Offenders (N=10,676[a], 2010)

Scale	Mean Scores First Offenders	Mean Scores Multiple Offenders	*t*-value	Cohen's *d* (effect size)
Truthfulness	8.20	7.04	4.13	*d*=0.49
Violence	26.68	54.32	-11.75	*d*=1.62
Stress Management	113.58	99.20	7.00	*d*=0.34

*Small effect; **Medium effect; ***Large effect

In the next two analyses, it was predicted that *Alcohol* Multiple Offenders (two or more alcohol-related arrests) would obtain significantly higher Alcohol Scale scores than *Alcohol* First Offenders (one or no alcohol-related arrests), and that *Drugs* Multiple Offenders (two or more drug-related arrests) would obtain significantly higher Drugs Scale scores than *Drugs* First Offenders (one or no drug-related arrests). Again, predictions were confirmed. (See Table 34 for *t*-test results).

Table 34. *T*-test Comparisons of *Alcohol* & *Drug* First and Multiple Offenders (N=10,676[a], 2010)

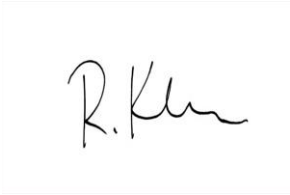
Scale	Mean Scores First Offenders	Mean Scores Multiple Offenders	*t*-value	Cohen's *d* (effect size)
Truthfulness	8.28	7.13	5.77	*d*=0.21
Alcohol	7.40	21.00	-51.16	*d*=1.14
Drugs	6.83	20.55	-38.48	*d*=1.23

*Small effect; **Medium effect; ***Large effect

These results demonstrate that the selected scales do accurately measure levels of severity. The scales effectively differentiate between offenders who are known to have more severe problems (Multiple Offenders) and First Offenders. In terms of the Truthfulness Scale, offenders with multiple arrests attained significantly lower mean scores than first-time offenders. This is possibly due to the Multiple Offenders' presumed familiarity with the court system and the checking of records. Multiple Offenders may have increased awareness of the ineffectiveness of denial and problem minimization, especially in court-related settings.

SUMMARY

This document "Outpatient Screen: An Inventory of Scientific Findings" summarizes many research studies supporting the reliability, validity and accuracy of the scales used in the Outpatient Screen. Based on research presented herein, it is reasonable to conclude the Outpatient Screen test will provide a sound empirical basis for responsible referral and decision-making. Annual program summary will provide program self-evaluation in the future.



Numa Khandaker
Research Analyst
Behavior Data Systems, Ltd.

* * *