ER

Employee Review

An Inventory of Scientific Findings

TABLE OF CONTENTS

Preface

Employee Review (ER) research and development began in 1980 and has continued to the present. The ER is designed for accurate, inexpensive and timely on-site employee screening. The copyrighted ER database ensures continued research and development. The ER is a brief, easily administered and automated (computer scored) test that is designed for employee screening. It includes true/false and multiple choice items and can be completed in 25 minutes. The ER has been standardized on college students, job employees, substance abuse clients, probationers, vocational rehabilitation participants and others.

This document summarizes much of the validity and reliability research that contributed to ER development. The ER has demonstrated reliability, validity and accuracy. It correlates impressively with both experienced staff judgment and other recognized tests. ER tests can be given directly on the computer screen or in paper-pencil test booklet format. All tests are computer scored on-site. ER reports are available within two minutes of test completion. Diskettes contain all of the software needed to score tests, build a database and print reports. By merging the latest psychometrics with computer technology the ER accurately assesses employee behavior and identifies employee risk as well as need. Employment staff can now objectively gather a vast amount of relevant information, identify employee strengths and formulate specific remediation strategies.

The ER identifies employee problems like growing negativity, disgruntled attitude, working under extreme stress (impaired stress management skills) or substance (alcohol and other drugs) abuse problems. When obstacles are identified, specific interventions or ways to deal with these problems are offered. Objective assessment and specific problem-related recommendations can help. When employers can offer corrective intervention to help employees keep their jobs employee turnover is often reduced and kept to a minimum. For these reasons ER research is ongoing, so that we can provide staff with the most accurate data possible.

The ER represents years of research and development, and integrates the latest psychometric procedures with computerized technology to provide a state-of-the-art employment screening instrument. The copyrighted ER database ensures ongoing research. And, on an annual basis the ER is essentially restandardized on a state-by-state basis.

This document describes the ER and gathers together research into one source. Its purpose is to provide understanding of the ER and the automated screening system it represents. ER research is ongoing, consequently no attempt was made to summarize all ER research. This document represents the evolution of the ER into a state-of-the-art employment assessment or screening instrument. It is emphasized that current studies are most representative of the present or revised ER.

Information on the Employee Review (ER) is available in the **ER Orientation & Training Manual**. Computer scoring information is contained in the **ER Computer Operating Guide**. Each of these manuals can be obtained upon request.

Behavior Data Systems, Ltd.
P.O. Box 44256, Phoenix, Arizona 85064-4256

EMPLOYEE REVIEW (ER)

Employee Review (ER) research and development began in 1980 and has continued to the present. Initially large item pools were collected for each ER scale by a group of psychologists and counselors involved in employment selection and screening. Subsequently, these item pools were administered to job employees and the items with the best statistical properties were retained. Final scale-item configurations were administered to job employees, substance (alcohol and other drugs) abusers, college students and other population samples in a series of reliability and validity studies. The ER's proven research continues to deliver the highest quality in job employee screening. And, the proprietary ER database ensures ongoing research and development.

The Employee Review (ER) is designed for employee screening. The ER has a fifth to sixth grade reading level, and requires 20 to 25 minutes to complete. It contains eight scales: Truthfulness, Alcohol, Drugs, Interpersonal Skills, Work Orientation, Aggressiveness, Security and Stress Coping Abilities. These eight scales represent important areas of employee screening--many of which are missed by other testing procedures. The ER is appropriate for adult employment screening.

<div align="center">

EMPLOYEE REVIEW
<u>MEASURES OR SCALES</u>
1. Truthfulness Scale
2. Alcohol Scale
3. Drugs Scale
4. Interpersonal Skills Scale
5. Work Orientation Scale
6. Aggressiveness Scale
7. Security Scale
8. Stress Management Scale

</div>

The ER is designed for accurate, inexpensive and timely on-site employee screening. It is an objective employment screening instrument designed to identify problem prone employees, or conversely, to recognize problem free individuals. The ER can be administered on a computer (IBM-PC compatibles) screen or by using paper-pencil test booklets. Regardless of how the ER is administered, all tests are scored and interpreted with a computer which generates ER reports.

The ER requires approximately 25 minutes for completion and is appropriate for adult males and females. The ER is composed of True-False and multiple-choice items. It can be administered individually or in groups. The language is direct, non-offensive and uncomplicated. Automated scoring and interpretive procedures help insure objectivity and accuracy. The ER is to be used in conjunction with a review of available records, a focused interview and experienced staff judgment.

How do you measure attitudes, predict complex behavior and establish employee risk? The answer: You use a computer! By merging the latest psychometrics with computer technology, the ER can accurately assess employee risk and needs. Staff can now objectively gather a vast amount of important information and identify employee problems so they can be worked through. The speed, accuracy and reliability of computers greatly increase employee screening efficiency.

TRUTHFULNESS SCALE: determines how truthful the employee was while completing the test. This scale identifies guarded, defensive or recalcitrant employees who minimize or deny problems and concerns. This scale identifies faking.

TRUTH-CORRECTED SCORES: A sophisticated psychometric technique permitted by computer technology involves "truth corrected" scores which are calculated individually for each of the eight ER scales each time a test is scored. Since it would be naive to assume everybody responds truthfully while completing any test or interview, the Truthfulness Scale was developed. **The Truthfulness Scale establishes how honest or truthful the employee is while completing the ER.** Correlations between the Truthfulness Scale and all other ER scales permit identification of error variance associated with untruthfulness. This error variance is then added back into the scale score, resulting in more accurate "Truth-Corrected" scores. Unidentified denial or untruthfulness produces inaccurate and distorted results. Raw scores may only reflect what the employee wants you to know. **Truth-Corrected scores reveal what the employee is trying to hide.** Truth-Corrected scores are more accurate than raw scores.

This procedure permits identification of faking, malingering and falsification of answers. Available research indicates that many job employees try to minimize their problems and concerns when applying for jobs. The ER detects these guarded, recalcitrant and defensive employees.

Risk Range Percentile Scores: Each ER scale is scored independently of the other scales. ER scale scoring equations combines client's pattern of responding to scale items and Truthfulness Scale scores. 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 percentile scores which are reported in the ER report.

ER 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. Percentile scale scores are derived from "obtained scale scores." The cumulative distributions of truth-corrected scale scores determines cut-off scores for each of the four risk range or severity categories. Individual scale score calculations are automatically performed and results are presented in the ER report numerically (percentile), by attained risk category (narrative) and graphically (ER profile).

DATABASE: Every time an ER is scored the test data is automatically stored on the diskette for inclusion in the ER database. This applies to ER diskettes used anywhere in the United States and Canada. When the preset number of tests are administered (or used up) on an ER diskette, the diskette is returned for replacement and the test data contained on these used diskettes is input, in a confidential (no names) manner, into the ER database for later analysis. This database is statistically analyzed annually, at which time future ER diskettes are adjusted to 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 enable employee assessment review and planning.

Confidentiality (Delete Client Names): Many agencies and programs are rightfully concerned about protecting their client's confidentiality. The proprietary Delete Client Names option is provided to allow deletion of client names from test diskettes prior to their being returned. This is optional and once the names have been deleted they are gone and cannot be retrieved. Deleting client names does not delete demographic information or test data. It only deletes the client names when the option is used. The option is available at any time and can be used whether the diskette is full or not. Once client names are deleted they are gone and cannot be recovered. **This ensures client confidentiality.**

The ER is a self-report test that can be completed individually or in group testing settings. There are no forms or questionnaires to be completed by the staff. ER reports eliminate the need for tedious, time consuming and error prone hand scoring. Specific problem identification can cut the waste associated with over-evaluation and expensive drug tests. The Employment Interview provides accurate information for use with a "focused" interview. Problems are identified so the interview can "zero in" or focus on topics of concern. Then, on an as needed or desired basis, individual employees can be interviewed, and in some cases when problems are identified, the employee can be referred for additional tests.

DESCRIPTION OF EMPIRICALLY BASED MEASURES OR SCALES

ER scales were developed from large item pools. Initial item selection was a rational process based upon clearly understood definitions of each scale. 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 ER familiarize themselves with the definition of each scale. For that purpose a description of each ER scale follows.

Truthfulness Scale: This scale provides a measure of the employee's truthfulness while completing the ER. As noted earlier, all interview and self-report tests are subject to the dangers of untrue answers. The Truthfulness Scale identifies these self-protective employees. As noted earlier, the Truthfulness Scale enables calculation of Truth-Corrected scores, which are more accurate than raw scores. A Truthfulness Scale is considered necessary - if not essential - to any self-report questionnaire or test.

Since the outcome of an employee's test score could affect their employment status, it would be naive to believe that job employees answer all questions truthfully. Many attempt to minimize their problems and concerns.

Alcohol Scale: This scale measures an employee's alcohol proneness and alcohol-related problems. Frequency and magnitude of alcohol use or abuse are important screening factors. Alcohol is a major licit or legal substance. Alcoholism is a significant problem in our society. Woolfolk and Richardson noted in their book "Stress, Sanity and Survival" that alcoholism costs industry over $15.6 billion annually due to absenteeism and medical expenses. In the new millennium, these expenses are much higher. Experienced staff are aware of alcoholics' job performance problems.

Drugs Scale: This empirically based scale is an independent measure of employee's drug use and abuse-related problems. The burgeoning awareness of the impact of illicit drugs on employees emphasizes the

need for a employee test to discriminate between licit and illicit drugs. Without a drug scale, many drug abusers would remain undetected. Increased public awareness of illicit drug (marijuana, cocaine, crack, amphetamines, barbiturates and heroin) abuse emphasizes the importance of including an independent measure of drug use or abuse.

The national outcry in the 1980's concerning cocaine use momentarily obscured the fact that Americans also abuse a number of other substances, including marijuana. Marijuana or pot has "come back" in the 1990's and new millennium. Marijuana can be an intoxicant, depressant, hallucinogen, stimulant, or all of the above. The principal mind-altering ingredient in marijuana (THC) may linger for days or even weeks. Studies have shown that THC intoxication can return, for no apparent reason, even when a person has not recently smoked marijuana.

Interpersonal Skills Scale: This scale measures how well the employee gets along with others. This scale incorporates friendliness, cooperation, interpersonal awareness and one's people orientation. Many attribute as much as 80 percent of job turnover to "people problems."

Cooperativeness is defined in terms of an employee's willingness to positively relate and work with others. It incorporates communication, acceptance, mutual assistance and affiliation motivation. On the other side of the coin, this scale identifies negative interpersonal relationship, negative attitudes toward authority and a high propensity towards "people problems."

Work Orientation Scale: This scale is an adjustment and work appraisal measure. It incorporates the employee's attitudes, employment history and overall adjustment. Work orientation describes an employee's attitudes towards work, work attendance, and problem free employment.

The Work Orientation Scale lends itself to incorporating various objective criteria, e.g., accidents, tardiness, sick leave, insurance payouts, early quit, grievance time, absenteeism, personal time off, and misconduct in future research.

Aggressiveness Scale: This scale measures aggressive and outgoing behavior. Aggressive people are often described as pushy and controlling.

Aggressiveness is a behavior consisting of strong self-assertiveness, social dominance and a tendency toward hostility. Extreme aggressiveness has been called "aggressive conduct disorder" and is characterized by domineering, punitive, or assaultive verbal or physical conduct. Aggressive types are now listed as subgroups of the antisocial personality. More descriptive information can be found in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).

Security Scale: This scale measures endangerment, unsocial attitudes and court-related history. Violence includes assaultive behavior. Unsocial behavior incorporates moral and ethical blunting. Court-related history incorporates violent crimes, substance-related offenses and misdemeanor convictions as well as probation. The Security Scale is a checklist of prior unsocial, violent or court-related offenses.

Stress Management Scale: This scale is a measure of the employee's ability to cope with stress. Stress exacerbates other symptoms of emotional problems. Seriously impaired stress coping abilities are usually associated with other identifiable emotional and psychological problems.

Stress is an increasingly significant concept in our society. The National Institute for Occupational Safety and Health (NIOSH) evaluated the health records of 22,000 workers in 130 organizations. Their conclusion: stress affects workers in all types of jobs at all levels. Unskilled laborers are equally susceptible as are top-line executives.

How effectively individuals cope with stress determines whether or not stress is a significant factor in their lives. Two concepts, "stress" and "coping abilities" dominate the literature on stress. The Stress Management Scale includes measures of both of these concepts in its Stress Quotient (SQ) equation. The better an individual's coping skills, compared to their amount of experienced stress, the higher the SQ score. In contrast, if an individual is experiencing more stress than he or she can cope with, the lower the SQ score. In the ER profile, Stress Quotient (SQ) scores were inverted to conform to the established risk levels ranging from low to high risk categories.

Stress exacerbates other symptoms of emotional, attitudinal, interpersonal and substance abuse related problems. Frequency and magnitude of impaired stress coping abilities 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.

ER DEVELOPMENT

ER scales 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, large item pools were collected for each ER scale by a group of psychologists and counselors involved in employment selection and screening. Subsequently, these item pools were administered to job employees and the items with the best statistical properties were retained. Final scale-item configurations were administered to job employees, substance (alcohol and other drugs) abusers, college students and other population samples in a series of reliability and validity studies. Thus, the ER has been researched, normed and validated on job employees.

The ER provides employers with two kinds of information--qualitative and quantitative. Qualitative information is obtained by self-report items that reflect the employees' opinions, perceptions and beliefs. Quantitative information is obtained by eight empirically based measures or scales. Scale development began in 1980 and ER research continues today. ER language is direct, non-offensive and uncomplicated.

RESEARCH STUDIES

The ER has been researched and normed on the job employee population. Reliability refers to consistency of results regardless of who uses the instrument. ER results are objective, verifiable and reproducible. Validity refers to a test measuring what it is purported to measure. The ER was validated in a series of studies that are summarized in this document. However, it should be emphasized that ER research is ongoing in nature.

The research which follows has been included in a chronological manner, so that the reader can observe the development of the ER into a state-of-the-art job employee assessment instrument or test. More recent studies (represented at the end of this document) are most representative of the ER's present or current statistical properties. For example the database research summarized on page 28 (1999, N = 3,513) demonstrates the ER's present statistical properties.

STRESS QUOTIENT

The Stress Quotient (SQ) 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 scores 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 coping abilities. 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 correlated -.70 with the Taylor Manifest Anxiety Scale and -.75 with the Cornell Index. Both correlation's 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 coping abilities. 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 correlation's were in predicted directions. The significant correlation's 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 coping abilities 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 correlation's 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 Wiggin's 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 correlation's 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) Reliability of the Stress Quotient (SQ) or Stress Management Scale was investigated in a sample of 189 job employees. There were 120 males and 69 females with an average age of 31. The SQ was administered at the time of employee 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 correlation's 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 coping abilities.

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 coping abilities.

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 coping abilities. 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**.

ER RESEARCH

The ER has a long history of research and development, much of which is contained in the following summary. **ER research is reported in a chronological format, reporting studies as they occurred.** This gives the reader the opportunity to see how the ER evolved into a state-of-the-art job employee assessment instrument. However, for current information refer to the more recent studies near the end of this research document.

10. A Study of ER Test-Retest Reliability

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 fair. Psychometric principles and computer technology insures ER accuracy, objectivity, practicality and cost-effectiveness.

Reliability is a measure of the consistency of a test in obtaining similar results upon re-administration of the test. One measure of test reliability, over time, is the test-retest correlation coefficient. In this type of study, the test is administered to a group and then the same test is re-administered to the same group at a later date.

Method
College students at two different colleges enrolled in introductory psychology classes participated in this study (1984). A total of 115 students participated and received class credit for their participation. The students were administered the ER in a paper-pencil test format. One week later they were re-tested with the ER again.

Results
The results of this study revealed a significant test-retest product-moment correlation coefficient of $r = 0.71$, $p<.01$. These results support the reliability of the ER. Test-retest consistency was very high and indicates that the ER scores are reproducible and reliable over a one week interval.

In another test-retest study (1985), the ER was administered on two occasions to the same people. Seventy outpatients were re-tested with the ER after a ten-day interval. The Pearson Product Moment Correlation Coefficient of .87 was highly significant. This study again demonstrates that the ER is a reliable employee screening instrument.

11. Validation of the Truthfulness Scale

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

The Truthfulness Scale identifies respondents who are 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 21 Truthfulness Scale items in the Employee Review 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 ER Truthfulness Scale, was administered to the subjects and the Truthfulness Scale was embedded in the test as one of the five scales. Truthfulness Scale scores were made up of the number of deviant answers given to the 21 Truthfulness Scale items.

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 reveal that the Truthfulness Scale accurately detects "Fakers" from those students that took the test honestly.

12. Validation of Six Employee Review 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, the six Employee Review scales (Truthfulness, Alcohol, Drugs, Interpersonal Skills, Work Orientation and Stress Coping Abilities) 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 ER 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 Psychopathic Deviant scales. The Drug Scale was validated with the MacAndrew and Psychopathic Deviant scales. The Interpersonal Skills Scale was validated with the Taylor Manifest Anxiety, Social Maladjustment and Social Alienation scales on the MMPI. The Work Orientation Scale was validated with the Manifest Hostility and Authority Conflict scales. The Stress Management Scale was validated with the Taylor Manifest Anxiety, Psychasthenia, Social Maladjustment and Social Alienation scales.

Content validity is a measure of how well test items (or scales) measure the factor they were designed to measure. As noted earlier, a large item pool was rationally developed for ER consideration. Consensual agreement among three psychologists and experienced personnel/selection staff familiar with ER scale definitions markedly reduced the initial item pool. Final item selection was empirical and based on each item's statistical properties. Selected items had acceptable reliability coefficients and correlated highest with their respective scales. The ER was then objectively normed and standardized.

Predictive validity measures how well a test can predict behavior the test was designed to measure. The Truthfulness Scale is important as these scores establish how truthful the respondent was while completing the ER. Truthfulness scale scores determine whether or not ER profiles are valid, and are integral to the calculation of truth-corrected scores.

Concurrent validity (criterion-related validity) correlates the scales of the test being validated with similar scales or measures from an established test which has demonstrated reliability and validity. This was done in the following 1985 (N = 100) study.

Method
One hundred (100) vocational rehabilitation clients (1985) were administered both the ER and the MMPI. Tests were counterbalanced for order effects -- half were given the ER first and half the MMPI first.

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

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 ER 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 correlation's 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 **Drugs Scale** correlates significantly with the MacAndrew (r = 0.62) Scale and the Psychopathic Deviant (r = 0.54) Scale.

Table 1. (1985) Product-moment correlations (N = 100)
between MMPI scales and Employee Review scales

MMPI SCALES (MEASURES)	Employee Review Scales (Measures)					
	Truthfulness	Alcohol	Drugs	Interpersonal Skills	Work Orientation	Stress Coping
L (Lie) Scale	0.72	-0.38	-0.41	-0.28	-0.29	0.53
Psychopathic Deviant	-0.37	0.52	0.54	0.35	0.27	-0.59
Psychasthenia	-0.34	0.38	0.41	0.28	0.37	-0.68
Social Maladjustment	-0.25	0.34	0.26	0.18	0.35	-0.54
Authority Conflict	-0.43	0.31	0.47	0.37	0.55	-0.46
Manifest Hostility	-0.45	0.34	0.47	0.37	0.57	-0.58
Taylor Manifest Anxiety	-0.58	0.47	0.46	0.48	0.50	-0.78
MacAndrew	-0.40	0.58	0.62	0.44	0.26	-0.33
Authority Problems	-0.32	0.36	0.42	0.35	0.18	-0.22
Social Alienation	-0.47	0.35	0.45	0.28	0.48	-0.67

NOTE: All correlations were significant at p < .001.

The **Interpersonal Skills Scale** is most highly correlated with the Taylor Manifest Anxiety (r = 0.48) scale on the MMPI. The Interpersonal Skills Scale also significantly correlates with the Social Maladjustment Scale and Social Alienation Scale on the MMPI.

The **Work Orientation Scale** is most significantly correlated with the Manifest Hostility (r = 0.57) and the Authority Conflict (r = 0.55) MMPI scales. These findings are consistent with the conceptual description of the Work Orientation Scale that was cited earlier.

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 coping Ability 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 Employee Review scales. All of the ER scales were highly correlated with the MMPI criterion scale they were tested against. The large correlation coefficients support the validity of the ER. All product-moment correlation coefficients testing the relation between ER scales and MMPI scales were significant at the p < .001 level.

13. Inter-item Reliability of the Employee Review

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 measures 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. Coefficient Alpha results are demonstrated in this study (1985, N = 389).

Method
This study (1985) included three separate groups of subjects: 100 outpatients in private practice, 100 substance abuse inpatients, and 189 job employees -- totaling 389 subjects. Separate inter-item reliability analyses were conducted to compare results across the three groups.

Results and Discussion
The inter-item reliability coefficient alpha and within-test reliability statistics are presented in Tables 2 and 3, respectively. All inter-item reliability coefficient alphas and within-test reliability F-values are significant at p<.001. These results support the reliability of the ER. The ER is a highly reliable instrument.

Table 2. Inter-item reliability, coefficient alpha. (1985)
Outpatients, Substance Abuse Inpatients and Job Employees (N = 389)

ER SCALES MEASURES	N ITEMS	Outpatients (N = 100)	Inpatients (N = 100)	Job Employees (N = 189)
Truthfulness Scale	21	0.81	0.79	0.81
Alcohol Scale	21	0.86	0.93	0.83
Drugs Scale	21	0.80	0.85	0.79
Interpersonal Skills	21	0.63	0.72	0.70
Work Orientation	21	0.74	0.74	0.78

13

Stress Coping Abilities	40	0.81	0.84	0.73

Table 3. Within-test reliability, F statistic.

<u>**All F statistics are significant at p<.001.**</u>

ER SCALES <u>MEASURES</u>	N <u>ITEMS</u>	Outpatients <u>(N = 100)</u>	Inpatients <u>(N = 100)</u>	Job Employees <u>(N = 189)</u>
Truthfulness Scale	21	21.73	53.15	45.91
Alcohol Scale	21	9.29	31.46	47.75
Drugs Scale	21	27.19	16.34	58.18
Interpersonal Skills	21	26.97	17.05	48.26
Work Orientation	21	15.97	19.21	28.67
Stress Coping Abilities	40	46.74	16.20	195.86

These results (Tables 2 and 3) demonstrate the impressive reliability of the ER. Reliability was demonstrated with three different groups of people (outpatients, inpatients and job employees) taking the ER. In each of these subject samples, all ER scales (measures) were found to be significantly independent of the other ER scales as shown by the highly significant within-test F statistics. The F statistic is obtained in within-subjects between measures ANOVA performed on each individual ER scale in each of the samples. The F statistics show that each ER scale measures essentially one factor (or trait). In addition, all ER scales show high inter-item reliability. This is demonstrated by the Standardized Cronbach's Coefficient Alpha - a widely used test of inter-item reliability when using parallel models. This measure reveals that all items in each ER scale are significantly related and measure just one factor. In other words, each ER scale measures one factor, yet the factor being measured is different from scale to scale. The inter-item reliability coefficients show very similar results across the three subject samples. The Truthfulness Scale, Alcohol Scale and Drugs Scale are in close agreement. The Stress Management Scale shows similar results for the chemical dependency groups but the job employee group had a slightly lower coefficient alpha. This difference might be accounted for by the fact that individuals applying for a job would not want to show themselves in a bad light by indicating they have an emotional, stress-related or mental health problem.

Because each sample may have scored differently from the other two samples, the data for all subjects were combined. For example, job employees may score low on the Alcohol and Drugs Scales and inpatient clients may score high. By combining the data, scale scores would likely be distributed from low to high and result in even better coefficient alphas than each sample separately. Table 4 presents the inter-item reliability analysis of all of these independent studies (N = 100, N = 100, N = 189) combined (N = 389).

Table 4. Inter-item reliability, coefficient alpha. All data combined (N = 389).

ER SCALES <u>MEASURES</u>	N <u>ITEMS</u>	COEFFICIENT <u>ALPHA</u>	F <u>VALUE</u>
Truthfulness Scale	21	0.82	96.93
Alcohol Scale	21	0.94	26.68
Drugs Scale	21	0.88	79.71
Interpersonal Skills	21	0.74	40.80
Work Orientation	21	0.77	53.03
Stress Coping Abilities	40	0.85	150.78

The combined data shows that all but one coefficient alpha increased in the combined data compared to coefficient alphas of each subject sample alone. These coefficient alphas in the combined data are very high and provide strong support for the reliability of the ER.

14. Relationships between Selected ER 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 Employee Review scales were chosen for this study: Truthfulness Scale, Alcohol Scale and Drugs Scale. The Truthfulness Scale was chosen because it is used in the ER to measure the truthfulness or honesty of the respondent while completing the ER. The Alcohol and Drugs Scales are well suited for comparison with the polygraph exam because of the situation specific nature of the scales. Alcohol and drug 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 the Truthfulness Scale items as used in the ER. The respondent's attitude, emotional stability and tendencies to fake good affect the Truthfulness Scale. It was expected that the Alcohol and Drugs Scales would be highly correlated with the polygraph results and the Truthfulness Scale would show a somewhat less but nonetheless significant correlation. The following study (1985, N = 189) demonstrates the nature of these polygraph-ER relationships.

Method
One hundred and eighty-nine (189) job employees (1985) were administered both the ER scales and the Polygraph examination. Tests were given in a counterbalanced order, half of the employees were given the ER scales first and the other half of the employees were administered the polygraph first. The subjects were administered the ER 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 ER scales demonstrated 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 Drugs Scale ($r = 0.56$, $p<.001$).

In summary, this study supports the validity of the ER Truthfulness Scale, Alcohol Scale and Drugs Scale. There were strong positive relationships between the selected ER scales and the Polygraph examination. The highly significant product-moment correlations between ER scales and Polygraph examinations demonstrate the validity of the ER Truthfulness, Alcohol and Drugs 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 ER. The fact that there was a very strong relationship between Polygraph results and ER scales shows that this type of information can be obtained accurately in self-report instruments.

These results indicate that the ER Truthfulness Scale is an accurate measure of the respondent's truthfulness or honesty while completing the ER. The Truthfulness Scale is an essential measure in self-report instruments. There must be a means to determine the honesty or "correctness" of the respondent's answers and there must be a means to adjust scores when the respondent is less than honest. The ER 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. The results of this study show that the ER is a valid assessment instrument.

15. Replication Study of ER Reliability

After this first sequence of studies, ER database research enabled statistical reliability analysis of each scale item. The original pool of ER scale items was statistically analyzed and only the items with the best statistical properties (item-whole correlation coefficients) were retained. Weak items were deleted and replaced in the revised ER. The following study (1987, N = 192) investigated the reliability of the revised ER.

This study (1987) was conducted to test the reliability (internal consistency) of the revised ER scales. In a replication of earlier ER research, the ER was administered to 192 chemical dependency inpatients. Age ranged from 18 to 56 years. This ER study (1987) is summarized in Table 5.

Table 5. Inter-item Reliability (1987, N = 192)
Chemical Dependency Inpatients

ER Scales	Coefficient Alpha	P< Value
Truthfulness	.79	.001
Alcohol Scale	.99	.001
Drugs Scale	.87	.001
Interpersonal Skills	.81	.001
Work Orientation	.81	.001
Stress Coping Abilities	.99	.001

The results of this study demonstrate the reliability (internal consistency) of the revised version of the ER. The reliability coefficients for the Interpersonal Skills and Work Orientation scales improved. All other scales retained their high reliability statistics. These results strongly support the reliability of the revised ER.

16. Validation of the ER in a Sample of Vocational Rehabilitation Clients

Reliability and validity of the revised ER scales continued to be studied in a sample of vocational rehabilitation clients. Consistently high reliability statistics of the ER have been found. Yet with the changes in scale items it was important to validate these changed scales. Vocational rehabilitation clients participated in this study; however, not all clients completed the MMPI that was used in the validation of the ER. This 1990 study involved 294 vocational rehabilitation clients.

Method
The participants in this study (1990) were vocational rehabilitation clients. 294 participants completed the ER and 171 completed the MMPI. The demographic composition of this sample was as follows: 203 (69

16

percent) males, and 91 (31 percent) females. Age: 16 to 25 years (71 males, 16 females); 26 to 35 years (93 males, 42 females); 36 to 45 years (32 males, 17 females); and 46 to 55 years (7 males, 16 females). Ethnicity: Caucasian (55 males, 32 females); Black (130 males, 58 females); Hispanic (9 males); Native American (7 males); and other (2 males, 1 female). Education: 8th grade or less (13 males, 1 female); some High School (43 males, 19 females); GED (16 males, 7 females); High School graduates (83 males, 24 females); some College (26 males, 21 females); Business/Technical School (1 male, 1 female); College graduates (13 males, 15 females); and Graduate/Professional degrees (8 males, 3 females). Reliability coefficient alphas are presented in Table 6. There were 294 participants in this study (1990).

These results support the reliability of the ER. **All scale reliability coefficient alphas were above the 0.80 level.** All coefficient alphas are significant at the p<.001 level of significance. The ER is a reliability assessment instrument for screening vocational rehabilitation clients.

Table 6. Reliability Coefficient Alphas (1990, N=294)
<u>Vocational Rehabilitation Clients</u>

<u>ER Scales</u>	**Cronbach** <u>**Alpha**</u>	**Significance** <u>**Level**</u>
Truthfulness	.83	.001
Alcohol Scale	.86	.001
Drugs Scale	.87	.001
Interpersonal Skills	.80	.001
Work Orientation	.81	.001
Stress Coping Abilities	.93	.001

In the validation part of this study (1990) there were 171 participants. Gender composition was 129 males and 42 females. This sample is described as follows. Age: Under 17 years (2); 18 to 21 (20); 22 to 25 (25); 26 to 29 (27); 30 to 33 (24); 34 to 37 (22); 38 to 41 (17); 42 to 45 (13); 46 to 49 (5); 50 to 53 (8); over 54 (8). Education: 8th grade or less (20); Partially completed High School (43); GED (16); High School Graduate (53); Some College (36); and College Graduates (3). The results of this study are summarized in Table 7.

As found in earlier studies, all ER-MMPI correlations were significant and in predicted directions. These empirical findings further support the validity of the ER. ER-MMPI results are presented in Table 7.

The ER Truthfulness Scale was invented and correlates most significantly with the MMPI-L Scale and MMPI-K scale. The ER Alcohol Scale correlates most significantly with the MMPI Psychopathic Deviant scale, MMPI Social Maladjustment scale, and MMPI Social Alienation scale. The ER Drug Scale correlates most significantly with the MMPI F scale, MMPI Psychopathic Deviant scale, and MMPI TSC-V (Resentment) scale. The ER Interpersonal Skills Scale correlates most significantly with the MMPI Taylor Manifest Anxiety scale, MMPI Psychopathic Deviant scale, MMPI Manifest Hostility scale, MMPI TSC-V (Resentment) scale and MMPI Social Alienation scale. The ER Work Orientation Scale correlates most significantly with the MMPI F scale, MMPI TSC-III (Suspiciousness) scale and the MMPI Social Alienation scale. The ER Stress Management Scale correlates most significantly with the MMPI F scale, MMPI Psychopathic Deviant scale, MMPI Psychasthenia scale, MMPI Taylor Manifest Anxiety scale, and MMPI Social Alienation scale.

MMPI SCALES (MEASURES)	Employee Review Scales (Measures)					
	Truthfulness	**Alcohol**	**Drugs**	**Interpersonal Skills**	**Work Orientation**	**Stress Coping**
L (Lie) Scale	.511**	.022	-.186*	-.034	.089	-.065
F (Validity)	-293**	.379**	.269**	.323**	.276**	.462**
K (Validity)	.458**	-.201*	-.151	-.340**	-.077	-.319**
Psychopathic Deviant	.241**	.312**	.190*	.300**	.065	.491**
Psychasthenia	-.279**	.202*	.115	.142	.069	.470**
Taylor Manifest Anxiety	-.394**	.288**	.151	.259**	.031	.536**
MacAndrew	.005	.051	.090	.188*	.127	.076
Social Maladjustment	-.335**	.273**	.174	.139	.033	.329**
Manifest Hostility	-.465**	.197*	.159	.380**	.176	.266**
TSC-III (Suspiciousness)	-.373**	.195*	.061	.292**	.209*	.247**
TSC-V (Resentment)	-.457**	.322**	.195*	.404**	.140	.402**
Social Alienation	-.377**	.283**	.171	.311**	.249**	.447**

Significance Level: ** $p<.001$, * $p<.01$

Similar ER-MMPI correlations were demonstrated earlier. The present study further supports the validity of the ER. The ER measures what it purports to measure. ER scales correlate significantly and in predicted directions with selected MMPI scales.

17. Reliability of the ER in Two Samples of Job Employees

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 fair. Psychometric principles and computer technology insures accuracy, objectivity, practicality, cost-effectiveness and fairness.

In 1991 research studies on the ER continued. Two studies (1991) were conducted to test the reliability of the ER scale in two different samples of job employees. The total N of the two studies was 971 participants. 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. These studies (1991, Group1 = 177, Group 2 = 794) are summarized below.

Method
There were two samples of job employees included in these studies (1991). **The subjects in Group 1 consisted of 177 job employees.** Of the 177 respondents, 171 were men and 6 were women. The demographic composition of this sample is summarized as follows: Age: 16 to 25 years (31, 17.5%); 26 to 35 (93, 52.5%), 36 to 45 (35, 19.8%); 46 to 55 (14, 7.9%); and over 55 (4, 2.3%). Ethnicity: Caucasian (152, 85.9%); Black (11, 6.2%); Hispanic (3, 1.7%); Native American (2, 1.1%); and Other (9, 5.1%). Education: 8th grade or less (15, 8.5%); Some High School (36, 20.3%); GED (36, 20.3%); High School

graduate (63, 35.6%); Some College (23, 13.0%); Business/Technical School (1, .6%); College Graduate (2, 1.1%); and Graduate/Professional Degree (1, .6%).

Group 2 consisted of 794 job employees. There were 677 (85.3%) males and 117 (14.7%) females. Age: Under 16 years of age (1 male); 16 to 25 years (229 males, 28 females); 26 to 45 years (460 males, 29 females); 46 to 55 years (33 males, 6 females); and over 55 (14 males, 4 females). Ethnicity: Caucasian (400 males, 71 females); Black (62 males, 14 females); Hispanic (151 males, 9 females); Native American (59 males, 21 females); Asian (1 female); and Other (5 males, 1 female). Education: 8th grade or less (8 males, 1 female); Some High School (182 males, 36 females); GED (69 males, 6 females); High School graduates (216 males, 34 females); some College (165 males, 34 females); Business/Technical School (8 males); College Graduates (27 males, 5 females); and Graduate/Professional degree (2 males 1 female).

Reliability coefficient alphas are presented in Table 8 for studies #1 and 2 combined (N = 971).

Table 8. Reliability coefficient alphas. Job employees. (1991, N=971)
All coefficient alphas are significant at p<.001.

ER Scales	1 Job Employees N = 177	2 Job Employees N = 794
Truthfulness	.85	.85
Alcohol Scale	.84	.90
Drugs Scale	.91	.89
Interpersonal Skills	.81	.84
Work Orientation	.92	.88
Stress Coping Abilities	.92	.94

The results of these studies support the reliability (internal consistency) of the ER. All coefficient alphas are significant at p<.001. All scale reliability coefficients attained very high levels. In both of these samples of job employees ER reliability coefficient alphas were very high and very similar between samples. These results show that the ER is a reliable job employee assessment instrument. ER scales are objective, verifiable, reproducible and reliable. The internal consistency (reliability) of the ER has been demonstrated. The ER has acceptable and empirically demonstrated reliability.

18. Reliability of the ER in a Sample of Vocational Rehabilitation Clients

This study (1992) tested the reliability of the ER in a sample of vocational rehabilitation clients. With expanded use of the ER across the US and Canada, the reliability of the ER continues to be investigated in different samples of job employees or individuals being assessed for employment. This study explored the applicability of the ER for different types of job employees. Vocational rehabilitation clients usually have special concerns regarding their employability. High reliability statistics of the ER with vocational rehabilitation clients would suggest that the ER has broad applicability for assessing a variety of job employees. This study (1992, N = 446) is summarized below.

Method

There were 446 vocational rehabilitation clients included in this study (1992). There were 347 (77.8%) males and 99 (22.2%) females. Age: 221 (16 to 25 years); 143 (26 to 35); 46 (36 to 45); 31 (46 to 55); and 5 (over 55 years). Ethnicity: Caucasian (370); Black (18); Hispanic (14); Asian (1); Native American (39); and Other (4). Education: Below 8th grade (24); Some High School (71); GED (64); High School graduates (155); Some College (92); Business/Technical School (9); and College Graduates (31).

Reliability coefficient alphas are presented in Table 9 for 446 vocational rehabilitation clients.

Table 9. Reliability coefficient alphas. Vocational rehabilitation clients (1992, N=446)
All coefficient alphas are significant at p<.001.

ER Scales	Cronbach's Alpha
Truthfulness	.84
Alcohol Scale	.88
Drugs Scale	.90
Interpersonal Skills	.84
Work Orientation	.85
Stress Coping Abilities	.91

This study supports the reliability (internal consistency) of the ER in vocational rehabilitation clients. All scales have highly significant reliability coefficient alphas. Cronbach Alpha is considered the most important index of internal consistency or reliability. Reliability refers to consistency of test results regardless of who uses the test. ER scales have been demonstrated in a number of studies to be both mutually exclusive and have high inter-item scale consistency. ER scales are objective; verifiable, reproducible and reliable. The internal consistency (reliability) of the ER has been demonstrated. The ER is a reliable assessment instrument for vocational rehabilitation clients.

19. Reliability of the ER and the ER Judgment and Security Scales

In 1994, the ER was revised to incorporate two new measures called the Judgment Scale and Security Scale. This required reorganization of the ER and statistical analysis of the ER scales. The Judgment Scale measures understanding and comprehension. Understanding incorporates logical and comprehension abilities. Judgment incorporates an employee's ability to compare facts or ideas to understand their relationships and to draw conclusions. Judgment is necessary for a person to evaluate their situation and decide upon future action. Judgment involves understanding and affects decision-making. Work-related problems increase as judgment decreases. Employees with good judgment would be considered less of an employment risk.

The Security Scale measures endangerment, violence, unsocial attitudes and ethical blunting. This is a security measure that integrates a checklist of prior unsocial, violent and court-related offenses. Reliability research was used to review all ER scales. The study (1994, N = 245) involving both new job employees and rehires follows.

Method
In this study (1994), the ER was administered to 191 New Employees and 54 Rehires (N=245). There were 126 males (51.4%) and 119 females (48.6%). Rehires had worked for the participating company, had been laid off, and were designated as appropriate for rehiring. New Employees had never been

employed at the participating company. The **demographic composition of the New Employees** sample is described as follows: Gender: There were 98 males (51.3%) and 93 females (48.7%). Age: 16-20 (54, 28.3%), 21-25 (49, 25.7%), 26-30 (32, 16.8%), 31-35 (20, 10.5%), 36-40 (15, 7.9%), 41-45 (11, 5.8%), 46-50 (4, 2.1%), 51-55 (3, 1.6%), over 55 (3, 1.5%). Ethnicity: Caucasian (57, 29.8%), Black (39, 20.4%), Hispanic (81, 42.4%), Asian (5, 2.6%), Native American (7, 3.7%), Other (1, 0.5%). Education: 8th Grade or Less (16, 8.4%), Some HS (59, 30.9%), GED (10, 5.2%), HS Graduate (82, 42.9%), Some College (14, 7.3%), Technical/Business School (3, 1.6%), College Graduate (1, 0.5%). Marital Status: Single (143, 74.9%), Married (44, 23%), Divorced (4, 2.1%), Separated (0).

The **demographic composition of the Rehires** sample is described as follows: Gender: There were 28 males (51.9%) and 26 females (48.1%). Age: 16-20 (15, 27.8%), 21-25 (8, 14.8%), 26-30 (11, 20.4%), 31-35 (6, 11.1%), 36-40 (4, 7.4%), 41-45 (5, 9.3%), 46-50 (3, 5.6%), 51-55 (1, 1.9%), over 55 (1, 1.9%). Ethnicity: Caucasian (17, 31.5%), Black (9, 16.7%), Hispanic (21, 38.9%), Asian (3, 5.6%), Native American (4, 7.4%), Other (0). Education: 8th Grade or Less (3, 5.6%), Some HS (19, 35.2%), GED (1, 1.9%), HS Graduate (21, 38.9%), Some College (7, 13.0%), Technical/Business School (0), College Graduate (0). Marital Status: Single (33, 61.1%), Married (19, 35.2%), Divorced (0), Separated (1, 1.9%), Widowed (1, 1.9%).

Gender comparisons indicated that age, race and education were not significantly different between males and females. However, marital status was significantly different where males were more often single and females were more often married, t=2.29, p=.02. Also, there was a gender difference in US citizenship status where more females than males were not US citizens. There was no difference between genders for possessing a driver's license. Comparisons between Rehires and New Employees indicated that the groups differed only in marital status where Rehires were more often married and New Employees were more often single.

The average age of all participants was 27.8 years. The youngest age was 18 and the oldest age was 63. The average age for Rehires was 29.3, and the youngest and oldest ages were 18 and 56, respectively. The average age for New Employees was 27.4, and the youngest and oldest ages were 18 and 63, respectively. Males were a little older (2 years on average) than females, however, the t-test comparison between gender groups showed that the difference was not significant. The Rehires were also 2 years older than New Employees, but again, the difference was not significant. A test for homogeneity of variance indicated that the distributions of age of the groups were not significantly different.

A t-test comparison between groups indicated that education level was not significantly different between Rehires and New Employees. The groups did not differ in terms of race, US citizenship or having a driver's license. The groups did differ in terms of marital status, where Rehires were more often married and New Employees were more often single.

The participants' prior history for court-related convictions, probation and arrests was obtained from information provided by the participants on the ER answer sheet. This **prior history for New Employees** was as follows: Misdemeanor Convictions: 0 (153, 80.1%), 1 (20, 10.5%), 2 (9, 4.7%), 3 or more (6, 3%). Felony Convictions: 0 (182, 95.3%), 1 (5, 2.6%), 2 or more (1, 0.5%). Probation: 0 (171, 89.5%), 1 (14, 7.3%), 2 (2, 1.0%), 3 or more (1, 0.5%). Probation Revocations: 0 (185, 96.9%), 1 (2, 1.0%), 2 or more (1, 0.5%). Total Number of Arrests: 0 (137, 71.7%), 1 (26, 13.6%), 2 (13, 6.8%), 3 (5, 2.6%), 4 or more (7, 3.7%). Alcohol Convictions: 0 (177, 92.7%), 1 (5, 2.6%), 2 (6, 3.1%), 3 or more (1, 0.5%). Drug

Convictions: 0 (185, 96.9%), 1 (2, 1.0%), 2 or more (1, 0.5%). DUI Convictions: 0 (178, 93.2%), 1 (7, 3.7%), 2 (2, 1.0%), 3 or more (2, 1.0%).

The **prior history for Rehires** was as follows: Misdemeanor Convictions: 0 (40, 74.1%), 1 (6, 11.1%), 2 (4, 7.4%), 3 or more (3, 5.6%). Felony Convictions: 0 (50, 92.6%), 1 (3, 5.6%), 2 or more (0). Probation: 0 (44, 81.5%), 1 (7, 13.0%), 2 (1, 1.9%), 3 or more (1, 1.9%). Probation Revocations: 0 (52, 96.3%), 1 (1, 1.9%), 2 or more (0). Total Number of Arrests: 0 (35, 64.8%), 1 (7, 13.0%), 2 (5, 9.3%), 3 (3, 5.6%), 4 or more (2, 3.8%). Alcohol Convictions: 0 (48, 88.9%), 1 (3, 5.6%), 2 (1, 1.9%), 3 or more (1, 1.9%). Drug Convictions: 0 (52, 96.3%), 1 (1, 1.9%), 2 or more (0). DUI Convictions: 0 (48, 88.9%), 1 (4, 7.4%), 2 (0), 3 or more (1, 1.9%).

Comparisons between New Employees and Rehires on court-related history show similar results. Similar percentages are reported for felony convictions, times on probation, probation revocations, alcohol-related convictions, drug-related convictions and prior DUI or DWI convictions. The largest differences between New Employees and Rehires are reflected in misdemeanor convictions (80.1% of New Employees had no incidents, whereas 74.1% of Rehires had no incidents). This means that 19.9% of New Employees had one or more misdemeanor convictions, and 25.9% of Rehires had one or more misdemeanor convictions. Similarly, 64.8% of Rehires, and 71.9% of New Employees did not report a prior arrest. In other words, 35.2% of Rehires, and 29.3% of New Employees had a prior arrest.

An analysis of variance was performed on each ER Work Orientation Scale item to determine Rehires and New Employees differences. Those items that discriminated between these two groups (significant difference of .10 or less) were selected for inclusion in the Work Orientation Scale. As noted earlier, the Work Orientation Scale is a work appraisal measure incorporating the employee's attitude, work history and overall work adjustment. It is important that this scale discriminates between "acceptable" and "unacceptable" employees. Since Rehires were identified as meeting the "rehire criteria", they were representative of the desired hiring criteria. All items on the Work Orientation Scale discriminate between the Rehire and New Employee groups at the .10 significance level or less. This discriminating ability is a very desirable feature of the ER.

Reliability coefficient alphas are presented in Table 10.

Table 10. Reliability coefficient alphas. New Employees and Rehires (1994, N=245)
All coefficient alphas are significant at p<.001.

ER Scales	Cronbach Alpha
Truthfulness Scale	.88
Alcohol Scale	.91
Drugs Scale	.89
Interpersonal Skills	.85
Work Orientation	.87
Security Scale	.85
Judgment Scale	.86
Stress Coping Abilities	.92

These results strongly support the reliability (internal consistency) of the ER. And in particular, the Judgment Scale and Security Scale were shown to have statistically highly significant reliability. All

reliability coefficient alphas were significant at p<.001. The ER is an objective, standardized and reliable job employee assessment instrument.

ER scales were found to be significantly independent of other ER scales. This mutual exclusivity was demonstrated by a within-subjects between measures ANOVA performed on each ER scale. Each ER scale measures one factor. Also, all ER scales demonstrate high inter-item consistency. This is demonstrated with the Cronbach Coefficient Alpha, which is the most widely used test of inter-item reliability. Items in each ER scale are highly related and each scale measures one factor, yet the factor being measured differs from scale to scale.

20. Reliability of the ER in a Large Sample of Job Employees

The reliability of the ER was studied (1995) in a large sample of job employees. As the ER becomes more widely used it is important to continue to investigate the reliability of ER in large samples of participants. This study (1995) involved 1,014 job employees. The Judgment and Security Scales were included in this study.

Method
This study (1995) included 1,014 job employees. There were 406 males and 608 females. All employees completed the Employee Review (ER) as part of their job application procedures. The demographic composition of this sample was summarized as follows: Age: 16 to 20 years (males 111, 27.3%; females 107, 17.6%); 21 to 25 (males 111, 27.3%; females 144, 23.7%); 26 to 30 (males 69, 17.0%; females 88, 14.5%); 31 to 35 (males 48, 11.8%; females 89, 14.6%); 36 to 40 (males 27, 6.7%; females 53, 8.7%); 41 to 45 (males 15, 3.7%; females 57, 9.4%); 46 to 50 (males 9, 2.2%; females 37, 6.1%); 51 to 55 (males 7, 1.7%; females 16, 2.6%); 56 to 60 years (males 3, 0.7%; females 12, 2.0%); and Over 60 (males 4, 1.0%; females 5, 0.8%). Ethnicity or Race: Caucasian (males 149, 36.7%; females 176, 28.9%); Black (males 54, 13.3%; females 105, 17.3%); Hispanic (males 160, 39.4%; females 232, 38.2%); Asian (males 8, 2.0%; females 15, 2.5%); Native American (males 24, 5.9%; females 66, 10.9%); and Other (males 10, 2.5%; females 11, 1.8%). Education: 8th grade or less (males 28, 6.9%; females 51, 8.4%); Some High School (males 103, 28.4%; females 157, 25.8%); GED (males 25, 6.2%; females 40, 6.6%); High School Graduate (males 197, 48.5%; females 280, 46.1%); Some College (males 43, 10.6%; females 68, 11.2%); Technical/Business School (males 3, 0.7%; females 1, 0.2%); College Graduate (males 7, 1.7%; females 9, 1.5%); Professional/Graduate School (males 0; females 1, 0.2%). Marital status: Single (males 288, 70.9%); females 327, 53.8%); Married (males 93, 22.9%; females 197, 31.9%); Divorced (males 17, 4.2%; females 51, 8.4%); Separated (males 6, 1.5%; females 27, 4.4%); and Widowed (males 0; females 7, 1.2%).

There were 1,014 employees included and 406 were males (40.0%) and 608 were females (59.9%). Gender comparisons indicated that education was not significantly different between males and females. However, age, ethnicity and marital status were significantly different for males and females. Males were younger than females (t = 5.49, p < .001). More males were Caucasian than females (t = 2.37, p < .018). Males were more often single and females more often married (t = 6.02, p < .001). There was no gender difference in U.S. citizenship status. Of the 1,014 employees, 367 (90.4%) males and 552 (90.8%) females had U.S. Citizenship. Similarly 285 males (70.2%) and 448 females (73.7%) had driver's licenses.

The employees' prior history for court-related convictions, probation and arrests was obtained from information provided by the employee on the ER answer sheet. This prior history for the participants in this study was as follows. Misdemeanor Convictions: 0 (866, 85.4%), 1 (71, 7.0%), 2 (39, 3.8%), 3 (18, 1.8%), 4 (7, 0.7%), 5 (3, 0.3%), 6 or more (6, 0.6%). Felony Convictions: 0 (975, 96.2%), 1 (30, 3.0%), 2 (4, 0.4%), 3 or more (2, 0.2%). Probation: 0 (926, 91.3%), 1 (76, 7.5%), 2 (5, 0.5%), 3 (3, 0.3%), 4 or more (1, 0.1%). Probation Revocations: 0 (993, 97.9%), 1 (14, 1.4%), 2 (3, 0.3%), 4 or more (1, 0.1%). Total Number of Arrests: 0 (823, 81.2%), 1 (83, 8.2%), 2 (45, 4.4%), 3 (34, 3.4%), 4 (4, 0.4%), 5 (8, 0.8%), 6 or more (12, 1.2%). Alcohol Convictions: 0 (954, 94.1%), 1 (37, 3.6%), 2 (13, 1.3%), 3 (5, 0.5%), 4 or more (0). Drug Convictions: 0 (987, 97.3%), 1 (17, 1.7%), 2 (4, 0.4%), 3 or more (1, 0.1%). DUI Convictions: 0 (949, 93.6%), 1 (46, 4.5%), 2 (11, 1.1%), 3 or more (2, 0.2%).

Reliability coefficient alphas are presented in Table 11 for 1,014 job employees that completed the ER.

Table 11. Reliability coefficient alphas. Job Employees (1995, N=1,014)
All coefficient alphas significant at p<.001.

ER Scales	Cronbach's Alpha
Truthfulness Scale	.887
Alcohol Scale	.906
Drugs Scale	.885
Interpersonal Skills	.847
Work Orientation	.866
Security Scale	.846
Judgment Scale	.854
Stress Coping Abilities	.906

These results show that all ER scales demonstrate high inter-item consistency. All coefficient alphas were significant at the p<.001 level. In this study (1995) a large sample (N=1,014) of job employees was used to investigate ER reliability. This study strongly supports the reliability (internal consistency) of the ER.

21. ER Reliability and Scale Risk Range Accuracy

This study (1998) was conducted to test the reliability and accuracy of the ER. Risk range percentile scores are calculated for each ER scale. These risk range percentile scores are derived from scoring equations based on responses to scale items and Truth-Correction. 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 ER risk range percentile scores involves comparing the risk range percentile scores obtained from employee ER test results to the predicted risk range percentages as defined above. The percentages of employees expected to fall into each risk range are the following: Low Risk (**39%**), Medium Risk (**30%**), Problem Risk (**20%**) and Severe Problem or Maximum Risk (**11%**). The actual percentage of employees falling in each of the four risk ranges, based on their risk range percentile scores, was compared to these predicted percentages in the following 1998 (N = 2,446) study.
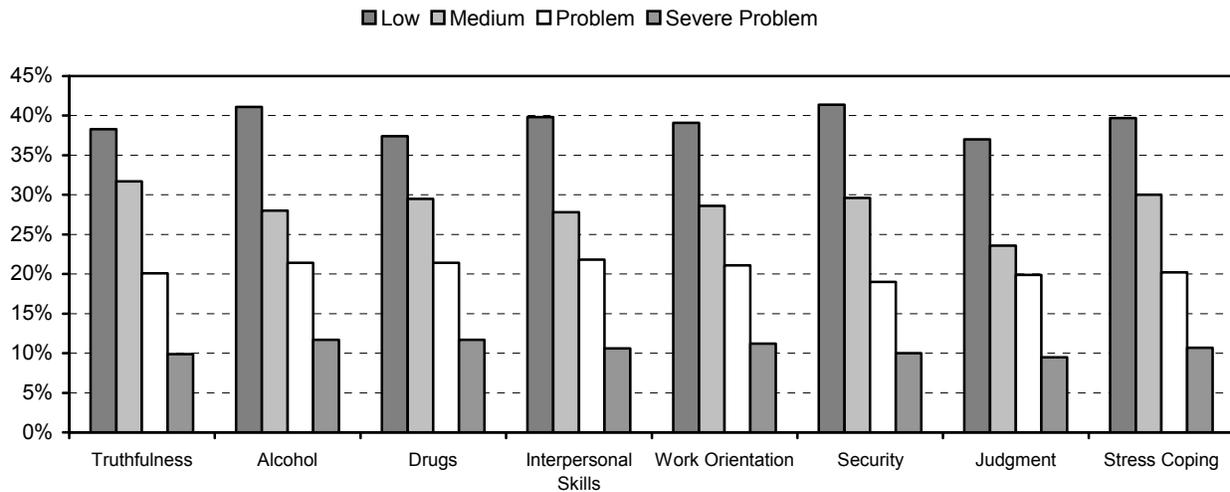
24

Method
This study (1998) included 2,446 job employees from two geographical areas. There were 119 participants from a southeastern location and 2,327 participants were from the southwest. The test sites were temporary employment services that administered the ER as part of routine screening procedures.

Accuracy
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 12 presents these comparisons. The differences between obtained and expected are shown in parentheses.

Table 12. ER Risk Range Percentile Scores (1998, N = 2,446).

■Low ■Medium □Problem ■Severe Problem

Scale	Low Risk (39%)		Medium Risk (30%)		Problem Risk (20%)		Severe Problem (11%)	
Truthfulness Scale	38.3	(0.7)	31.7	(1.7)	20.1	(0.1)	9.9	(1.1)
Alcohol Scale	41.1	(2.1)	28.0	(2.0)	21.4	(1.4)	11.7	(0.7)
Drugs Scale	37.4	(1.6)	29.5	(0.5)	21.4	(1.4)	11.7	(0.7)
Interpersonal Skills	39.8	(0.8)	27.8	(2.2)	21.8	(1.8)	10.6	(0.4)
Work Orientation	39.1	(0.1)	28.6	(1.4)	21.1	(1.1)	11.2	(0.2)
Security Scale	41.4	(2.4)	29.6	(0.4)	19.0	(1.0)	10.0	(1.0)
Judgment Scale	37.0	(2.0)	23.6	(3.6)	19.9	(0.1)	9.5	(1.5)
Stress Coping Abilities	39.7	(0.7)	30.0	(0.0)	20.2	(0.2)	10.7	(0.3)

As shown in the graph and table above, the ER scale scores are very accurate. The objectively obtained percentages of participants falling in each risk range are very close to the expected percentages for each risk category. **All of the obtained risk range percentages were within 3.6 percentage points of the expected percentages and many (18 of 32 possible) were within one percentage point.** Only four obtained percentages were more than two percent from the expected percentage and these were off by 2.4%, 2.2%, 3.6% and 2.1%. This is very accurate assessment.

For those participants who are identified as having problems (Problem and Severe Problem risk ranges or 31% of the participants), the obtained percentages were extremely accurate. The comparisons between obtained and expected percentages are the following: Truthfulness Scale 30%, Alcohol Scale 30.9%, Drugs Scale 33.1%, Interpersonal Skills 32.4%, Work Orientation 32.3%, Security Scale 29%, Judgment Scale 29.4%, Stress Management Scale 30.9%. The problem risk ranges for all ER scales are in close agreement to the expected percentage. These results demonstrate that the ER scale scores accurately identify job employee risk.

Reliability coefficient alphas are presented in Table 13.

Table 13. Reliability coefficient alphas. Job Employees (1998, N=2,446)
All coefficient alphas significant at p<.001.

ER Scales	Cronbach's Alpha
Truthfulness Scale	0.86
Alcohol Scale	0.85
Drugs Scale	0.84
Interpersonal Skills	0.82
Work Orientation	0.88
Security Scale	0.85
Judgment Scale	0.85
Stress Coping Abilities	0.92

These results support the reliability of the ER. All coefficient alphas were significant at p<.001. All coefficient alphas for ER scales are above the generally accepted level of 0.80 for assessment tests. The ER is a reliable employee screening instrument.

Taken together, these results demonstrate that the ER is a reliable and accurate job employee assessment instrument. All ER scales have very high reliability coefficient alphas and ER scale risk range percentile scores closely approximate their predicted percentages. The ER is an accurate employee assessment tool.

22. Reliability, Validity and Accuracy of the ER

ER scales were reassessed to comply with users needs. For example, many ER users requested the Judgment Scale be replaced with an Aggressiveness Scale. It was felt that the Judgment Scale did not enhance screening enough, whereas the inclusion of the Aggressiveness Scale does. As noted earlier, aggressiveness incorporates self-assertiveness, social dominance and a tendency toward hostility. These aggressive behaviors contribute to understanding "people problems" which account for as much as 80% of employee turnover. The revised ER was shortened to 136 items. This scale replacement greatly expands the comprehensiveness of the ER.

Research on many Behavior Data Systems assessment instruments has revealed gender differences on some of their scales. For this reason gender information has been included on the ER answer sheet. When gender differences exist on any test scale, separate male/female scoring methodologies are developed. When appropriate, separate scoring procedures for males and females helps ensure fairness and accuracy. Similarly, recent research has concluded that race or ethnicity accounts for some differences in scale scores. Race information has also been included on the ER answer sheet.

A study (1999) was conducted to investigate the reliability and accuracy of the revised ER. ER test data were compiled from agencies located in the Southwest, Midwest and Southeast. There were a total of 3,513 job employees participating in this study.
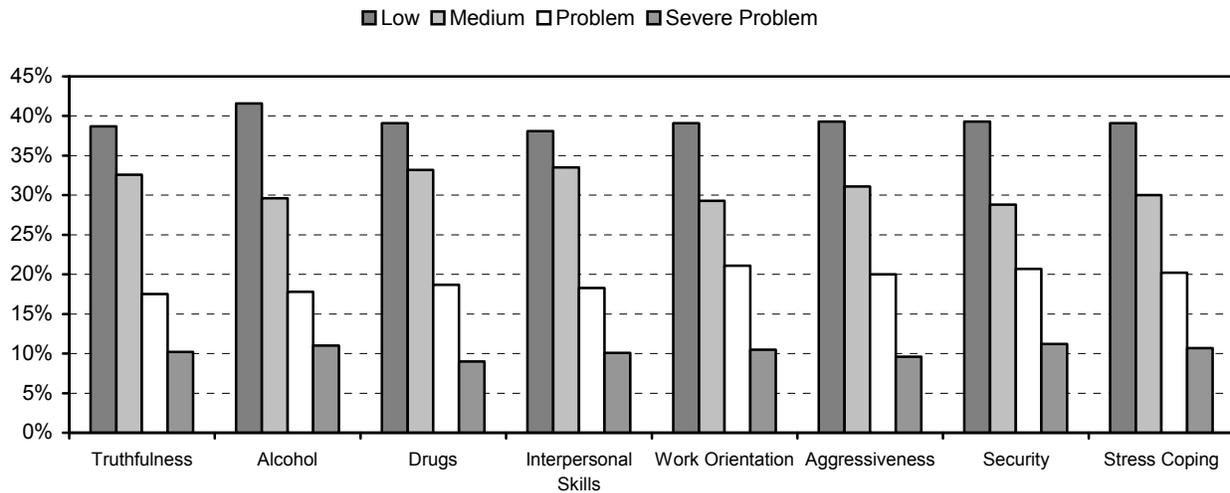
Method

There were 3,513 job employees included in this study (1999). The participants were from three geographical areas: Southwest, Southeast and Midwest. Each job employee completed the ER on a voluntary ER research basis.

Accuracy

Accuracy of the ER is determined by the close approximation of obtained risk range percentages to predicted percentages. There are four risk range percentages that clients are assigned to based on their scale scores. The risk range percentages and the predicted (shown in parentheses) are: **Low Risk** (39%), **Medium Risk** (30%), **Problem Risk** (20%) and **Severe Problem or Maximum Risk** (11%).

Table 14. ER Risk Range Percentile Scores (1999, N = 3,513).

■Low ▨Medium □Problem ■Severe Problem

Scale	Low Risk (39%)		Medium Risk (30%)		Problem Risk (20%)		Severe Problem (11%)	
Truthfulness Scale	38.7	(0.3)	32.6	(2.6)	17.5	(2.5)	10.2	(0.8)
Alcohol Scale	41.6	(2.6)	29.6	(0.4)	17.8	(2.2)	11.0	(0.0)
Drugs Scale	39.1	(0.1)	33.2	(3.2)	18.7	(1.3)	9.0	(2.0)
Interpersonal Skills	38.1	(0.9)	33.5	(3.5)	18.3	(1.7)	10.1	(0.9)
Work Orientation	39.1	(0.1)	29.3	(0.7)	21.1	(1.1)	10.5	(0.5)
Aggressiveness Scale	39.3	(0.3)	31.1	(1.1)	20.0	(0.0)	9.6	(1.4)
Security Scale	39.3	(0.3)	28.8	(1.2)	20.7	(0.7)	11.2	(0.2)
Stress Coping Abilities	39.1	(0.1)	30.0	(0.0)	20.2	(0.2)	10.7	(0.3)

These results show that obtained risk range percentages closely approximated the predicted percentages for each of the eight ER scales. All obtained risk range percentages were within 3.5 percentage points of the predicted. There were 19 out of the 32 possible risk ranges that were within one percentage point of the predicted. Only six obtained risk range percentages deviated from the predicted by more than two percentage points and these were 2.6%, 2.6%, 3.2%, 3.5%, 2.5% and 2.2%. This is very accurate

assessment. The ER accurately measured risk in this sample of job employees. The ER is an accurate job employee assessment instrument.

Reliability coefficient alphas are presented in Table 15 for the 3,513 job employees studied.

Table 15. Reliability coefficient alphas. Job Employees (1999, N=3,513)
All coefficient alphas significant at p<.001.

ER Scales	Cronbach's Alpha
Truthfulness Scale	0.86
Alcohol Scale	0.85
Drugs Scale	0.84
Interpersonal Skills	0.83
Work Orientation	0.88
Aggressiveness Scale	0.85
Security Scale	0.81
Stress Coping Abilities	0.92

Validity

The ER Alcohol and Drugs Scales are measures of alcohol and drug abuse or severity of abuse. Alcohol and Drugs Scales scores predict when an individual has an alcohol or drug problem. **The criteria in this analysis for identifying employees as problem drinkers and drug users is having been in treatment (alcohol or drugs).** Having been in treatment identifies employees 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. In this ER study, treatment information was obtained from the employee. Thus, employees are separated into two groups, those who had treatment and those who have not had treatment. Then, employee scores on the Alcohol and Drugs Scales were compared. It is predicted that employees with a treatment history will score in the problem risk range (70th percentile and above) or higher on the Alcohol and Drugs Scales. Alcohol and drug treatment information was obtained from employee answers to ER test items (#50, #29 & #92) regarding alcohol and drug treatment.

Predictive validity results for the Alcohol Scale (using scale scores) and alcohol treatment show that for the 124 employees who reported having had alcohol treatment, all 124 individuals, or 100 percent, had Alcohol Scale scores at or above the 70th percentile. These results show that the ER Alcohol Scale accurately identified employees with alcohol problems. There is a very strong positive correlation between Alcohol Scale scores and alcohol treatment. **These results strongly validate the ER Alcohol Scale.**

The predictive validity of the Drugs Scale was done in the same way using drug treatment as the criterion. Of the 195 employees who reported having had drug treatment all 195 individuals or 100 percent had Drugs Scale scores in the 70th percentile or higher (Problem Risk and above). **These results strongly validate the ER Drugs Scale.**

These results strongly support the reliability, validity and accuracy of the ER. The ER achieves very high statistical reliability. It accurately classifies employees into risk range categories and it accurately identifies employees who have drinking and/or drug abuse problems. The ER does what it purports to do and that is to accurately screen job employees.

23. Reliability of the Revised ER

This study (2010) examined the inter-item reliability of the revised Employee Review (ER). The ER was revised in the latter part of 2010; this revised version of the ER replaced the previous version. The Interpersonal Scale, Security Scale, Judgment Scale and Aggressiveness Scale were omitted from the revised version and the Work Orientation Scale was substantially expanded. Individual items from each scale were revised or replaced with new items. The ER now has the following five (5) scales: 1) **Truthfulness Scale**, 2) **Work Orientation Scale**, 3) **Alcohol Scale**, 4) **Drugs Scale**, and 5) **Stress Management Scale**. The reliability, validity and accuracy analyses presented in previous studies were closely approximated. This study represents ongoing ER database research.

Method and Results

The participants in this study (2010) were 2,496 employees. 94.7% of the sample was male and 5.3% were female. The demographic composition of this group is as follows: Age: Under 20 (7.1%); 20 through 29 (29.0%); 30 through 39 (30.6%); 40 through 49 (21.9%); 50 through 59 (6.5%); 60 and older (4.8%). Ethnicity: Caucasian (72.3%); Black (14.1%); Hispanic (9.2%); Asian (1.1%); Native American (1.4%); Other (1.9%). Education: 8th grade or less (4.4%); Some High School (22.8%); High School Graduate (40.8%); Some College (21.2%); College Graduate (7.7%); Professional/Graduate Degree (3.1%). Marital Status: Single (47.7%); Married (28.3%); Divorced (17.1%); Separated (5.9%); Widowed (1.0%).

Table 6. Reliability Coefficient Alphas (2010, N=2,496)
Vocational Rehabilitation Clients

ER Scales	Cronbach's Alpha	Significance Level
Truthfulness	.86	.001
Alcohol Scale	.80	.001
Drugs Scale	.82	.001
Work Orientation	.80	.001
Stress Management	.92	.001

SUMMARY

This document "ER: An Inventory of Scientific Findings" is not intended as an exhaustive compilation of ER research. Yet it does summarize many research studies supporting the reliability, validity and accuracy of the Employee Review (ER). Moreover, ongoing ER database research ensures an increasingly accurate picture of ER job employees and the risk they represent. It is reasonable to conclude the ER provides a sound empirical basis for responsible employee-related decisions. The ER does what it purports to do.

The ER acquires a vast amount of relevant information for staff review prior to decision making. Empirically based scales are objective and accurate. Assessment has shifted from subjective opinions to objective accountability.

It should be noted that in this research document studies are presented chronologically – when the research was done. This enables the reader to see the evolution of the ER into a state-of-the-art employee screening instrument or test. The ER provides an empirical foundation for responsible decision making. The ER is a reliable, valid and accurate instrument for employee screening.

Areas for future research are many and complex. ER research will continue to be ongoing in nature. Wherever possible, emphasis will be placed on local standardization research. The advantages of local test standardization are many and include accuracy, reliability, validity and fairness. When Behavior Data Systems (BDS) standardizes a test on a client population, the instrument becomes uniquely appropriate to that user's needs. BDS offers to standardize the ER on a new (or prospective) client's employee (or potential employees) workforce. This typically involves one month of free testing along with a free database analysis for companies/agencies that are the first to use the ER in their industry and/or state. This procedure insures employee/employee ER standardization. This unique and very desirable feature is only possible because of BDS' proprietary ER database. Consistent with the foregoing, BDS encourages other scientists to participate in ER research. Few fields of assessment represent such important opportunities for creative discovery.

In summary, this document is a cumulative record of the evolution of the Employee Review (ER). Studies are presented chronologically – in the same sequence they were completed. Current studies are most representative of the ER. Behavior Data Systems is committed to ongoing research. Interested parties should contact Behavior Data Systems, PO Box 44256, Phoenix, Arizona 85064-4256.

Donald D. Davignon, Ph.D.
Senior Research Analyst

041500