National Center for Intermodal Transportation
&
Economic Development

Safe and Competent Intermodal Transportation Workers:

The Interaction of Shift Work and Worker Resilience on Health & Safety of Intermodal Transportation Workers

NCITEC Project: 2012_18

Patrick Sherry Ph.D.
Alison Bondanza , Ph.D.

National Center for Intermodal Transportation
University of Denver

Revised May 15, 2016

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Safe and Competent Intermodal Transportation Workers: The Interaction of Shift Work and Worker Resilience on Health & Safety of Intermodal Transportation Workers

Patrick Sherry Ph.D.

Principal Investigator: Alison Bondanza, PhD

National Center for Intermodal Transportation
University of Denver
2400 S. Gaylord
Denver, CO 80208

Abstract

Workplace accidents have a devastating effect on the transportation industry. Recently, the United States Department of Transportation acknowledged that over 80% of transportation accidents are human factors caused. In an effort to develop a pre-employment screening tool for selecting individuals who will have a lower risk of being involved in an accident or incident, investigators administered a survey instrument to 459 intermodal shift workers in a large mid-western rail company. The survey instrument assessed attitudes, personality characteristics, work habits and practices, etc. related to safe and effective shift work. Data from work histories was then correlated with survey results to establish concurrent validity of the assessment instruments for predicting for workplace safety. Results of analyses indicated that Spatial Reasoning, Conscientiousness and Neuroticism were significantly correlated with the number of personal injuries sustained by persons in the sample. A number of significant correlations were obtained between scales on the Denver Lifestyle Inventory (DSI) and personal injuries. In addition, Derailments were significantly correlated with behavior and attitudinal predictors reflecting physical and mental distractibility and activity, possibly indicating a possible pattern or individual characteristic of distractibility or lack of concentration. Distractibility was also significantly correlated with Reaction Time and Digit Span Trials Correct from the Web Neuro Battery. Lastly, Absenteeism was significantly correlated with items related to comfort and satisfaction with shift work schedules. The data gathered in the study provide validity consistent with that of the Uniform Personnel Selection Guidelines (EEOC, 1964, 1978). Implications of these results suggest a higher probability for selecting and maintaining a work force with a low accident injury rate by using the pre-employment selection tools.

Table of Contents

Abstract 4
Introduction 7
Causes of Accidents 8
Resilience Theory and Safety 8
Shift Work 9
Strengthening Workforce Capacity Through Selection 9
Approaches to Selection 10
Previous Work in Personnel Selection 12
General Research in Predicting Job Performance 13
Purpose of the Study 14
Method 15
Participants 15
Instrument Construction 15
Cognitive Interviews 15
Data Analysis 15
Procedures 16
Instruments 16
Denver Life Style Inventory (DLSI) 16
Differential Aptitude Test (DAT) 16
AON Workplace Practices Test (WPT) 18
Employee Quality Measurement System (EQMS) 19
Web Neuro Tests 19
Results 21
Participants 21
EQMS Scores 21
Personal Injuries 24
Decertifications 26
Derailments 28
Absenteeism 32
Efficiency Testing 33
Shift Work and Safety 35
Summary of Findings 36
Discussion 37

Study Limitations 37
Conclusion 38
References 39

List of Figures

Figure 1. Causal pathways affecting worker safety and health. (NIOSH, 2009) 8
Figure 2. Hypothesized model of job performance. 10
Figure 3. Steps in the develpment of a selection measurement. (*from FRA, 2003*) 11
Figure 4. Adapted from (Gordon et. al, 2015). 20
Figure 5. Frequency of number of Personal Injuries for railroad employees. 24
Figure 6. Personal Injuries and Job Tenure. 26
Figure 7. Dercertificaitons for employees in the study. 27
Figure 8. Derailments over time of employment. 29
Figure 9. Frequency of Derailments in the sample. 30

List of Tables

Table 1. DAT for PCA Subtests and Administration Time. 17
Table 2. DAT Subtest Reliablitites. 18
Table 3. Correlations between predictor and criterion measures 21
Table 4. Characteristics of study participants. 22
Table 5. Correlations between predictive measures and EQMS scores. 23
Table 6. Correlations between predictive measures and number of Personal Injuries (N=338). 25
Table 7. Correlatnions between assesment measures and Decertifcations. 27
Table 8. Regression analysis of predictors on to Decertifcations. 28
Table 9. Regression analysis predictors of Derailments in work history. 30
Table 10. Significant Correlations of Web Neuro Measures with Derailments 31
Table 11. Correlations of measures with Absenteeism. 32
Table 12. Regression of assessment battery measures on to Absenteeism. 33
Table 13. Correlation between assessment measures and number of efficiency test failrures. 34
Table 14. Assessment of shift characteristics and safety outcomes. 35
Table 15. Correlation between assessment instrument scores and criterion measures. 42

Safe and Competent Intermodal Transportation Workers: The Interaction of Shift Work and Worker Resilience on Health & Safety of Intermodal Transportation Workers

Introduction

The goal of identifying "safe and competent" transportation workers and utilizing tools to reduce the number of workplace accidents has been identified since the formation of OSHA in 1971. Later work established that congruence between the attitudes and beliefs of supervisors and employees was related to the occurrence of workplace accidents and injuries (Sherry, 1991). Since then a number of different approaches have been offered and suggested as a means of reducing workplace accidents. While the fatality rate on U.S. highways has steadily declined over the last several years, the number of fatalities has remained relatively constant. The nation, as a whole, lost 41,059 lives in 2007 to traffic crashes. In fact, in 2005, motor vehicle traffic crashes were the leading cause of death for every age from 3 through 6 and 8 through 34. In fact:

> Among the critical reasons attributed to drivers, about 41 percent were recognition errors, about 34 percent were decision errors, about 10 percent were performance errors, and about 7 percent were nonperformance errors. About 18 percent of the drivers were involved in at least one non-driving activity, with the majority (about 12%) engaged in conversing either with other passengers or on a cell phone. The effectiveness of emerging crash avoidance technology that use existing vehicle systems such as adaptive cruise control, braking systems, seat belt pre-tensioners, motorized seats, sunroofs, etc., in mitigating the effects of various driver performance, recognition, and decision errors can be assessed using this information. (NHTSA, 2008)

In many cases the approach has been to try to "engineer" the human factor out of the workplace to essentially make it impossible for an accident of injury to occur. However, unfortunately, this has not proven possible such that even in a 150 year old industry such as intermodal rail roads that a significant number of railroad accidents occur annually.

Consequently, the focus has recently shifted to one in which the individual, and the individual characteristics of the person have been focused on as key components of the safety equation. An article by Geller & Wiegand (2005) discussed the importance of behavioral and personality factors as important aspects of behavior that need to be addressed. These behavioral

tendencies include conscientiousness, or the attention to one's environment and a desire to carefully follow procedures and guidelines as an important behavioral characteristic likely to lead to safe work behavior. Also related to the importance of behavioral characteristics and tendencies is the overarching quality of resilience, which is supported by behavioral characteristics of individuals and also the environmental characteristics of the organization. (Norris, et. al, 2008).

Causes of Accidents

Accidents, injuries and human error is thought to result from a number of different causes. For the most part a theoretical model that focuses on the main contributing factors has been promoted. For example, one very popular approach is the utilization of the B=f(P*E) or the interactional approach in which the behavior "B" is thought to be the result of the interaction between the Persona "P" and the Environment "E".

$$B = f(P*E)$$

In this model "B" Behaviors, or in other words accidents and job performance such as operational procedures, lead to injury, are consider "B" or the Behavior. These are thought to be greatly influenced by the environment "E" which consists of the hazards, the opportunities and safeguards. Next the characteristics of the individual "P" or Person are also thought to influence the occurrence of the behavior.

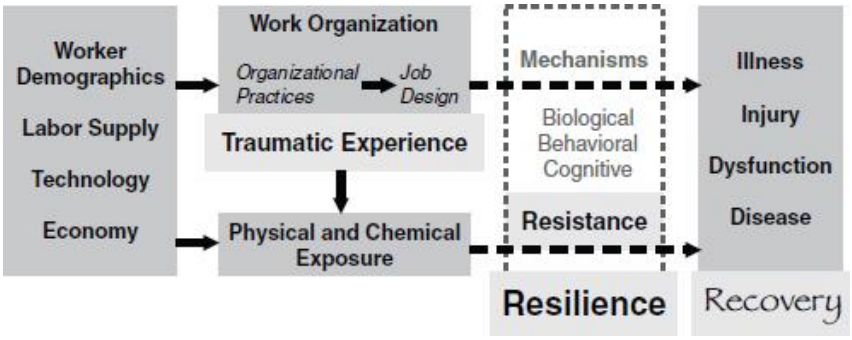

Figure 1. Causal pathways affecting worker safety and health. (NIOSH, 2009)

Resilience Theory and Safety

More recently, NIOSH (Colvin & Taylor, 2012) proposed a model in which a number of factors affect the individual and which then lead to stress and ultimately to poor health, injury and accidents. This model from NIOSH depicted in Figure 1 shows that worker characteristics are input into the organization which then mediates those characteristics via on the job experiences, organizational practices, job design, with a resulting throughput affecting worker

health and safety in the form of injury, illness, incidents and accidents. Accordingly, one way, albeit there are others, is to identify workers significantly more likely to be able to handle and tolerate the organization and work environment experiences without the resultant negative impact on health and safety. In the present study, the role of personality and cognitive factors are hypothesized to have appositive impact on maintaining health and safety. Conversely, those without the requisite personality, attitudinal and cognitive ingredients and capabilities will be ore likely to suffer the negative effects of the work environment.

In addition, a call for organizations to increase the workforce capacity to deal with stress and strain, that can lead to accidents has also been proposed. (Colvin & Taylor, 2012) This can be done by selecting persons who are likely to have characteristics that will contribute to resilience. Personality characteristics such as conscientiousness and emotional stability. In addition, the cognitive abilities and skills associated with verbal. Spatial and mechanical reasoning have also been shown to be useful in predicting successful rail road dispatchers (Sherry, 1993).

Shift Work

Railroad work is particularly challenging from a health and safety standpoint due to the fact that much of the operational work is not done on any particular schedule, especially freight operations. Much of the work of the freight railroad operates on a variable schedule with no definite start time. Consequently, it is difficult to assess the type of shift as there is no definite start time. The employees of the railroad often have to be "on-call" for long periods of time awaiting a call to work. Because of the this fact one characteristic that may be related to resilience when working in a railroad environment might be a favorable attitude towards working consecutive days, working midnights and dealing with the lack of predictability. An appropriate selection assessment tool then, will likely have measures which assess comfort with and interest in the variable shift work found in railroad operations.

Strengthening Workforce Capacity Through Selection

In an attempt to fill this gap in the literature, the current project proposed to develop measures that would assist hiring managers in identify "safe and competent" workers who have less likelihood of being involved in unsafe acts and behaviors such as accidents leading ot personal injury thereby ensuring more resilient individuals and a more resilient work environment. Typical unsafe acts or events in intermodal and railroad work involve derailments, violations of operating rules, etc. Assessing attitudes, personality characteristics, work habits and practices, etc. that are shown to be related to safe and effective shift work in the intermodal industry *before the person enters the workforce* will be beneficial in the long run and likely lead to healthier and safer employees. Given the unique shift work nature of the intermodal industry

in the present study a set of measures that identify the effects that shift work has upon the occurrence of accidents and incidents will also be examined. These shift work identifiers will be combined with the those measures for identifying the work attitudes and personality characteristics associated with resilience in the face of shift work will also be identified. In the realm of psychological theorizing regarding the various factors that influence the occurrence of safe and productive work behavior a tripartite model has been proposed. As shown in figure 2, the combined characteristics of the person, the situation they have been exposed to, and the extent of their training is thought to be related to the total of job performance.

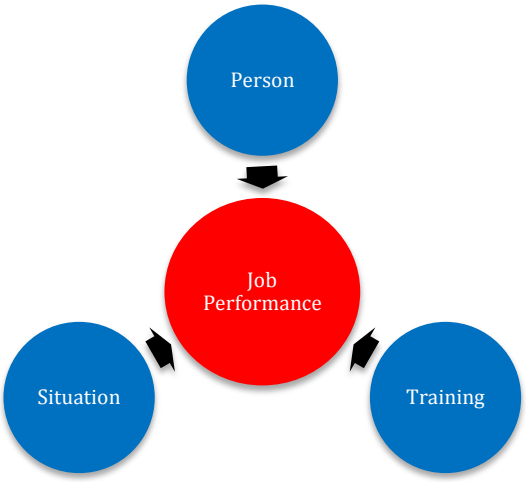

Figure 2. Hypothesized model of job performance.

The unique contribution of the present study is the attempt to identify a combination of personality and behavioral orientation measures along with cognitive and problem solving competencies of persons being considered for the work force, in an applied shift work environment, to predict safety performance and behavior as a means of strengthening workforce capacity and resilience in response to the demands and hazards of the work environment as a means of improving safety and job performance in the long run over time.

Approaches to Selection

Personnel selection approaches have attempted to address the steps that should be followed when developing a tool designed to select personnel for key position. Typically the development begins with a job analysis designed to determine the prerequisite knowledge skills and abilities required of incumbents in a particular position. Gatewood & Field (1994) devoted an entire text to the topic of personnel selection in human resources. The key components of selection are to be performed under legal and other constraints designed to protect the rights of

both employees and employers. In addition, the process should be designed in order to maximize the fairness of the approach to the participants. Having an objective and reliable and valid method of selection ensures greater and greater confidence in the outcomes. Ultimately the purpose is to select individuals who are very likely to successfully perform the task and duties o the job in question.

Previous attempts at developing measure that will improve job performance, of which safety is a key component, have been conducted over the past 30 years. In particular, the Federal Railroad Administration (FRA, 2003) sponsored a study investigating the effectiveness of various measures at predicting job performance. Results of this investigation demonstrated that a sequenced step by step process could lead to the development of standard selection instruments. As shown in Figure 2 identification of appropriate instruments and validation of the instruments should be undertaken before utilizing in a selection program.

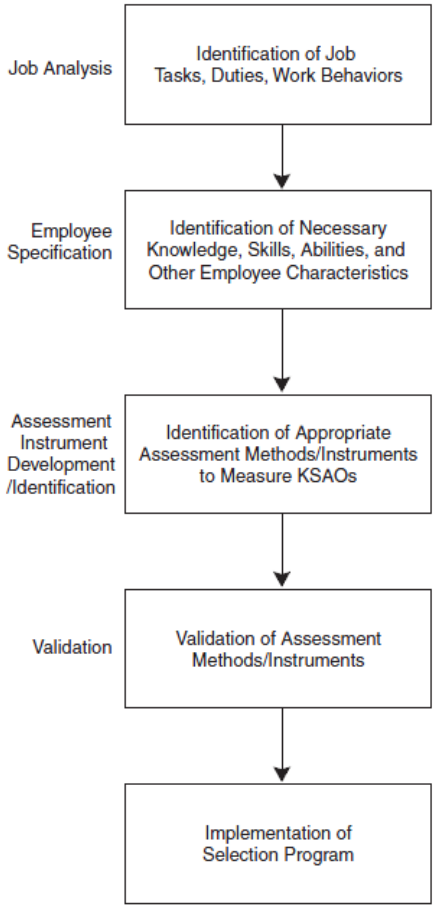

Figure 3. Steps in the develpment of a selection measurement. (*from FRA, 2003*).

Previously, a job analysis, a set of specified knowledge, skills and abilities (KSAs) has been specified for railroad workers. These were identified for engineers and conductors in a large railroad company in the Midwest. Further specification of the KSA's is also located in the requirements for certification which include a number of very specific knowledge domains. These have been further specified and used to develop criteria for successful performance of the job of railroad engineer.

The more difficult task in recent time has been the need to identify consistent criteria across individual performers. Previously there was a limited amount of data available on job performance and only an infrequent amount of data on individual employees. With the advent of large scale computer systems and data bases it has become more possible to gather and maintain data on the performance of large numbers of employees thus creating the opportunity for more robust analyses of potential measures for railroad engineers and conductors. The present study was made possible due to the fact that one large railroad in the Midwest was able to maintain a sufficiently large data base on nearly four hundred and fifty railroad employees on a wide range of performance indicators. Thus, the opportunity presented for the ability to test the validity of several predictors in the extent to which they were significantly related to several readily available standard measures.

Validity is an important question in that it requires the demonstration of a statistically significant relationship between scores on a measure and key indicators of job performance. The following section outlines the efforts that have been made to demonstrate the concurrent and predictive validity of various measures in personnel selection.

Previous Work in Personnel Selection

As noted the FRA (2003) report identified that several different types of measures were available for use in selecting top level performance. Ability tests for example have a long history in other fields of being useful and valid indicators of performance. General ability tests which measure a broad range of verbal, numerical and abstract reasoning skills are commonly used for selection purposes for medical, legal, and engineering profession. However, they have some issues with respect to adverse impact which must be managed carefully. Personality inventories have been shown to be very helpful in many occupations. Considerable predictive validity evidence is available and there is limited risk of adverse impact.

Geller & Wiegan (2005) argued that the use of personality-based measures could provide increased awareness and understanding of the diversity of individual differences related to injury

prevention. They noted that personality factors can affect both injury prevention and injury proneness. For example "entitlement thinking" could stifle personal responsibility to achieve peak performance, including safety. (Geller & Wiegand, 2005).

General Research in Predicting Job Performance

Several studies over the years have suggested that personality and behavioral styles are significantly related to workplace health and safety. For example, Cellar, et. al. (2001) studied two hundred and two undergraduate participants (134 female, 68 male) who completed a personality inventory (NEO-PI-R) and other measures of prior workplace accident involvement. Significant inverse relationships were found between the factors of Agreeableness and Conscientiousness and the number of work-related accidents. Further, regression analyses indicated that both Agreeableness and Conscientiousness factors may be useful for predicting certain types of workplace accidents. Also, Christian, et. al (2009) examined person- and situation-based antecedents of accidents and injuries in the workplace. Interestingly, knowledge and motivation towards safety outcomes were strongly related to safety outcomes, as were organizational and work group safety climate. With regard to accidents and injuries, however, group safety climate had the strongest association.

In research with another occupation devoted to operating vehicles, a personality questionnaire was added to the regression equation for the prediction of job success of airline pilots that was previously based on performance on a check flight simulator and previous flying experience. A total of 274 pilots applying for employment completed a standard personality measure along with other requirements. After selection, the pilots were classified as performing either at standard or below standard after about 3 years of employment. Using multiple-regression analysis, job success was predicted with 73.8% accuracy through the simulator checkflight and flying experience prior to employment, but was raised to 79.3% a statistically significant increase in the total variance accounted for by the measures. Overall, successful pilots scored substantially higher on personality dimensions related to interpersonal behaviors and lower on emotionality scales. This is an interesting finding and suggests that personality adds significantly to the predictive efficiency of a selection battery in addition to the already substantial contribution of other practical and operational testing.

Another study in the aviation industry of 485 Chinese airline employees including pilots, flight attendants, engineers, and service employees Baba, et al (2009) found that a measure of Proactive personality positively predicted a measure of organizational citizenship and individual performance. In addition, a personality measure of emotional exhaustion and perceived safety climate moderated the relationship between proactive personality and individual performance both independently and jointly. Lastly, Barron, et. al. (2016) found that the Big Five personality

domains along with cognitive aptitudes, including aptitude spatial and numerical reasoning as well as aviation knowledge predicted both manned and unmanned aircraft pilot performance on-the-job, as measured by supervisor and senior rater stratification on 3 years of Officer Performance Reports (OPRs). Results were generally consistent in showing that the same aptitudes, knowledge, and personality traits that predict successful job performance for manned aircraft pilots also predicted successful job performance for pilots of unmanned aircraft as well.

Originally, Hurtz and Donovan (2000) reviewed the literature on the relationship between the Big Five personality characteristics and job performance and demonstrated a significant predictive relationship. However, they also pointed to the need for the inclusion of the occupational context as well. Later, Clarke and Roberson (2008) used a meta-analytic technique to show a significant relationship between accident involvement and the Big-Five personality dimensions (extraversion, neuroticism, conscientiousness, agreeableness, and openness). Specifically, low levels of conscientiousness and agreeableness were found to be significantly related to accident involvement validity indices of .27 and .26, respectively. They also found differential validities depending on the context of the occupational performance. Lastly, Kanani, et al. (2014) studied 40 drivers who were hospitalized in 2013 for auto accidents in comparison to a sample of 40 drivers who were not involved in accidents. Results demonstrated that the drivers who had accidents scored higher on several scales of the NEO-PI (Costa, & McCrae, 1992) including Emotionality (N) , Openness (O), Agreeableness (A).

Taken together these studies support the rationale for exploring the measurement of personality and behavioral traits and cognitive abilities in persons seeking to enter the intermodal rail transportation workforce. A consistent finding across studies has been that personality and cognitive ability has been significantly correlated with the safe performance of vehicle operation for aircraft and automobiles.

Purpose of the Study

The present study attempts to identify what combination of personality and behavioral orientation measures along with cognitive and problem solving competencies of persons being considered for the work force, in an applied shift work environment, to predict safety performance and behavior as a means of strengthening workforce capacity and resilience in response to the demands and hazards of the work environment as a means of improving safety and job performance in the long run over time. The objective of the present study was to identify measures that could be used to examine railroad shift worker's attitudes, personality characteristics, work habits and practices, etc. related to safe and effective shift work. A sample of persons working on various schedules in the railroad industry was gathered along with their work history including accidents, incidents and various other performance indicators. At the same time study participants also completed several research instruments and questionnaires

designed to assess their attitudes, personality and behavioral orientate towards their work. Analyses were conducted to determine the relationship between the shiftwork history and their personality and attitudinal data.

Method

Participants

The experimental survey was administered to all of the operating employees of a large western railroad engaged in transportation operations. The participants in the study had the designation of operator, conductor, or engineer in training.

Instrument Construction

Instruments were constructed for the study in several ways: use of existing instruments and construction of new instruments. In addition, a new instrument, the Denver Life Style Instrument, was constructed using items generated from cognitive interviews with various experts and job incumbents.

Cognitive Interviews

As was recommended by DeVellis (1991), the investigators met with several subject matter experts from a local transportation organization and its safety committee to ensure that the survey items had face and content validity. Members of this committee held expertise in safety, culture, risk, survey development, within the transportation industry. In addition, the investigators met with several members of the training staff who were assigned the responsibility for reviewing and certifying the qualifications of engineers and conductors in the field. Engineers and conductors are required to comply with the US Code of Federal Regulations under Title 49, , Subtitle B, Chaper II, Part 240 and 243 which specifies criteria for hiring and certification. Trainers and supervisors are also required to be knowledgeable and competent to ensure that to ensure that each engineer "maintains the necessary knowledge, skill and ability concerning personal safety, operating rules and practices, mechanical condition of equipment, methods of safe train handling (including familiarity with physical characteristics as determined by a qualified Designated Supervisor of Locomotive Engineers), and relevant Federal safety rules." (CFR - 240.123) Thus, these individuals were judged to be competent to evaluate the characteristics of individuals who could safely and efficiently perform the various duties of

Data Analysis

Data was analyzed using descriptive statistics and multiple regression analyses procedures using SPSS version 21. Data were first scanned to determine the presence of missing or partially completed surveys.

Procedures

The survey was administered electronically to all eligible employees of a large transportation organization in a large western state. Study participants received the survey and the consent form approved describing participant anonymity by the IRB. The final version of the survey was electronically distributed to the approximately 459 employees of the transportation division. The number of available employees varied due to the work demands of the business, more customer demand would necessitate more employees. Seasonal fluctuations in business demand affected the number of participants available at any one time.

Data from participants was obtained electronically. Participants were invited to complete the survey instruments following their participation in a previously scheduled training meeting or safety briefing or as they became available for work and signed in for their daily assignment. Participants were approached by a research assistant and invited to participate in the study. Completion of the online instruments took approximately 60 minutes.

Participants work histories were obtained from the human resources department of the large transportation company. The data were housed in a standard database. Participants employee id numbers were used to identify individuals.

Databases from the online instrument administration were matched with the archival human resources database by using employee id numbers. A new combined data base was constructed that contained key work history information and new results from online surveys.

Instruments

Denver Life Style Inventory (DLSI)

This instrument was constructed from items generated from discussions with trainers, job incumbents and job supervisors. The instrument consisted of 90 questions covering a range of work attitudes, values, and work preferences, plus demographic questions. Items wwere responded to using a Likert format response ranging from 1 = To a little or no degree; 2= To a slight degree; 3= To a moderate degree; 4= To a Considerable degree; and 5= To a great degree. Due to the fact that the instrument was newly constructed, and that the purpose of the study was to validate and gather initial data on the instrument, no psychometric characteristics were available at the outset. The DLSI consisted of several clusters of items related to personality characteristics of: Conscientiousness, Concentration, and Adjustment

Differential Aptitude Test (DAT)

The Differential Aptitude Test for Personnel and Career Assessment (DAT for PCA) (Psychological Corporation, 1991). was developed from the original Differential Aptitude Test first published in 1947 revised in 1962 , 1972 1980 and 1990. In 1987 a computerized edition

was developed for personal computers. The DAT for PCA consists of eight subtests, none of which is longer than 20 minutes in length to administer.

Table 1. DAT for PCA Subtests and Administration Time.

Test	# items	Admin Time (Minutes)
Verbal Reasoning	30	20
Numerical Ability	25	20
Abstract Reasoning	30	15
Mechanical Reasoning	45	20
Space Relations	35	15
Spelling	55	6
Language Usage	30	12
Clerical Speed & Accuracy	100	6

The Verbal Reasoning (VR) test is a measure of the ability to understand concepts framed in words. The VR test is aimed at the evaluation of the ability to think constructive and to find commonalties about apparently different concepts. The Numerical Reasoning test (NR) is design to assess the ability to understand numerical relationships and numerical concepts. The Abstract Reasoning Test is a nonverbal measure of reasoning ability. It involves the ability to present relationships between abstract patterns of visually presented material. The AR is designed to assess the underlying principle or relationship in a series of l changing diagrams. The Mechanical Reasoning (MR) test is based on the Bennett Mechanical Comprehension Test (Bennett & Fry,1943). It was designed as an alternate form of the Bennett. Each item consists of a visual presentation of a mechanical situation and a question. The successful test taker will apprehend the underlying principle involved in the pictorial representation. An individual who scores well on BMCT demonstrates an aptitude for learning mechanical skills. The test is also shown to predict training performance (r = .52) and job performance (r = .31) across many jobs and industries. Space Relations (SR) is designed to assess the ability to visualize a three dimensional pattern and to project how the object would look if rotated. Items present various choices and the test taker must select the correct visual representation of the rotation. No other tests from the DAT for PCA were used in this study.

The DAT for PCA is a reliable instrument and the reliability coefficients for the PCA are estimated as follows:

Table 2. DAT Subtest Reliablitites.

Test	Kuder – Richardson Formula #20
Verbal Reasoning	.90
Numerical Reasoning	.88
Abstract Reasoning	.90
Mechanical Reasoning	.90
Space Relations	.92

The DAT for PCA is considered to be a valid test due to the fact that there are a number of studies that have demonstrated the relationship between scores on the test and various aspects of criterion validity and job performance. Selection tests are considered valid if they are able to demonstrate a relationship between scores on the test and some particular criterion. By collecting test scores from the DAT for PCA and demonstrating a correlation between the test and some criterion such as: supervisor ratings, success in training programs job proficiency ratings and the like we can assume that there is a significant relationship between scores on the test and job performance. The DAT for PCA technical manual (Psychological Corporation, 1991) cites over 60 studies involving the DAT and the DAT for PCA demonstrating relationships between test scores and job performance. Thus, the AT for PCA is considered to be a valid predictor of job performance.

AON Workplace Practices Test (WPT)

Aon Hewitt's Workplace Practices Test (WPT) is designed to predict the safety behaviors of job candidates. This test targets the propensity to take risks, impulsiveness, beliefs about controllability of accidents, desire to focus on following rules and do things correctly, general responsibility, and work ethic. The Workplace Practices Test (WPT) requires approximately 30 minutes completing 54 items. This test contains three types of items. The first type requires candidates to indicate the extent to which they agree or disagree with a statement, the second type requires candidates to indicate which answer most reflects what they would do in a particular situation, and the third type requires candidates to indicate which answer best represents their experiences, judgments, and opinions. The WPT consists of three subscales entitled: Safety and rule orientation; Risk aversion and impulse control; and Work Orientation. Reliability and internal consistency data have not been made available. However, the test was validated using a large transportation company and found that a WPT was able to measure

attitudes and validated the selection test with the current workforce. Results demonstrated that first-year candidate injury rates dropped from 14.4% to 4.5% within the first year of implementation. Additionally, short-term and long-term disability expenses were reduced along with workers' compensation costs. A second validation study on a large railroad found that scheduled and hourly transportation and mechanical jobs. Individuals who scored low on the test had 78% more safety incidents on average. Additionally, low scorers had 21% more absences on average than other scorers. It is estimated that applicants who failed the test—using a pass rate of 67%—would have had even more incidents and absences than applicants with low test scores.

Employee Quality Measurement System (EQMS)

The EQMS is a compilation of an employee's job performance indicators. The EQMS is a composite of several key dimensions of job performance including: Start of Service Date; Total Active Years of Service; Overall EQMS Combined Score; Number of Personal Injuries; Number of Human Factor Derailments; Number of Decertifications; Number of Unexcused Absences; Total Incident; Frequency of Incidents; Number of Efficiency tests Failed, and Scores on Rules Exams. These were part of the employee's permanent employment record with the company and were used to evaluate overall performance. The exact formula for calculating the EQMS is proprietary and was not released to the investigators.

Web Neuro Tests

Cognitive assessment tasks were administered through the Web Neuro online instrument platform. (Silverstein, et. al., 2007) The assessment batter consists of a number of different neuropsychological tests that have been utilized for nearly 75 years in different settings. The current battery has been normed on both normal and clinical populations assessing cognitive functioning, memory, executive function and emotional recognition.

WebNeuro consists of the following domains of cognitive function: sensorimotor, memory, executive planning, attention, and emotion perception (social cognition). The battery includes a number of measures:

Sensorimotor Domains Simple motor tapping test. For this test participants tap the space bar on the keyboard quickly for 60 sec. The dependent variable is total number of taps with the dominant hand.

Choice reaction time test. Respondents view the computer screen as one of four target circles is randomly presented over a series of trials. For each trial, the respondent is required to use the mouse and click on the illuminated circle as quickly as possible following presentation. The dependent measure is the mean reaction time across trials.

Attention Domain Digit span test. Respondents were presented with a series of digits on a computer screen, separated by a one second interval. The respondent is immediately asked to enter the digits using the mouse. In the first part of the test, respondents are required to recall

the digits in forward order. In the second part, they are required to recall them in reverse order. In each part, the number of digits in each sequence is gradually increased from 3 to 7, with two sequences at each level. The dependent measure for each part is the maximum number of digits the respondent recalled without error.

Continuous Performance Test. A set similar looking letters (B, C, D, or G) are presented to the respondent on the computer screen (for 200 msec), separated by an interval of 2.5 sec. If the same letter appears twice in a row, the respondent is required to press the space bar. Speed and accuracy of response are equally stressed in the task instructions. There are 125 stimuli presented in total, 85 being non target letters and 20 being target letters (i.e., repetitions of the previous letter). The dependent variables are the number of errors of omissions and false positives.

Executive Function Domain Switching of attention test. This task is a computerized adaptation of the "Trail Making Test" Part B (Reitan, 1958). The respondent is presented with a pattern of 13 numbers (1–13) and 12 letters (A–L) on the screen and is required to click inside the appropriate circles for numbers and letters alternatively in ascending sequence (i.e., 1 A 2 B 3 C . . .). As each number or letter is clicked in correct order, a line is drawn automatically to connect it to the preceding number or letter in the sequence. This allows the respondent to visualize the path touched. This task tests the ability of the respondent to switch attention between mental tasks, in this case number and letter sequence checking, and thereby alternate between the respective mental sets induced. The dependent variable is time to completion.

Verbal interference test. This task taps the ability to inhibit automatic and irrelevant responses and has similarities to the Stroop task (Golden, 1978). The respondent is presented with colored words, one at a time. Responses are made on the screen by clicking the mouse on the appropriate word on the response pad. The dependent variable in each part is the number of words correctly identified.

Maze test. This is a standard maze problem where the respondent uses the arrow keys on the keyboard to navigate around the grid. The dependent variable is the total number of errors.

Table 1 Cognitive and emotion tasks

Task name	Assesses	Test duration	Measures of task performance
General cognition tasks			
Switching of attention	Information processing speed	3 minutes	Completion time
Verbal interference	Cognitive flexibility	2 minutes	Reaction time – incongruent trials
Continuous performance test	Attention	6 minutes	Reaction time
Go/no-go	Response inhibition	4 minutes	Reaction time
Digit span	Working memory	6 minutes	Total number of trials correct
Motor tapping	Motor coordination	1 minute	Number of taps – dominant hand Number of taps – non-dominant hand
Maze	Executive function	10 minutes	Completion time Number of errors Number of overrun errors
Emotional cognition tasks			
Emotion identification	Emotion identification accuracy Speed of emotion identification	4 minutes	Number of correct responses Reaction time for correct responses

Figure 4. Adapted from (Gordon et. al, 2015).

Results

Participants

A total of 459 participants, this included all of the operating employees in the southern region of a large transportation company. The group consisted of 98% males and 64.19% were 40 over. With respect to length of time on the job, the average was 14.8 years, a fairly seasoned workforce, but ranging from .2 years to 45 years. The racial and ethnic composition of the sample is also displayed in Table 4 and shows that the majority of the participants were Caucasian (59%), followed by Hispanic (12.5%), Latino (0.28%), Asian (0%), American Indian (1.3%) and Multi-racial (2.23%).

EQMS Scores

The EQMS scores were obtained from the historical data available on study participants. Interestingly, the EQMS score, which are a composite of all indicators, was not significantly correlated with personal injuries. The EQMS score combines the various safety indices available to a railroad employer on safety. The key indicators of personal injury, derailments, decertification, absenteeism, total number of incidents and frequency of incidents. Table 3 shows the correlation between the predictors and the accident injury (criterion) summary variables.

Table 3. Correlations between predictor and criterion measures

	Derailments	Decerts	Absenteeism	Total Incident Count	Frequency Index	EQMS Score
Personal Injuries	-.037	.114*	-.065	.821**	.293**	-.007
Derailments		-.028	.058	.276**	.484**	-.321**
Decertification			.065	.349**	.231**	-.534**
Absenteeism				.383**	.611**	-.200**
Total # Incidents					.686**	-.306**
Frequency Index						-.347**

Table 4. Characteristics of study participants.

Variable	Result
Gender	
Male	98%
Female	2%
Age	
18 to 29	9.37%
30-39	26.45%
40-49	35.26%
50-59	22.87%
60 or older	6.06%
Over 40	64.19%
Race	
Caucasian	59.89%
Hispanic	12.53%
African- American	23.68%
Asian	0.00%
American Indian	1.39%
Multi-racial	2.23%
Education	
Less than High School	0.83
High School	26.17
Some College	52.62
College Degree (BA/BS)	15.98
Some Graduate School	1.93
Graduate Degree	2.48
Tenure	
Mean	14.8
Mode	7.4
Min	0.2
Max	45.8
Median	13.15

Results of the regression analysis of the measures in the assessment battery indicate that there were only three measures that have a significant correlation with the EQMS Score: Spatial Reasoning (r=-.240, p<.014), not preferring midnight work (r=-.093, and having a preference for competitive fringe benefits (r=.251, p<.001). (See Table 5). Multiple regression analysis also revealed that only Spatial Reasoning makes a sufficiently strong contribution to the overall variance to warrant its inclusion in a predictive equation. Further analysis of the other performance measures with the EQMS score shows that the other measures have significant relationships with EQMS with the exception of Personal Injuries. Given that the correlations between the other measures and EQMS are low it will make more sense conceptually to develop prediction equations, and therefore selection batteries that predict the individual performance measures.

Table 5. Correlations between predictive measures and EQMS scores.

	EQMS	P<	N
Differential Aptitude Test			
Verbal Reasoning	ns		83
Mechanical Reasoning	ns		83
Spatial Reasoning	-.240	.014	83
AON			
Risk Aversion	ns		192
Safety Orientation	ns		192
Work Orientation	ns		192
Denver Life Style Inventory			
q12. Starts quarrels with others...	ns		338
q25. Is inventive...	ns		338
q51. It's inevitable, I will probably get injured or have an accede	ns		338
q62. I like working midnights	-.093	.044	338
q74. Competitive fringe benefits are important to me in my job…	.251	.001	328

No significant correlations were obtained between any of the measures on the Web Neuro Battery and the EQMS scores. However it should be noted that there were only n=49 observations for some of the correlations.

Personal Injuries

Looking at the specific subcomponents of the EQMS index provides some additional revealing insights. A number of significant correlations were obtained between scales on the Denver Lifestyle Inventory (DSI) and the personal injury indicator. As can be seen in Table 6 there were quite a few correlations. The combined summary scores of Conscientiousness and Neuroticism were also significant predictors of the scores on the Personal Injuries Scale.

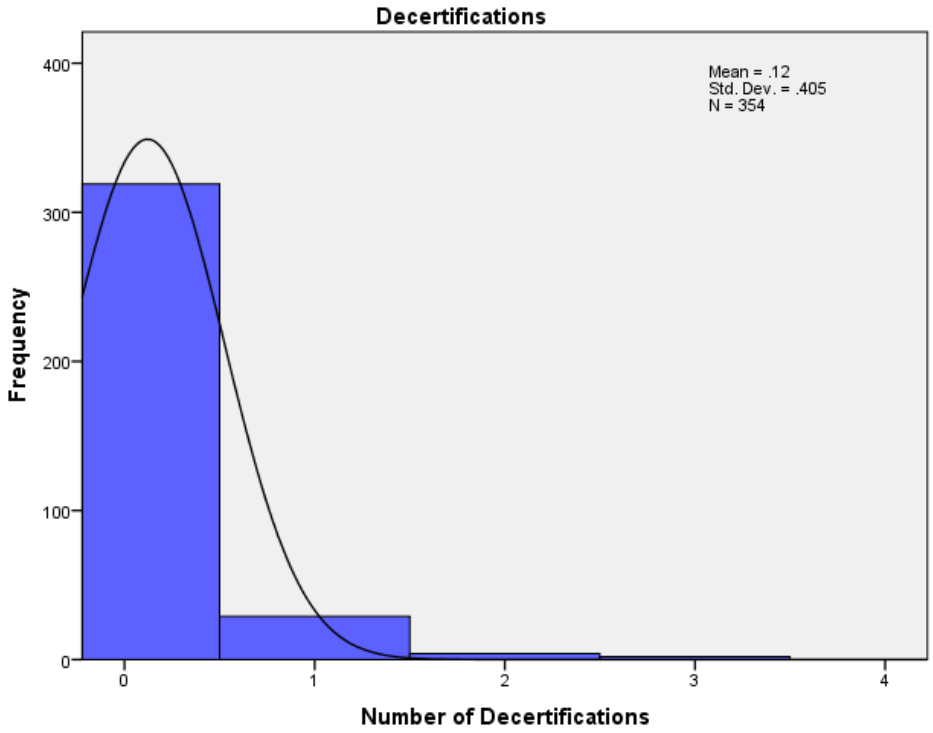

Figure 5. Frequency of number of Personal Injuries for railroad employees.

Table 6. Correlations between predictive measures and number of Personal Injuries (N=338).

	Personal Injuries
q4. Is depressed, blue...	.115*
q13. Is a reliable worker...	-.099*
q14. Can be tense...	.095*
q19. Worries a lot...	.106*
q29. Can be moody...	.138**
q38. Makes plans and follows through with them...	-.121*
q56. I believe that I will be able to prevent an accident from occurring to me.	.148**
q60. I have no problem with working 14 days straight.	.113*
q62. I like working midnights	.125*
q63. My spouse or partner is ok with me working midnights.	.098*
q65. I prefer to lay off from work for family events if they are important.	-.155**
q66. My family resents the amount of time I am away from home.	-.171**
Personality Measures	
Conscientiousness (Summary Score)	-.090*
Neuroticism (N7) (Summary Score)	.145**

It should be noted that the number of Personal Injuries scores increase with increasing longevity in the position. In other words, there is a risk associated with simply being on the job longer that is associated with a higher number of personal injuries. Consequently, this measure may have some as yet undetermined confounds. The gradually increasing line in Figure 6 shows that the number of injuries increases at an increasing rate over time from 0-5 years of employment to greater than 20 years of employment broken down by 20 percent intervals of frequency of years of employment. Clearly there is a connection between the two and as the person continues their employment the risk of exposure to workplace hazards increases and thus the increased number of injuries recorded. However, not all of those persons with a long tenure reported injuries so while there may be a connection and even a correlation there may also be other factors, such as distractibility and concentration and conscientiousness that contribute to the lower accident rates in some individuals.

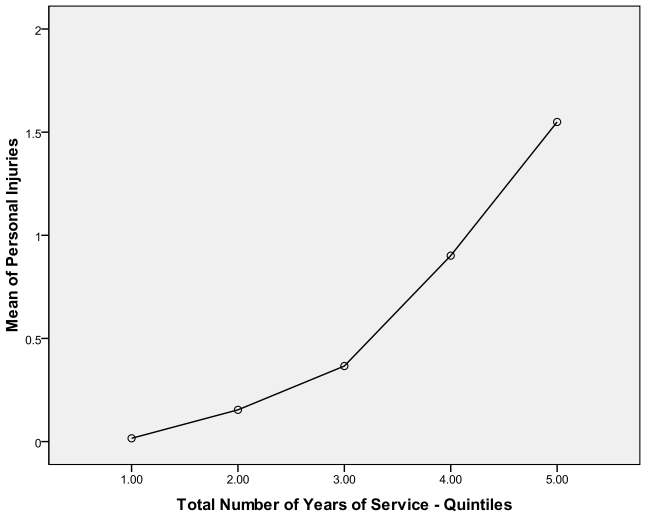

Figure 6. Personal Injuries and Job Tenure.

Several significant correlations were obtained between the Web Neuro Battery – Total Immediate Recall test (a memory recognition test summing the total number of recognized words over three trials) and the Personal Injury Scores. In addition there were significant correlations between successful completion of the Maze test on the Web Neuro inventory and number of personal injuries (Maze learning time` or time taken to the end of the final trial with at least one error- just before successfully completing the maze without error twice consecutively). However it should be noted that there were only n=51 observations for some of the correlations.

Also of interest, the combined set of variables, entered simultaneously into a regression equation does result in a significant prediction of personal injuries (R^2 = .083, F(14, 323)= 2.084, p<.012). The most significant predictors in the set of variables appear to be related to family concern for the employee and the presence of safety locus of control (items 66 and 56 respectively). Based on these results it is recommended that a Conscientiousness and Neuroticism measure be included in the selection assessment battery.

Decertifications

The Decertification score was correlated with a number of the other key predictors in the assessment battery. Decertifications occur when an individual breaks an operating rule such as running a red signal, which is like running a red light in a vehicle. Decertifications for this group occur Table 7 shows the results of the correlations. As can be seen in Figure xx, Decertifications occur relatively infrequently, but they are serious and can lead to immediate dismissal from the railroad.

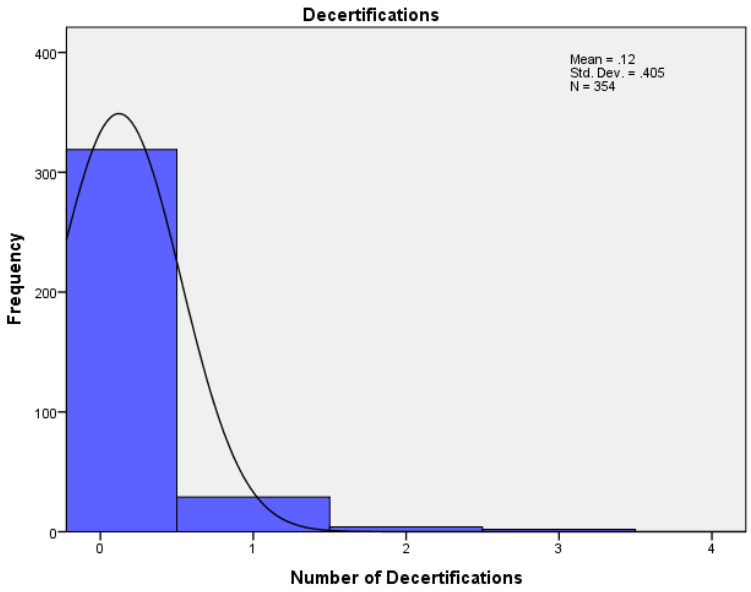

Figure 7. Dercertificaitons for employees in the study.

Table 7. Correlatnions between assesment measures and Decertifcations.

	Decerts	P<	N
Differential Aptitude Test			
Verbal Reasoning	.263**	.008	83
Mechanical Reasoning	.262**	.008	83
Spatial Reasoning	.421**	.000	83
AON			
Risk Aversion	.078	.141	192
Safety Orientation	.146*	.022	192
Work Orientation	.024	.372	192
Denver Life Style Inventory			
q12. Starts quarrels with others...	-.108*	.023	338
q25. Is inventive...	.094*	.042	338
q51. its inevitable, I will probably get injured at … work.	.104*	.027	338
q62. I like working midnights	.139**	.005	338
q74. Competitive fringe benefits are important to me in my job…	-.145**	.004	328

All of the measures in Table 7 are significantly related to the occurrence of Decertifications. Interestingly, when we look at persons who have only been working for 8 years or less the correlation between spatial relations (r=.309, p<.02), safety orientation (r=.205, p<.03), is significant. The relationships are nearly doubled in the group of workers who have worked more

than 19 years suggesting that these factors influence job performance in even experienced personnel.

The results are less clear with respect to the relative or comparative effectiveness or predictive power of the measures due to the fact that there fewer subjects to study as not all participants completed all of the measures. The results of the regression analysis are shown in Table 8, and indicate that there is a significant overall prediction of the occurrence of Decertifications (R^2 =.345, df = 36, F(6,30) = 2.637, p<.036). However, the variable that contributes the most is Spatial Relations (t=3.502, p<.001).

Table 8. Regression analysis of predictors on to Decertifcations.

Model		Standardized Coefficients Beta	t	Sig.
1	(Constant)		.177	.860
	SR – DAT	.535	3.502	.001
	Risk A – AON	-.123	-.616	.543
	Safety O – AON	.054	.291	.773
	q12. Starts quarrels with others...	-.238	-1.249	.221
	q51. I will probably get injured	-.130	-.735	.468
	q62. I like working midnights	.144	.924	.363

Thus, it is recommended that an assessment battery that would be used to select employees with a lower likelihood of having accidents such as a decertification would have a Spatial Relations Reasoning test be included in a selection battery.

Derailments

A number of measures in the assessment battery were significantly associated with the number of derailments reported in the work history. As can be seen in Table 9 the majority of these indicators were derived from the Denver Life style Inventory (DLSI). In the interest of simplicity and to reduce the number of variables depicted to a more manageable number, the correlation matrix is omitted here and only the results of the regression analysis are presented. Results of the stepwise analysis of variance resulted in a six-step procedure which identified six main variables that were significantly predictive of the occurrence of derailments. These predictors were largely behavior and attitudinal and consist of items apparently reflecting physical and mental distractibility and activity, possibly indicating a characterological pattern of

lack of concentration or distractibility. In addition, the results also identify the impact of family considerations. However, taken together these measures accounted for nearly 10% of the variance in reported derailments (R2=.091, F(6,331)=5.512, p<.000).

Also of interest is the fact that there is a significant difference in the number of derailments depending upon the length of time with the railroad. Derailments increase after the first five years and then decline gradually. (see Figure 8). The graph shows a slight rise in derailments as the individual approaches 30 years on the job. This data reflects two issues. First, persons who have a derailment during the first few years of working are unusual. This may reflect the lack of opportunity due to lack of seniority or other factors.

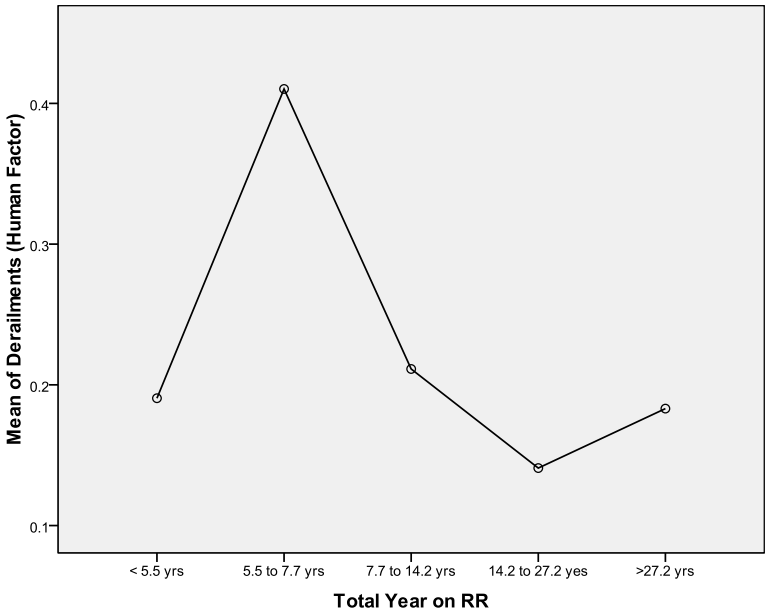

Figure 8. Derailments over time of employment.

Nevertheless, the assessment selection battery data are fairly robust with statistically significant results and large N's. The results suggest that personality factors may play a fairly large role in the prediction of these events.

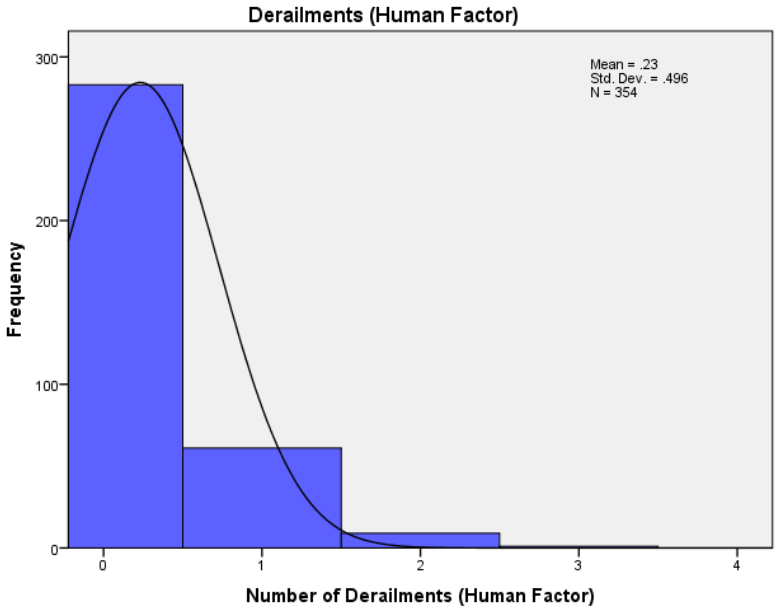

Figure 9. Frequency of Derailments in the sample.

Table 9. Regression analysis predictors of Derailments in work history.

Model		Beta	t	Sig.
	(Constant)		3.692	.000
	q43. Is easily distracted...	.107	1.908	.057
	q63. My spouse or partner is ok with me working midnights.	-.169	-3.066	.002
	q65. I prefer to lay off from work for family events	-.129	-2.366	.019
	q1. Is talkative...	-.113	-2.141	.033
	q42. Likes to cooperate with others...	-.117	-2.112	.035
	q41. Has few artistic interests...	.111	2.080	.038

Table 10. Significant Correlations of Web Neuro Measures with Derailments

Web Neuro Results	Derailments
Digit Span Trials Correct	.293*
Digit Span Forward	.241*
Attention and Concentration - Reaction Time	-.343**
Info Processing Efficiency - Switching	-.291*
Choice Reaction Time	-.264*
Maze Completion Time	-.249*
Maze – Path Learning Time	-.244*
Emotion Recognition - Fear Faces	.257*
Delayed Emotion Recognition – Avg Response Time	-.254*

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).

A correlational analysis of the Web Neuro subtests is shown in table 10. When these variables are compared to each other and entered simultaneously in the regression analysis only two of the measures have sufficient strength to warrant inclusion in the final stepwise equation, namely Reaction Time ($t=-2.88$, $p<.006$) and Digit Span Trials Correct ($t=2.56$, $p<.014$). Again, both of these are measures of distractibility and concentration. Taken together these Web Neuro Measures suggest a strong distractibility component affecting the occurrence of Derailments. Therefore, it is recommended that the assessment selection battery include a personality measure of distractibility. The Web Neuro Digit Span test could be used but its practicality will be an issue.

Absenteeism

The following table shows the results of the correlation of the various measures with Absenteeism. The majority of the measures reaching significance were from the DLSI.

Table 11. Correlations of measures with Absenteeism.

	Absenteeism
q2. Tends to find fault with others...	-.103*
q3. Does a thorough job...	-.095*
q13. Is a reliable worker...	-.175**
q23. Tends to be lazy...	.218**
q46. All accidents are preventable.	-.107*
q50. Inform their co-workers if they believe that work is being done unsafely.	.090*
q51. its inevitable, I will probably get injured or have an accident at work.	.115*
q60. I have no problem with working 14 days straight.	.131**
q66. My family resents the amount of time I am away from home.	-.114*
q84. A job that has activities that involving ideas and problem solving.	-.093*
Conscientiousness	-.093*

*. Correlation is significant at the 0.05 level (1-tailed).

**. Correlation is significant at the 0.01 level (1-tailed).

Results of the regression analysis of these variables produced a significant equation with four measures from the DLSI. These measures together produced a significant equation (R^2 =.106, $F(4,325) = 9.65$, $p<.001$) that accounted for a total of 11% of the variance in Absenteeism as recorded in the history file.

Table 12. Regression of assessment battery measures on to Absenteeism.

Absenteeism.		Standardized Coefficients	t	Sig.
		Beta		
	q23. Tends to be lazy...	.243	4.550	.000
	q2. Tends to find fault with others...	-.166	-3.075	.002
	q51. its inevitable, I will probably get injured or have an accident at work.	.144	2.693	.007
	q66. My family resents the amount of time I am away from home.	-.137	-2.546	.011

These results suggest that there is a significant association between the occurrence of absenteeism and general personality characteristics of reliability and conscientiousness as measured by the four items that were significant in the regression analysis. However, the relationships are not as robust as the other measures suggesting that we are tapping a fairly mature workforce who likely have learned to show up for work. Most likely, these measures would be more successful with new-hires as persons low in conscientiousness would have been eliminated from the workforce by the time the sample was taken. Therefore, it is recommended that a measure of conscientiousness including the items listed above be included in the assessment battery.

Efficiency Testing

It is standard practice, mandated by the railroad safety improvement act, that operations or efficiency testing be conducted on an annual basis and that the results of that testing be maintained by the railroad company and available for inspection. (CFR › Title 49 › Subtitle B › Chapter II › Part 239 › Subpart D › Section 239.301). These tests contribute to the efficiency of rail operations and help to ensure safe operating practices. The results of the operational testing program were also maintained in the employee database and were available for analysis and examination of likelihood of being predicted by the assessment measures included in the assessment battery.

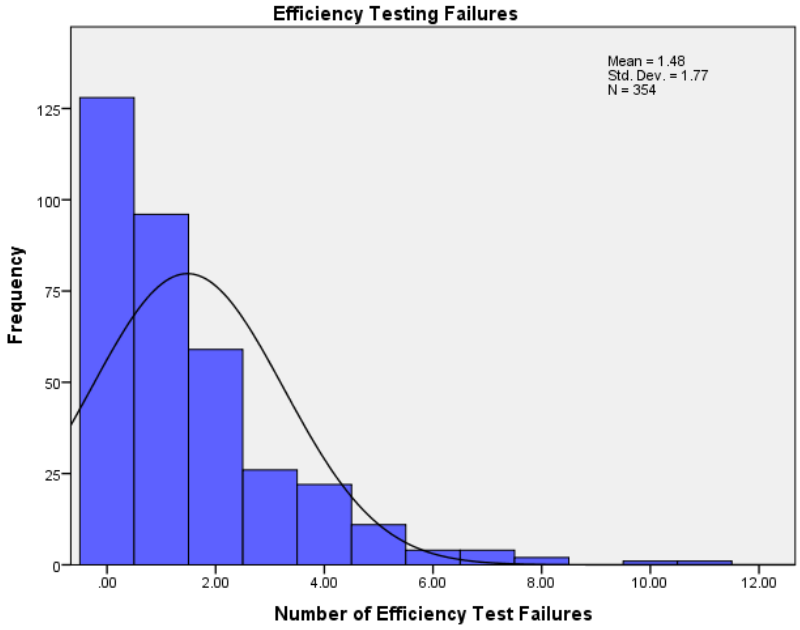

Table 13. Correlation between assessment measures and number of efficiency test failrures.

	Efficiency Testing Failures
Differential Aptitude Test	
Verbal Reasoning (N=82)	-.231*
Mechanical Reasoning (N=82)	-.231*
Spatial Reasoning (N=82)	-.134
AON	
Risk A (N=61)	0.04
Safety Orientation (N=61)	0.043
Work O (N=61)	-0.132
Personality	
Conscientious (N=338)	-.141**
Extraversion (N=338)	-0.062
Emotionality (N=338)	0.053
** Correlation is significant at the 0.01 level (1-tailed).	
* Correlation is significant at the 0.05 level (1-tailed).	

Attempts to conduct regression analyses were not possible due to the overall sample size. While there were sufficient size to conduct some of the test when all of the variables were entered into the equation fewer than 20 cases were available for whom all of the predictors were available.

Shift Work and Safety

Results of the analysis of the measures included in the assessment battery that determine intermodal employees reactions to characteristics of shift work on health and safety are included in Table xx below. Ten items were constructed which were related to various aspects of shift work I the operational environment. These included working consecutive shifts, working away from home, working midnight hours and impact of shift work on family & social activities. Results indicate that a lack of tolerance of consecutive work days was negatively correlated with several safety indices (q64 – r=-.153, p<.001). Also, being uncomfortable working midnights was also associated with negative safety outcomes such as Injuries, Decertifications and Total Number of Incidents (q62 – r=-. 143, p<.001).

Table 14. Assessment of shift characteristics and safety outcomes.

	Injuries	Derailments	Decerts	Absences	Total Incidents	Efficiency Test Failures	EQMS Score
q58. My family is able to accommodate and adjust to the amount of time I have to work away from home.	0.045	-0.082	0.045	0.082	0.056	-0.065	-0.064
q59. My family would be able to manage successfully if I had to work 14 consecutive days without a day off.	0.075	-0.042	0.025	0.029	0.065	-0.068	-.091*
q60. I have no problem with working 14 days straight.	-.113*	0.032	-0.072	-.131**	-.153**	.095*	0.085
q61. I have been satisfied with jobs that have required me to be away from home for extended periods of time.	0.04	-0.06	0.025	0.005	0.022	-0.076	-0.062
q62. I like working midnights	-.125*	0.046	-.139**	-0.053	-.143**	-0.043	.093*
q63. My spouse or partner is ok with me working midnights.	-.098*	.093*	-.093*	-0.065	-.099*	-0.009	0.05
q64. I am comfortable choosing to work over attending family events on a regular basis.	-0.069	0.017	-0.074	-0.086	-.104*	0.044	0.081
q65. I prefer to lay off from work for family events if they are important.	-.155**	-.111*	-0.087	-0.027	-.191**	0.059	0.077
q66. My family resents the amount of time I am away from home.	-.171**	0.046	-0.084	-.114*	-.191**	0.01	0.075
q67. I can only tolerate being away from home for a day or two.	-0.041	-0.02	-0.04	-0.067	-0.076	-0.005	.106*

* Correlation is significant at the 0.05 level (1-tailed).

** Correlation is significant at the 0.01 level (1-tailed).

Summary of Findings

1. 486 total study participants
    a. 346 completed the Denver Life Style Questionnaire
    b. 354 on whom the EQMS Scores were available
    c. 194 AON assessments
    d. 92 Differential Aptitude Tests
    e. 50 Web-Neuro tests
2. EQMS scores component measures not correlated with personal injuries
    a. EQMS correlated with Derailments, Decertifications, Absenteeism, Total Number of Incidents, Frequency Index and Number of Efficiency Test Failures.
3. EQMS scores significantly correlated with Spatial Reasoning
4. Personal Injuries significantly correlated with personality measures of Conscientiousness and Neuroticism
5. Decertifications were significantly correlated with DAT-Spatial Relations and AON-Safety Orientation ($r=.205$, $p<.03$), is significant.
6. Derailments were significantly correlated with behavior and attitudinal predictors reflecting physical and mental distractibility and activity, possibly indicating a possible pattern or individual characteristic of distractibility or lack of concentration.
7. Distractibility significantly correlated with Reaction Time and Digit Span Trials Correct
8. Absenteeism significantly correlated with four items from the DLSI

Discussion

An attempt to define and measure an a set of measures that would be associated with employees safe work performance was attempted. A battery of assessment measures deemed to measure the skills and abilities and characteristics associated with the tasks of the job of railroad engineer and conductor was constructed. The measures in the battery assessed the personality, attitudes, and cognitive skills and abilities of the study participants. Results indicated a number of significant relationships between the measures and the criterion measures which purported to measure job safety.

The important finding from this study is that there are significant relationships between psychological measures and historical indices of job performance. These statistically significant relationships are sufficient to demonstrate a correlation between psychological testing and job performance. Thus, for all practical and scientific purposes the psychological measures can be considered valid indicators of job performance. From the standpoint of the EEOC guidelines, these tests then have been validated for use in the workplace. Additional testing, to demonstrate that accommodations and considerations have been made to reduce adverse impact are needed, but, the tests have been validated. Adverse impact is a separate, and very important issue, that can be addresses subsequently. However, it should be noted that in most cases no adverse impact has been found when using attitudinal and personality measures.

The components of the successful battery will likely include both cognitive, attitudinal and personality measures. Results of the assessment showed that verbal reasoning and spatial reasoning were predictive, as were conscientiousness, and other attitudes towards the work. While all three did not consistently predict each of the work criterion measures (i.e. Derailments, Decertifications, etc.) However,

Study Limitations

Limitations of this study are that despite the large sample size missing data contributed to a lack of power with some of the instruments. The other issue was that this is not a predictive validity study but rather a concurrent validity study conducted on job incumbents. The utilization of job incumbents, and in this case they had been in place for many years, may create a bias in the results in that the person who may have scored differently on the selection measures have been eliminated from the workforce. Thus the sample and these results may suffer from restriction of range. Such a finding can under represent the true relationship between the characteristics identified in pre-employment screening and those actually on the job. *In other words, the true validity coefficients would probably be much higher, between the pre-*

employment testing and the criterion measures. In real life, those persons who are not performing adequate are usually terminated before statistics can be developed which show their elimination and the relationship between the test scores and the job performance. In addition, those persons who are underperforming are not kept in the sample so that when the study is done there are very few underperformers remaining in the workforce therefore liming the range of score that can be predicted from the pre-employment assessment measures.

Conclusion

In conclusion, this study has described the development and potential use of a measure of corporate safety culture. The instrument can be used in any corporate environment, but has most technical relevance for transportation focused entity. The instrument has demonstrated psychometric properties of reliability and validity. Moreover, it has been show to measure characteristics of corporate culture which differentiate employees who have received recognition for safe work behaviors or accidents and those who have not.

References

Baba, V., Tourigny, L., Wang, X. and Liu, W. (2009). Proactive personality and work performance in China: The moderating effects of emotional exhaustion and perceived safety climate. *Canadian Journal of Administrative Sciences*, 26, 1, 23–37.

Barrick, M. R. Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta–analysis. *Personal Psychology*, 1991؛ 44, p: 1-26

Barron, L. G. ; Carretta, T.R. ; & Rose, M. R. (2016). Aptitude and Trait Predictors of Manned and Unmanned Aircraft Pilot Job Performance. *Military Psychology*, 2016

Bennett , G.K. & Fry, D.E. (1943). Mechanical Comprehension Test. The Psychological Corporation, New York,

Bennett, G. K., & Wesman, A. G. (1947). Industrial test norms for a southern plant population. *Journal of Applied Psychology*, 31(3), 241-246. doi:10.1037/h0056062

Cellar, D.F., et al. (2001). The Five-Factor Model and Safety in the Workplace: Investigating the Relationships Between Personality and Accident Involvement. *Journal of Prevention & Intervention in the Community*, 22: 43-52.

CFR › Title 49 › Subtitle B › Chapter II › Part 239 › Subpart D › Section 239.301). https://www.law.cornell.edu/cfr/text/49/239.301

Christian, M. S.; Bradley, J. C.; Wallace, J. C.; Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology*, Vol 94(5), 1103-1127.http://dx.doi.org/10.1037/a0016172

Clarke, S. & Robertson, I. (2008) An Examination of the Role of Personality In work Accidents Using Meta-analysis. *Applied Psychology*, 2008, Vol.57(1), pp.94-108

Clarke, S. and Ivan T. Robertson, I.T. (2005) A meta-analytic review of the Big-Five personality factors and accident involvement in occupational and non-occupational settings. *Journal of Occupational and Organizational Psychology*. 78, 3, 355–376.

Colorado Department of Transportation (2011). CDOT. http://www.coloradodot.info/about

Colvin, H. and Taylor, R. (2012). Building a Resilient Workforce: Opportunities for the Department of Homeland Security: Workshop Summary. Planning Committee on Workforce Resiliency Programs; Board on Health Sciences Policy; Institute of Medicine.

Costa, P.T. Jr. & McCrae, R.R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, FL: Psychological Assessment Resources.

De Vellis, R. F., & Dancer, L. S. (1991). Scale development: theory and applications. *Journal of Educational Measurement*, 31(1), 79-82.

Gatewood, R.D, Field H. (1994). Human Resource Selection, 3rd Edition, Fort Worth Harcourt Brace.

Geller, E. S., & Wiegand, D. M. (2005). People-based safety: Exploring the role of personality in injury prevention. *Professional Safety*, 50(12), 28-36. Retrieved from http://du.idm.oclc.org/login?url=http://search.proquest.com.du.idm.oclc.org/docview/200369783?accountid=14608

Golden, C. J. (1978). *Stroop color and word test*. Wood Dale, IL: Stoelting Co.

Gordon, E, Rush A.J., Palmer, D.M., Braund, T.A., Rekshan, W. (2015). Toward and online cognitive and emotional battery to predict treatment remission in depression. *Neuropsychiatric Disease and Treatment*, 11: 517-531. https://www.dovepress.com/toward-an-online-cognitive-and-emotional-battery-to-predict-treatment--peer-reviewed-fulltext-article-NDT

Gordon, E., Cooper, N., Rennie, C., Hermens, D., Williams, L.M. (2005). Integrative neuroscience: the role of a standardized database. *Clinical EEG Neuroscience*. 36(2):64–75.

Hewitt, A. (2013). *Hiring for a Safer Workplace*. AON Corporation. http://www.aon.com/attachments/human-capital-consulting /2013_ Hiring_ for_ a_Safer_Workplace.pdf

Hormann, HJ. & Maschke, P. (1996). On the Relation Between Personality and Job Performance of Airline Pilots. *The International Journal of Aviation Psychology*. 6, 2.

Hurtz, G.M.; Donovan, J. J. (2000) Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85(6), 869-879.

Kan'Ani, K., Hor, M., Hadi, S., & Taiiebi, P. (2014). A study on personality features of accident agents through applying NEO-PI-R questionnaire. *Interdisciplinary Journal of Contemporary Research In Business*, Vol.6(1), pp.87-98.

National Highway Traffic Safety Administration (2008). National Motor Vehicle Crash Causation Survey: Report to Congress. US Department of Transportation. DOT HS 811 059.

NHTSA, (2008). National Motor Vehicle Crash Causation Survey Report to Congress. DOT HS 811 059, July 2008.

NIOSH (National Institute of Occupational Safety and Health). (2009) Causal pathways affecting worker safety and health. http://www.cdc.gov/niosh/programs/workorg/

Norris F., Stevens S., Pfefferbaum B., Wyche K., & Pfefferbaum R. (2008). Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *American Journal of Community Psychology*. 41(1):127–150.

Paul, R.H., Gunstad J., Cooper, N., et al. (2007). Cross-cultural assessment of neuropsychological performance and electrical brain function measures: additional validation of an international brain database. *International Journal of Neuroscience.* 117, (4) :549–568.

Paul, R.H., Lawrence, J., Williams, L.M., Richard, C.C., Cooper, N., Gordon, E. (2005). Preliminary validity of "integ-neuro": a new computerized battery of neurocognitive tests. *International Journal of Neuroscience.* 115(11):1549–1567.

Psychological Corporation, (1991). *The Differential Aptitude Tests for Personnel and Career Advancement*. New York: Harcourt Brace Jovanovich, Inc.

Rail Safety Improvement Act (RSIA), 49 CFR , §, 20156, (2008).

Reitan, R.M. (1958). The validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, pp. 271–276.

Sherry, P. (1991). Person-environment fit and accident prediction. *Journal of Business and Psychology*, 5, 3, 411-416. doi:10.1007/BF01017711

Sherry, P. (1993). *Validation of Selection Procedures for Railroad Dispatchers*. Technical Report. Intermodal Transportation Institute: University of Denver, March 16, 1993.

Silverstein, S.M., Berten, S., Olson, P., Paul,R., Williams, L., Cooper, N., & Gordan, E. (2007). Development and validation of a World-Wide-Web-based neurocognitive assessment battery: WebNeuro. *Behavior Research Methods*, Vol.39(4), pp.940-949.

Tabachnick, B. G., & Fidell, L. S. (2001). Using Multivariate Statistics. Boston: Allyn and Bacon.

U.S. Department of Labor. (2010). U.S. Bureau of Labor Statistics, U.S. Department of Labor. Retrieved from http://www.bls.gov/iif/oshwc/cfoi/ cfch0007.pdf

Williams, L.M., Mathersul, D., Palmer, D.M., Gur, R.C., Gur, R.E., Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*. 31(3):257–277.

Williams, L.M., Mathersul, D., Palmer, D.M., Gur, R.C., Gur, R.E., Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*. 31(3):257–277.

Table 15. Correlation between assessment instrument scores and criterion measures.

	Personal Injuries	Derailments (HF)	Decertifications	Absenteeism	Total Incident Count	Frequency Index	Efficiency Test Failures	EQMS Score
Differential Aptitude Test (N=82)								
Verbal Reasoning	0.114	0.048	.262**	0.013	.188*	0.117	-.231*	-0.088
Mechanical Reasoning	0.018	-0.024	.258**	0.02	0.083	0.086	-.231*	-0.029
Spatial Reasoning	-0.016	0.053	.419**	0.07	0.139	.226*	-0.134	-.232*
AON (N=61)								
Risk Aversion	-0.112	-0.207	0.036	-0.054	-0.168	-.225*	0.04	0.066
Safety Orientation	-0.077	-0.148	0.156	0.119	-0.02	-0.071	0.043	-0.128
Work Orientation	-0.059	-0.072	0.049	-.218*	-0.156	-.312**	-0.132	0.046
Personality Scales (N=338)								
Conscientiousness	-.090*	-0.056	0.000	-.093*	-.128**	-.091*	-.141**	0.036
Extraversion	0.011	-.132**	-0.023	-0.06	-0.063	-.108*	-0.062	0.087
Emotionality	.147**	0.023	-0.049	0.037	.126*	0.031	0.053	0.041

** Correlation is significant at the 0.01 level (1-tailed).

* Correlation is significant at the 0.05 level (1-tailed).