THE TREE TO SUCCESS: USING DECISION TREE ANALYSIS TO PREDICT EMPLOYEE PERFORMANCE AND HEALTH

by

Justin M. Raines

December, 2015

Director of Dissertation: Dr. Mark C. Bowler

Major Department: Department of Psychology

The purpose of the current study was to evaluate the relative advantages and disadvantages of a novel statistical technique within the context of personnel decision-making. Specifically, the use of decision tree analysis was examined regarding its potential benefits over binary logistic regression. Using Monte Carlo simulation, a series of data sets were generated based on meta-analytic correlation matrices representing the topics of (1) employee performance and (2) employee health. Each data set was analyzed via both decision tree analysis and binary logistic regression with subsequent comparisons being made concerning model validity, adverse impact, and interpretability. Overall, decision tree analysis demonstrated a variety of benefits over the more traditional method. In general, decision tree analysis produced predictive models that possessed nearly equivalent levels of validity as models produced by logistic regression. Of greater importance, the majority of decision tree analysis models produced no adverse impact, whereas logistic regression models were largely associated with discriminatory results. Lastly, decision tree analysis models were generally more parsimonious and interpretable than the competing logistic regression models. The practical implications of these results are discussed and suggest that the use decision tree analysis holds the potential to greatly improve the way in which organizations make decisions regarding the productivity and health of employees.

THE TREE TO SUCCESS: USING DECISION TREE ANALYSIS TO PREDICT EMPLOYEE PERFORMANCE AND HEALTH

A Dissertation

Presented to the Faculty of the Department of Psychology

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Health Psychology

by

Justin M. Raines

December, 2015

THE TREE TO SUCCESS: USING DECISION TREE ANALYSIS TO

PREDICT EMPLOYEE PERFORMANCE AND HEALTH

by

Justin M. Raines

APPROVED BY:

DIRECTOR OF
DISSERTATION: _____
Mark C. Bowler, PhD

COMMITTEE MEMBER: _____
Shahnaz Aziz, PhD

COMMITTEE MEMBER: _____
Karl L. Wuensch, PhD

COMMITTEE MEMBER: _____
Hamid Fonooni, PhD

CHAIR OF THE DEPARTMENT
OF PSYCHOLOGY: _____
Susan L. McCammon, PhD

DEAN OF THE
GRADUATE SCHOOL: _____
Paul J. Gemperline, PhD

This dissertation is dedicated to the memory of my hero and grandfather, Clifford Campbell.

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the guidance and support of my dissertation chair, Dr. Mark Bowler. I am grateful for the knowledge, encouragement, and autonomy he provided me throughout this process. I would also like to thank the members of my committee, Dr. Shahnaz Aziz, Dr. Karl Wuensch, and Dr. Hamid Fonooni, for dedicating their time and expertise. Next, I must express immeasurable gratitude to my remarkable parents, Mike and Karen, my brother, Matt, and my sister, Andrea. I will be forever thankful for their unwavering love, support, and patience. I must also thank my amazing grandmother for her endless support of my ambitions. Finally, none of my achievements would be possible without the love of my life, Anna, who introduced me to work psychology, and motivates me to follow my dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Personnel Decision-Making

Personnel decision-making has gained mainstream popularity in the global business community and is a key management activity for any successful organization (Scullion, Collings, & Caligiuri, 2010). Furthermore, the process of making personnel decisions is no longer a function exclusive to Human Resources (HR), as Chief Executive Officers (CEOs) report spending a substantial amount of time on personnel activities (Economist Intelligence Unit, 2006). Personnel decision-making is part of an organization's efforts to strategically attract, select, develop, and retain high-performing employees (Farndale, Scullion, & Sparrow, 2010; Scullion et al., 2010). Moreover, employers utilize strategic personnel decision-making to remain competitive in the current global market, which considers highly-skilled workers to be one of the most important corporate resources available (Chambers, Foulon, Handfield-Jones, Hankin, & Michaels, 1998; Farndale et al., 2010; Vaiman, Scullion, & Collings, 2012).

Unfortunately, personnel decision-making is an area in which many HR professionals lack competence and expertise (Mellahi & Collings, 2010; Vaiman et al., 2012). Additionally, challenges, such as talent shortages, have continued to develop and place further strain on organizations seeking to remain competitive (Scullion et al., 2010). For example, in 2000, approximately 90 percent of organizational leaders reported experiencing difficulty attracting and retaining talented employees (Axelrod, Handfield-Jones, & Welsh, 2001) and as of 2008, talent was still listed as a top priority of organizations (Beechler & Woodward, 2009). New issues have also emerged including a demand for organizations to focus on employee-centered outcomes (e.g., well-being; Van De Voorde, Paauwe, & Veldhoven, 2012). Research on employee well-being suggests that such a focus may not only improve employee perceptions

(e.g., commitment, satisfaction; Nishii & Wright, 2007), but also enhance performance (Hoffmeister, Gibbons, Schwatka, & Rosecrance, 2015). As a result, organizations must now attend to a variety of targets and be capable of engaging in effective personnel decision-making in order to ensure future success.

Objectives of Personnel Decision-Making

Within an organization, personnel decision-making may address a variety of issues included in the process of attracting, selecting, developing, and retaining talented employees (Vaiman et al., 2012). For example, leaders within an organization may use data (e.g., an employee's tenure and previous performance) to ascertain if an employee should either be sent to training courses or is prepared for a shift into management. Although all personnel decisions can have a significant effect on both the employee and organization (Singh, Darwish, Costa, & Anderson, 2012), two specific objectives of personnel decision-making are particularly important to the workplace for their impact on organizational outcomes. Specifically, both employee performance (within the context of personnel selection) and employee health are essential to organizational success. The former objective is considered critical as it characterizes the degree to which an employee contributes to an organization and serves as the criterion of interest when selecting new employees. Subsequently, it has been the focus of organizational researchers for decades (Dunnette & Borman, 1979; Giffin, 1989; Morris, Daisley, Wheeler, & Boyer, 2015). In contrast, the latter is becoming increasingly more important to organizations as its influence on performance becomes better understood (Schwartz & Riedel, 2010). Although personnel selection boasts an extensive history of usefulness within organizations, employers have only recently begun to embrace the potentially equal importance of employee health (Schultz & Edington, 2007).

Personnel selection. Personnel selection has played a role in the lives of employees, organizations, and researchers for nearly 100 years. The overall goal of any personnel selection system is to produce and maintain high levels of organizational productivity by selecting individuals who will demonstrate high levels of job performance (Hunter, 1983). This task is typically accomplished through the application of psychological tests that systematically assess and predict an individual's potential future performance (Dunnette & Borman, 1979). Furthermore, practices such as job analysis and construct validation add rigor and accuracy to this endeavor (Freyd, 1923). Researchers predominately view personnel selection as a means for making accurate decisions regarding an employee's expected level of future performance. Present issues surrounding personnel selection include the use of alternative predictors of employee performance (e.g., personality, emotional intelligence, integrity) and novel methods of statistical analysis to address differential relationships with a criterion across employees (Sackett & Lievens, 2008).

Employee health. More recently, employers have begun focusing their attention on employee health as a mechanism for maintaining high levels of productivity and retaining top talent (Sears, Shi, Coberley, & Pope, 2013). Not only can physical and mental health affect an individual's status with regard to morbidity and health care costs (Harrison, Pope, Coberley, & Rula, 2012), it can also interfere with work-related factors, such as the ability to be present and perform at an acceptable level (Stewart, Ricci, Chee, & Morganstein, 2003). Although it was originally established as an important aspect of personnel management via its link with organizational health care costs (Vickery, Golaszewski, Wright, & McPhee, 1986; Yen, Edington, & Witting, 1991), the focus on employee health has continued to grow. For example, employers now address issues such as forecasting when employees are a health-related risk to the

organization and are in need of workplace intervention (Schultz & Edington, 2007). Furthermore, modern interventions, such as changing an organization's ergonomics climate via job redesign and elimination of perceived hazards, are being incorporated into the workplace in an attempt to reduce physical and mental strain and increase employee productivity (Hoffmeister et al., 2015). The impact of well-being on organizational outcomes and the efficacy of workplace interventions are also being studied longitudinally in order to better estimate health costs and inform future interventions (Sears et al., 2013). As with personal selection, issues regarding the application of novel statistical analyses to issues of employee health must also be examined. Although the benefits of such analyses may be best demonstrated first within the context of personnel selection, which serves as the entry point into an organization and represents a disproportionate amount of empirical research, additional examination within the context of employee health is critical when identifying benefits for fields of study such as Occupational Health Psychology (OHP).

Statistical Techniques for Decision-Making

In order to obtain and retain top talent, numerous personnel selection measures have been developed to help organizations forecast a candidate's potential job performance (Schmidt & Hunter, 1981). Such measures include work samples, cognitive ability tests, integrity assessments, and personality inventories (Sackett & Lievens, 2008; Thornton & Kedharnath, 2013; Van Iddekinge, Roth, Raymark, & Odle-Dusseau, 2012). Following the collection of candidate data via some or all of these measures, statistical techniques are utilized to develop predictive models that determine the probability that a given job candidate will demonstrate an acceptable level of performance. Traditionally, organizational researchers have relied on techniques such as regression to assist them in developing these types of models (Raju,

Steinhaus, Edwards, & DeLessio, 1991; Stauffer & Ree, 1996). However, research suggesting that potential improvements can be made with regard to the validity and interpretability of derived solutions has promoted the use of novel statistical techniques, such as decision tree analysis (Lemon, Roy, Clark, Friedmann, & Rakowski, 2003; Lewis, 2000; Loh, 2009).

Multiple linear regression is one of the most frequently employed statistical techniques in the social sciences (Sanders & Brynin, 1998). In general, regression models use information about a given set of predictor variables to construct a simple linear equation that predicts the value of a given criterion variable. In order to evaluate the derived solution, researchers use an R^2 statistic that indicates the proportion of variance in the criterion explained by the model (Sweet & Grace-Martin, 2012). Multiple linear regression is best used when the outcome of interest demonstrates the characteristics of continuous, interval-style data. However, if the criterion is dichotomous, as may often be the case in personnel decision-making, binary logistic regression is a more appropriate analysis. Although both of these techniques are widely used within the field of personnel decision-making, each method is characterized by a variety of important assumptions and potential weaknesses.

One statistical method that demonstrates particular promise for personnel decision-making is data mining. Data mining procedures, such as decision tree analysis, are currently utilized within the fields of statistics, medical diagnosis, and fraud detection (Lavanya & Rani, 2011). For example, medical researchers utilize various decision tree analysis techniques, such as Iterative Dichotomiser 3 (Quinlan, 1986), C4.5 (Quinlan, 1993), and Classification and Regression Trees (Breiman, Friedman, Olshen, & Stone, 1984), to make important clinical decisions regarding diagnosis and treatment (e.g., classifying patient risk). Similarly, financial institutions also rely on decision tree analysis and other data mining procedures, such as neural

and Bayesian networks, to detect fraudulent financial transactions and behaviors (e.g., manipulation of financial records; Zhou & Kapoor, 2011).

Of particular relevance to making personnel decisions is decision tree analysis. Decision tree analysis utilizes binary recursive partitioning to construct multi-pathway predictive models of categorical outcomes. This method is preferred by many researchers for its ability to provide accurate and interpretable classification solutions in the midst of data characteristics that typically violate the statistical assumptions of traditional techniques (Karels, Bryant, & Hik, 2004; Lewis, 2000). As with medical data, personnel data can often be non-normal, demonstrate non-linear relationships, and have a substantial amount of multicollinearity (Schmidt, Ones, & Hunter, 1992). As a result, traditional modeling techniques may provide predictive solutions with decreased levels of validity (Lumley, Diehr, Emerson, & Chen, 2002). Furthermore, traditional techniques may also limit the ability for certain candidates to demonstrate compensatory abilities, resulting in solutions with increased levels of adverse impact. In contrast, decision tree analysis is both non-parametric and compensatory (Karels et al., 2004).

Despite such benefits, the use of decision tree analysis in the practice of personnel decision-making is largely non-existent. Potential reasons for the underutilization of decision tree analysis include a general lack of understanding and experience with the technique, as well as the technique's origin within the field of machine learning, which may be unfamiliar to individuals who make personnel decisions (Lewis, 2000; Loh, 2011). Furthermore, the statistical software necessary to conduct decision tree analysis was, until recently, more difficult for practitioners to obtain (Lewis 2000). However, as there have been substantial gains in both the understanding of the technique and the accessibility of necessary software in recent years, decision tree analysis holds clear promise for personnel decision-making in organizations.

Current Study

In the current study, the relative advantages and disadvantages of applying decision tree analysis to predict employee performance and health were investigated. Specifically, decision tree analysis was evaluated in comparison to binary logistic regression regarding the ability to provide organizations with optimal decision-making solutions. By incorporating employee performance and health, the current study consisted of both a historically significant component of personnel decision-making, as well as an emerging aspect that is expected to play a growing role in the success of organizations. Furthermore, the current study included key outcomes of interest within the field of OHP. In order to compare the efficacy of each technique, a variety of comparison criteria were utilized (i.e., validity, adverse impact, interpretability). In particular, decision tree analysis was investigated for its potential to model multiple pathways to organizational outcomes and provide fair and accurate statistical solutions.

Research Questions

Given the many benefits provided by the use of decision tree analysis in other areas of study (e.g., statistics, medicine), it is imperative that the technique be investigated for potential advantages within the context of personnel decision-making. In conducting such an investigation, the following research questions (RQ) were proposed:

RQ_1: In comparison to traditional methods (i.e., binary logistic regression), can the use of decision tree analysis provide benefits to the practice of personnel decision-making?

RQ_2: When modeling key organizational outcomes (i.e., employee performance and health), what potential benefits to classification accuracy are provided by decision tree analysis in comparison to binary logistic regression?

RQ_3: When modeling key organizational outcomes (i.e., employee performance and health), what potential benefits to non-discriminatory hiring (i.e., adverse impact) are provided by decision tree analysis in comparison to binary logistic regression?

RQ_4: When modeling key organizational outcomes (i.e., employee performance and health), what potential benefits to interpretability are provided by decision tree analysis in comparison to binary logistic regression?

CHAPTER II: LITERATURE REVIEW

Personnel Selection

Faced with today's competitive global environment, organizations are increasingly aware of the impact that personnel decision-making can have on organizational success (Carless, 2009; Vaiman et al., 2012). Although a variety of decisions and strategies may be used by an organization to increase the performance of its employees, none may be more important than efficient and effective personnel selection (Thomas & Scroggins, 2006). Personnel selection is the process of collecting and utilizing data about a job candidate in order to make important employment decisions, such as whether or not to extend an initial offer of employment or whether or not a particular individual should be promoted (Guion & Gibson, 1988). This process is of critical importance to the success of an organization (Chien & Chen, 2008) as personnel selection is the primary mechanism for evaluating an individual's ability to perform a particular job and placing him or her within the organization (Zhang & Liu, 2011). In comparison, personnel-related decisions made after selection are largely aimed at improving upon an employee's skills, abilities, and performance (Gorman, Thibodeaux, Eisinger, & Overstreet, 2012). However, if an employee does not possess the initial traits, skills, and abilities necessary for successful performance, such strategies may be misguided and result in failure and wasted resources (Gowan & Gatewood, 2013). Consequently, it is essential that organizations maximize personnel decisions prior to the onset of employment.

Fortunately, decades of research on personnel selection have provided organizations with a wealth of knowledge and guidance (Schmidt & Hunter, 1998). For example, research has highlighted the importance of first conducting a robust evaluation of the job of interest (i.e., job analysis) in order to accurately understand the nature of the job and identify the knowledge,

skills, and abilities (KSAs) that are necessary for success (Borman, Hanson, & Hedge, 1997). Following the collection of such information, organizations must identify methods for determining if a job candidate has the appropriate KSAs required to perform the job effectively (Robertson & Smith, 2001). Personnel selection research has promoted the use of psychological assessment devices to assist organizations in accomplishing this task (e.g., Carless, 2009; Terpstra & Rozell, 1993; Thomas & Scroggins, 2006). The use of psychological assessments has rapidly increased over time (Anderson, 2005). For example, as of 2004, nearly half of all Fortune 100 companies reported using some form of personality assessment when making employment decisions (Erickson, 2004). Moreover, integrity tests, a particular type of personality assessment aimed at measuring an employee's predisposition towards dishonesty and counterproductive work behavior (Hogan & Brinkmeyer, 1997), are reportedly given to millions of job applicants per year in an attempt to ascertain their suitability for employment (Heller, 2005). Additional methods recommended to organizations when making personnel decisions include structured interviews, work samples, and assessment centers (Borman et al., 1997; Guion & Gibson, 1988; Robertson & Smith, 2001). Overall, the use of these methods seeks to facilitate accurate prediction of job performance, improve employee-organization fit, and reduce turnover (Rothstein & Goffin, 2006).

Despite the rapid growth and success of selection procedures, improvements are still needed (Sackett & Lievens, 2008). One area in which innovations can benefit the field is in regard to how personnel decisions are ultimately made. Specifically, relatively little emphasis is placed on the statistical analysis process of personnel decision-making (Hunsley & Meyer, 2003). Due to this lack of attention, critical review and innovation is required.

Traditional Decision-Making Methods

Statistical Methods

Multiple linear regression. Multiple linear regression is one of the most commonly used statistical methods throughout the social sciences and within personnel selection (Halinski & Feldt, 1970; Kriska & Milligan, 1982). In general, linear regression uses information about a given set of independent or "predictor" variables (i.e., $X_1, X_2, X_{3, \ldots}$) to predict the value of a given dependent or "criterion" variable (i.e., Y; Flynn & Peterson, 1972). Although the primary purpose of regression is to predict future values of a criterion variable using only information about a set of independent variables, in the context of personnel selection, regression models are commonly first constructed on a concurrent sample that provides information on both predictor and criterion variables (Peterson & Wing, 2001). In either case, multiple regression models can be expressed as a simple linear equation:

$$Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n + e \qquad (1)$$

Within the linear equation, the term, a, represents the intercept, b represents the slope or change in the criterion variable given a one unit change in a predictor variable, and as previously mentioned, Y and X represent the criterion and predictor variables, respectively (Sanders & Brynin, 1998). Values of the intercept and slope are defined in order to minimize the sum of squared residuals and form the best-fitting line to the sampled data (McCornack, 1970). Subsequently, values of the criterion variable can be predicted using only the values of a select number of predictor variables and their estimated intercept and slope coefficients (Halinski & Feldt, 1970; Sanders & Brynin, 1998).

Unfortunately, predictions based on regression equations are rarely ever void of error. In turn, an error term, e, is also included in the linear equation and represents the variance in Y that

is not explained by the linear combination of the included predictor variables (Rossi, 2010). However, as the strength of the relationships between the predictor variables and *Y* increases, the magnitude of error in the model decreases. Moreover, as error, or residuals for each case in the model, becomes smaller, the regression model will be characterized by a tight clustering of cases around the estimate line, as shown in Figure 1 (Sanders & Brynin, 1998). Lastly, as a mechanism for evaluating linear regression models, researchers typically utilize the R^2 statistic, which ranges from 0 to 1 and indicates the strength of the relationship or proportion of variance in the criterion variable that is explained by the model (Kriska & Milligan, 1982; Sweet & Grace-Martin, 2012).

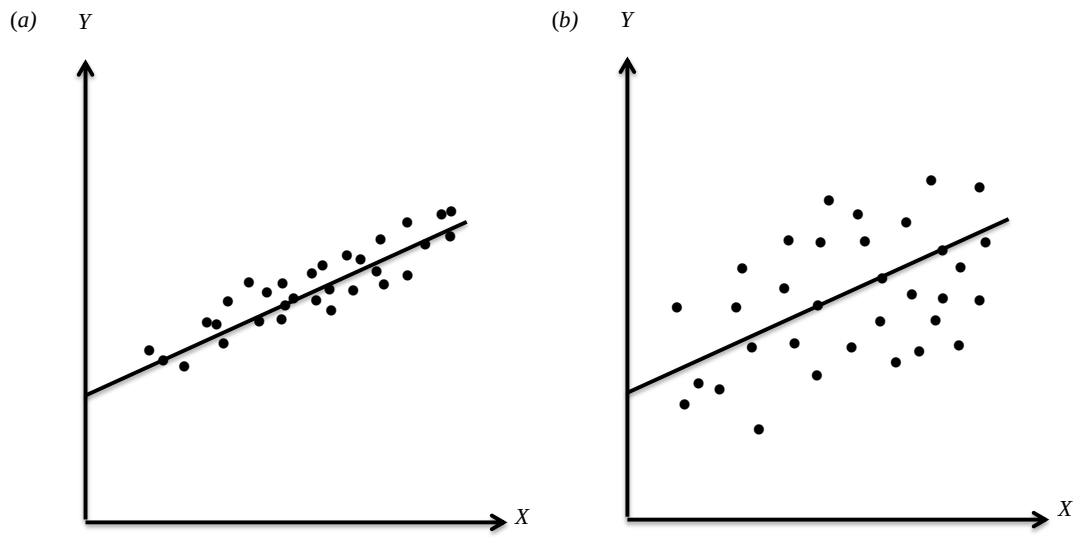


Figure 1. Strong relationship with tight clustering of cases around estimate line (a) and weak relationship with loose clustering of cases around estimate line (b).

When developing a personnel selection system, multiple regression is typically included in the final stage of data analysis (Sweet & Grace-Martin, 2012). First, job incumbents are assessed using a battery of selection instruments (i.e., predictor variables) from which predictor scores are derived (Weissmuller & Damos, 2014). Criterion scores are also collected for each incumbent during this stage. Next, organizations conduct univariate and bivariate analyses in

order to assess the assumptions of the linear model and better understand the structure and distribution of the predictor variables (Rossi, 2010). Following this stage, a multiple regression model is fit using the least squares method—determining the equation of the best fitting line that minimizes the sum of squared deviations (McCornack, 1970). If the derived model is significant, it is then utilized as a predictive model for job candidates (Rossi, 2010). This process is completed by first assessing job candidates via a battery of selection instruments, identical to those used to develop the predictive model (Peterson & Wing, 2001). Organizations then utilize both the predictor variable scores and the linear model's regression equation to generate a single predicted criterion score. Finally, selection decisions are made based on each candidate's predicted level of job performance (Weissmuller & Damos, 2014).

Strengths and weaknesses. Linear regression is one of the most well-known and well-received statistical techniques in organizational psychology (Raju et al., 1991; Stauffer & Ree, 1996). This may be due to the fact that regression allows individuals to statistically support theoretical explanations for social phenomena (Sanders & Brynin, 1998). Beyond this aspect, regression boasts a variety of other strengths that help to explain its popularity. For example, it is characterized by its ability to demonstrate the collective effect of multiple predictors (Sweet & Grace-Martin, 2012). Specifically, when multiple variables are included in a model, regression identifies the portion of variance collectively explained by the variables, each variable's specific effect on the outcome variable (i.e., semi-partial correlation), and the portion of variance that remains unexplained (Abdi, 2007). Such information is critical to researchers who must measure the amount of explained variance, in order to develop the most parsimonious model possible and accurately predict future outcomes (Halinski & Feldt, 1970; McCornack, 1970).

Furthermore, regression is powerful in its ability to account for the effect of spurious relationships through the use of control variables (Sweet & Grace-Martin, 2012). By including control variables, regression is able to determine which predictor variables demonstrate a unique effect on the outcome and which predictor variables only demonstrate an effect on the outcome due to their relationship with other variables (Jaccard & Turrisi, 2003). For example, researchers may control for demographic characteristics, such as gender and age, to ensure that predictor variables contribute unique amounts of variance and are useful for inclusion in selection systems. Similarly, researchers may find that although a variable appears to have a particular relationship with a criterion, in the presence of another variable, the original relationship disappears (Cohen, Cohen, West, & Aiken, 2003). This facet of regression allows researchers to identify the most valid and parsimonious model (McCornack, 1970). Overall, regression is a useful tool for constructing predictive equations that form the best fitting line for, theoretically, all individuals in a sample. Nevertheless, in addition to regression's many strengths, there are also several weaknesses that limit its use.

Although regression performs relatively well when modeling linear relationships, when data are non-linear or "curvilinear," regression may fail to identify the most accurate solution (Sharma, 2005). Specifically, when variables have a linear relationship with a criterion, scores on the predictor variable will have a consistent relationship with scores on the criterion variable (Cohen et al., 2003). For example, if positively correlated, low scores on the predictor variable are associated with low scores on the criterion. Similarly, if negatively correlated, low scores on the predictor variable are associated with high scores on the criterion. However, not all variables demonstrate linear relationships. One example of a curvilinear relationship is found between employee turnover and variables such as intelligence and job performance (see Figure 2; Hurley

& Estelami, 2007; Stone & Kendall, 1956; Trevor, Gerhart, & Boudreau, 1997). Specifically, research demonstrates that very high or very low levels of intelligence and job performance are associated with high risk of turnover for jobs that require average levels of intelligence or present limited opportunities to above-average employees (Williams & Livingstone, 1994). As a result, regression may fail to identify the curvilinear relationship and produce an inadequate linear model. Moreover, advanced regression techniques that can be used to model curvilinear data, such as polynomial regression, may surpass the capabilities of an organization whose primary function is not related to statistical analysis.

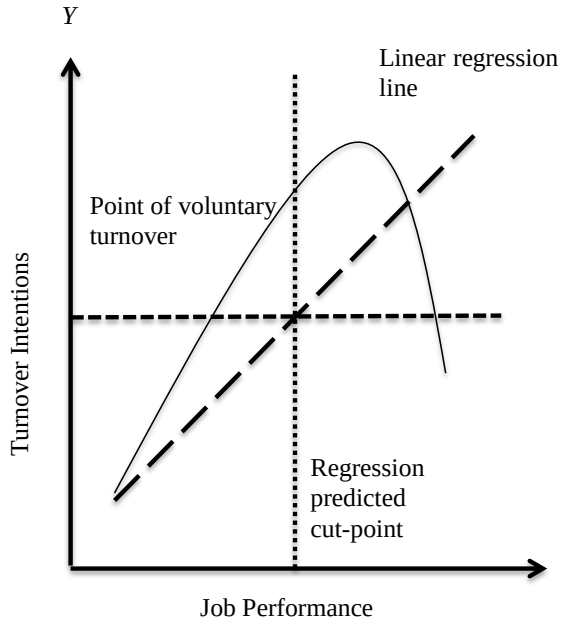


Figure 2. Sample curvilinear relationship between turnover and performance.

Additionally, regression is limited in its inability to deal with non-normal data, as well as data that consist of characteristics that violate statistical assumptions made by regression, such as the assumption of homoscedasticity (Long & Ervin, 2000). Specifically, linear regression models produce residuals, or deviations between the actual and predicted values of a criterion for each case, that are non-random. The residuals produced by linear regression assume homoscedasticity, or a stable and systematic distribution of residuals at all levels of the linear

combination (Leslie, Kohn, & Nott, 2007). However, some data do not demonstrate such consistency. Rather, for data that demonstrate heteroscedasticity, regression is much better at predicting values at a certain point in the distribution than it is at predicting values at other points in the distribution (Sweet & Grace-Martin, 2012). For example, the data may produce a "fan effect" that is characterized by cases that are increasingly less clustered around the regression line at higher values of the criterion than at lower values (see Figure 3). Due to such characteristics, regression fails to produce the best model of the data and is limited in its ability to accurately describe and predict the criterion (Long & Ervin, 2000). Nonetheless, linear regression stands as one of the most common statistical methods for predicting important organizational outcomes.

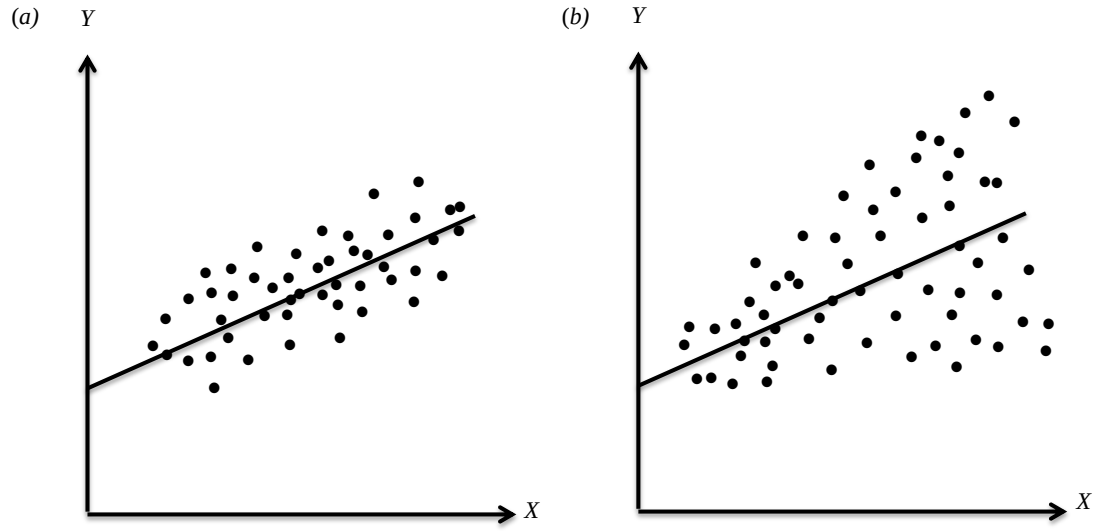


Figure 3. Evenly distributed clustering of cases around estimate line, "homoscedastic" (a) and fan-like clustering of cases around estimate line, "heteroscedastic" (b).

Binary logistic regression. In addition to multiple linear regression, researchers may also utilize binary logistic regression when making important personnel decisions. Logistic regression has become increasingly popular within the social sciences and is the preferred form of regression when modeling dichotomous criteria (Aldrich & Nelson, 1984; Pohlmann & Leitner, 2003). Compared to logistic regression, other forms of regression, such as multiple

linear, are limited when modeling data that consists of heteroscedasticity and nonlinearity. Such issues can be common when using dichotomous criteria and become even more pronounced when participants are not evenly split on a criterion (Stauffer & Ree, 1996). Furthermore, when attempting to clearly classify employee performance as either acceptable or not acceptable, linear regression is criticized for providing organizations with predicted values that are difficult to interpret (Raju et al., 1991). In contrast, logistic regression yields predicted values that can be interpreted as probabilities of success within a fixed 0 to 1 range. Logistic regression is also capable of being evaluated using easy to understand classification criteria such as accuracy, specificity, and sensitivity (Stauffer & Ree, 1996).

In general, logistic regression uses information about a given set of predictor variables to estimate the probability of a binary outcome occurring (Ngo, Govindu, & Agarwal, 2015). Similar to linear regression, logistic regression can also be expressed via the following equation:

$$\text{Log}\,[p / (1-p)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + e \qquad (2)$$

Within the equation, the term p represents the probability of the outcome, β_0 represents the intercept, and β_i represents the logistic regression coefficient for each predictor variable (Ngo et al., 2015). In comparison to linear regression, which uses a least squares method to determine the equation of the best fitting line that minimizes the sum of squared deviations (McCornack, 1970), logistic regression uses maximum likelihood estimation to find the set of coefficients for which the probability of the observed data is greatest (Czepiel, 2002). Furthermore, by utilizing a logit transformation, logistic regression is capable of performing such estimation despite the presence of non-normal data characteristics, such as non-linearity and heterogenous error variances (Peng, Lee, & Ingersoll, 2002; Pohlmann & Leitner, 2003).

***Strengths and weaknesses*.** Binary logistic regression is characterized by a variety of strengths associated with its relative lack of strict statistical assumptions that limit other forms of regression (Peng et al., 2002). This leniency makes logistic regression the current technique of choice when modeling dichotomous outcomes (Ngo et al., 2015). Nevertheless, logistic regression is still characterized by weaknesses that can limit its overall usefulness to organizations. For example, as with linear regression, logistic regression is limited when modeling data that demonstrate multicollinearity (i.e., when two or more predictor variables included in a model are highly correlated with one another; DeMaris, 2013; Vaughan & Berry, 2005). Unfortunately, personnel data are often comprised of predictors that consist of a substantial amount of multicollinearity (e.g., the five factor model; Schmidt et al., 1992). As a result, logistic regression can produce a model that possesses inflated estimates and standard errors, as well as counterintuitive signs of coefficients (Schaefer, 1986; DeMaris, 2013). Although the weaknesses of logistic regression are comparatively fewer than linear regression, areas for improvement still exist in regard to accurate classification results.

Selection Techniques

Cutoff scores. Beyond regression, researchers in personnel selection may employ a variety of other techniques when developing selection systems. One such technique is the setting of cutoff scores (Truxillo, Donahue, & Sulzer, 1996). Cutoff scores refer to a specified point on a distribution of scores, derived from either a single measure or battery of measures, above which candidates' scores are considered acceptable and below which they are considered unacceptable (Cascio, Outtz, Zedeck, & Goldstein, 1995). Such cutoff scores have traditionally played an important role in personnel decision-making and have assisted practitioners in classifying candidates for processes such as selection, promotion, and training (Cascio,

Alexander, & Barrett, 1988). Although cutoff scores may be assigned somewhat arbitrarily, legal and academic trends have placed pressure on organizations to use standardized practices for setting scores (Maurer, Alexander, Callahan, Bailey, & Dambrot, 1991). Consequently, it is important to review the importance of cutoff scores, their legal implications, methods for selecting quality cutoff scores, and their overall strengths and weaknesses.

Although a select number of organizations are required to set cutoff scores by their governing bodies (Truxillo et al., 1996), most are not. Furthermore, many organizational researchers (e.g., Schmidt, Mack, & Hunter, 1984) recommend using top-down selection procedures (i.e., measuring performance as a continuous variable and selecting candidates in order of their scores). Unfortunately, top-down selection may be complicated by conflicting organizational goals and responsibilities (e.g., validity, diversity; Cascio, Outtz, Zedeck, & Goldstein, 1991). As a result of such issues, some researchers recommend the use of statistical bands within which all candidates can be considered equal and selected upon an additional unbiased criterion (Cascio et al., 1995). However, banding has also received significant criticism on both statistical and logical grounds (Schmidt, 1991; Schmidt & Hunter, 1995). Therefore, the setting of cutoff scores is an important technique for organizations to consider when balancing valid selection decisions and workforce diversity.

As most organizations are faced with limiting factors, such as a specific number of vacancies, it is important for researchers to be able to distinguish between candidates who justify further consideration and those who do not. In order to accomplish this task, organizations may select cutoff scores that separate desirable candidates from those who are undesirable (Corte, 1995). Furthermore, although many employee outcomes are frequently considered continuous, most organizations will eventually dichotomize such outcomes in practice as either meeting or

failing to meet a pre-specified standard (Farrington & Loeber, 2000). Under such circumstances, there are a variety of ways in which cutoff scores can be used (Cascio et al., 1988). For instance, when faced with few vacancies or a low need for new incumbents, organizations can establish a relatively high cutoff score and only select individuals who exceed the high standard. Next, when faced with many vacancies and a high need for new incumbents, organizations can establish a relatively low cutoff score and select all individuals who demonstrate basic proficiency. In this second instance, the defined cutoff score is of less importance, as job openings are likely to be filled before reaching the low cutoff score (Cascio et al., 1988). Finally, organizations may strategically incorporate cutoff scores into a multiple hurdle selection system that only permits a select group of candidates to proceed to more costly selection procedures, such as assessment centers and on-site interviews (Weissmuller & Damos, 2014). For all of these circumstances, cutoff scores are currently the preferred method for determining which candidates have an adequate level of ability to move forward in the selection process and which ones do not (Truxillo et al., 1996).

Legally, cutoff scores have received significant attention. Overall, court decisions have been largely supportive of cutoff scores and place few restrictions on their use. For example, in *Board of Regents of the University of the State of New York v. Tomanio* (1980), Justice Stevens noted that testing "is a permissible method of determining qualifications, and lines must be drawn somewhere" (p. 446). Furthermore, when questioned on the grounds of constitutional rights and discrimination, courts have consistently upheld the method and support the use of cutoff scores as long as they are reasonable and are selected with some degree of rationale (Cascio et al., 1988). For example, classic adverse impact cases, such as *Griggs v. Duke Power Co.* (1971), *Albemarle Paper Co. v. Moody* (1975), and *Rogers v. International Paper Co.*

(1975), all call into question the selection of cutoff scores that were perceived as exceedingly stringent or discriminatory. Nevertheless, all organizations involved in such cases were able to justify the use of the particular cutoff score and the scores were permitted for use, with one portion of the *Griggs v. Duke Power Co.* (1971) ruling stating, "an employer may set his qualifications as high as he likes, he may test to determine which applicants have these qualifications, he may hire, assign, and promote on the basis of test performance" (p. 401). Despite seemingly vague standards and restrictions surrounding cutoff scores, researchers must be familiar with techniques for selecting appropriate cutoff scores, as court cases regarding discrimination and constitutionality have historically had a significant influence on the field of personnel selection and can be costly to organizations.

Formal recommendations for setting cutoff scores typically lack specificity regarding the type of procedure or information an organization should consider when making such decisions. For example, the *Uniform Guidelines on Employee Selection Procedures* (1978) merely state that a cutoff score should be both reasonable and consistent with the expectations for acceptable workforce proficiency (EEOC, 1978). Nevertheless, several general recommendations exist to help guide organizations in establishing appropriate cutoff scores. First, court cases, such as *Vulcan Pioneers v. New Jersey Department of Civil Service* (1985), have established the notion that cutoff scores should (1) be consistent with information derived from a job analysis, (2) permit the selection of qualified candidates, and (3) allow an organization to achieve its workforce diversity goals (Cascio et al., 1988). Next, the creation and implementation of cutoff scores should always be completed with the assistance of subject matter experts (SMEs; Cascio et al., 1988). Finally, it is suggested that, for organizations to best utilize SMEs and produce

cutoff scores that are effective and defensible, standardized methods should be used for setting the appropriate cutoff score. One such technique is the Angoff method (Angoff, 1971).

The Angoff method is considered to be the best-suited technique for producing cutoff scores that have a high level of content-related validity, technical adequacy, and practicality (Maurer et al., 1991). Furthermore, researchers (Berk, 1986; Cascio et al., 1988; Shepard, 1980) have established the Angoff method as the preferred technique for setting cutoff scores, due to its high degree of reliability and ease of use. Specifically, the Angoff method requires multiple raters, who are typically SMEs, to individually rate the probability that a minimally competent candidate would answer a test item correctly (Maurer et al., 1991). Test items may come from a variety of different job-related measures used for a selection system. After all raters have estimated the appropriate probability for each item, a cutoff score is determined by calculating the sum of the item proportions (i.e., the average of the item proportions multiplied by the number of test items; Maurer et al., 1991). Lastly, if an organization's selection system is not complementary to the Angoff method (e.g., uses a limited number of selection measures featuring test items), it is recommended that other standardized procedures be used, such as having SMEs determine an appropriate cutoff score based on a distribution of applicant performance (e.g., at the mean, one standard deviation above the mean, one standard deviation below the mean; Cascio et al., 1988).

***Strengths and weaknesses*.** Cutoff scores boast a variety of strengths and are considered an important tool for personnel decision-making. In general, they are easy to implement, legally defensible, and useful in categorizing candidates by their level of competence. Furthermore, cutoff scores help to facilitate the use of multiple hurdle selection systems by eliminating candidates who clearly do not have the minimum level of qualifications required for future

consideration in the selection process. In turn, cutoff scores can save organizations financial and time-based resources. Finally, due to the relatively vague standards surrounding the use of cutoff scores, organizations have the flexibility necessary to modify and select cutoff scores that best meet their changing needs (Cascio et al., 1988).

Organizations should, however, be aware of the limitations regarding the use of multiple cutoff scores. Although every attempt should be made to produce valid and reliable cutoff scores via standardized methods and quality SMEs, cutoff scores can be relatively subjective. As a result, organizations may find that personnel decisions made using a particular cutoff score are not well correlated with subsequent job performance. Despite legal guidelines that do not require mathematical precision or a high degree of accuracy, such issues should still remain important to personnel decision-makers (Cascio et al., 1988). Lastly, given the recommended procedure for establishing appropriate cutoff scores (e.g., Angoff method), it can be assumed that the quality of a cutoff score is a direct reflection of the SMEs who establish the score. Consequently, organizations should carefully consider each SMEs level of competence and qualifications for making assessments within the area of interest (Maurer et al., 1991).

Multiple hurdles. Organizations rarely use only one predictor variable when making important personnel decisions. In turn, they are faced with the task of determining the best technique for combining candidates' predictor scores in order to make personnel decisions (Salgado, Viswesvaran, & Ones, 2001). One of the most frequently used techniques for grouping candidates' predictor scores and forming selection decisions is the multiple hurdle system.

Within a multiple hurdle system, job candidates are assessed using a sequential series of job predictors that are nested into distinct groups or "hurdles" (Salgado et al., 2001). Hurdles

may consist of either a single job predictor (e.g., work sample) or a group of job predictors (e.g., cognitive ability and personality measures). Regardless of the number of predictors contained within each hurdle, when assessed by a multiple hurdle system, job candidates must ultimately meet the requirements of one hurdle before proceeding to the next (Salgado et al., 2001). Similar to a multiple cutoff method, multiple hurdle systems assume that a minimum level of ability is required to perform any particular job (Finch, Edwards, & Wallace, 2009). Consequently, when designing a multiple hurdle system, organizations must decide both the number and type of predictors to be nested within each hurdle, as well as the minimum level of acceptable performance on each hurdle (Finch et al., 2009). Overall, multiple hurdle systems are a useful and popular tool for incorporating the techniques of multiple regression and multiple cutoffs into a robust selection system.

***Strengths and weaknesses*.** One strength offered by a multiple hurdle system is the financial savings it provides to organizations. As most selection assessments vary in cost, organizations may be hesitant to use certain techniques (e.g., assessment centers) or administer all selection assessments to an entire candidate pool, which is often very large in comparison to the number of available positions (Salgado et al., 2001). Fortunately, multiple hurdle systems allow organizations to sequentially administer selection assessments to candidates and only utilize expensive techniques on quality applicants who have progressed through the hurdle system (Mendoza, Bard, Mumford, & Ang, 2004). Additionally, multiple hurdle systems may be preferred when selecting for positions that demand an applicant to have a particular skill or ability. For example, when selecting for airline pilots, organizations must ensure that a candidate has the visual abilities to pilot an aircraft (Temme, Still, & Fatcheric, 1995). If a candidate does not have such abilities, then there is no reason for the organization to administer additional

selection measures to the individual (Finch et al., 2009). However, it can be assumed that currently, such positions are the exception, rather than the rule, and that most modern day positions require a candidate to display a variety of skills and abilities. It is for such positions that multiple hurdle systems, as they are presently employed, demonstrate a critical weakness.

Specifically, multiple hurdle systems, as well as multiple cutoff systems, are non-compensatory. Non-compensatory selection systems assume that all predictor variables are equally important and thus, a job candidate cannot compensate for a deficit in one trait with strengths in another (Salgado et al., 2001). Selection techniques such as multiple hurdle systems are non-compensatory in that they eliminate job candidates for their lack of performance on specific variables prior to assessing them on additional variables that may allow the individual to still demonstrate acceptable levels of job performance. Although multiple hurdle systems can conserve financial resources for an organization if the job is truly non-compensatory, in the likely case that job performance can be achieved through compensatory means, then multiple hurdle systems will result in error. Specifically, error occurs in non-compensatory selection systems when candidates with high scores on initial screening predictors but unexceptional scores on later hurdles and measures of job performance are selected, whereas candidates with low scores on initial screening predictors but high scores on later hurdles and measures of job performance are rejected (Sackett & Roth, 1996).

Given that the majority of modern jobs would in some way facilitate compensatory routes to job performance, multiple hurdle systems and related non-compensatory techniques (e.g., multiple cutoff systems) are potentially inappropriate for continued use. The use of compensatory systems that combine aptitude-based measures with various other predictors such as work samples, structured interviews, and personality assessments as a means to increase

validity, decrease adverse impact, and increase utility has been explored in the past (Hoffman & Thornton, 1997; Sackett & Roth, 1996); however, the use of data analysis techniques that facilitate compensatory modeling has yet to be adequately investigated. Overall, traditional selection systems are associated with a variety of strengths, as well as very important weaknesses. Consequently, it is vital that novel decision-making methods be investigated for their ability to better model compensatory variables and to provide organizations with valid and fair selection systems.

Novel Decision-Making Methods

A novel decision-making method that demonstrates particular promise for personnel decision-making originates from the field of data mining. Data mining techniques have become increasingly popular in recent years and are currently utilized within the fields of statistics, medicine, and fraud detection (Lavanya & Rani, 2011). Specifically, decision tree analysis is an important statistical data mining technique that may prove extremely useful for personnel selection and the modeling of organizational outcomes. Decision tree analysis is suggested to be useful within the context of personnel decision-making due to its ability to construct compensatory, multi-pathway models that overcome the many weaknesses of traditional techniques. The construction of such compensatory models has the potential to assist organizations in reaching both their accuracy and workforce diversity goals. Furthermore, decision tree analysis has become a preferred method of analysis for many clinical researchers who value the technique's ability to provide accurate and interpretable solutions, despite the presence of data characteristics that violate common statistical assumptions (Lewis, 2000).

A defining feature of decision tree analysis is its use of categorical outcomes. Although organizational outcomes are often conceptualized as continuous, many outcomes are eventually

dichotomized as either achieving or failing to achieve a predefined standard. Job performance is an example of such an outcome. Specifically, job performance is often considered a continuous variable and may be measured via objective performance data (e.g., number of sales) or subjective supervisor ratings that use multiple scales of performance (Behrman & Perreault, 1982). However, in practice, organizations likely dichotomize job performance. For example, organizations may classify candidate performance as either "acceptable" or "unacceptable," incumbents as either "high performers" or "low performers," and employees as either displaying a certain behavior (e.g., counter-productivity) or not (Farrington & Loeber, 2000; Herrero & Villar, 2013; Stone & Kendall, 1956; Trevor et al., 1997). Furthermore, organizations may use such classifications to make decisions and take action, such as deciding whether to terminate an employee who consistently demonstrates poor performance or disruptive behavior (Viswesvaran, Ones, & Schmidt, 1996).

Consequently, although operationalized as being continuous variables, many organizational criterion measures may in fact be treated as dichotomous and then not truly have the continuous, interval-style characteristics required by multiple linear regression. However, as previously noted, binary logistic regression is specifically designed to model dichotomous outcomes (Sanders & Brynin, 1998). As a result, binary logistic regression is the most appropriate statistical method with which to compare to data mining techniques, such as decision tree analysis.

Decision Tree Analysis

Four primary components are necessary to address any given classification problem. These include: (1) A categorical outcome variable (e.g., job performance), (2) a set of predictor variables (e.g., cognitive ability, conscientiousness), (3) a "training" data set (i.e., a concurrent

sample), and (4) a future data set to which the derived model will be applied (Lewis, 2000). As with traditional regression, decision tree analysis uses the aforementioned components to construct a model for predicting values of the outcome variables from specifically defined values of the predictor variables (Loh, 2011). However, in contrast to the single linear solution provided by traditional methods, decision tree analysis identifies multiple pathways to a particular outcome (Lewis, 2000). This solution is not achievable by traditional regression techniques, which provide single linear combination solutions that assume a homogenous relationship between predictor and outcome variables across all participants (Breiman, Friedman, Olshen, & Stone, 1993).

Although this may seem counterintuitive, the identification of multiple pathways conforms to the general notion that predictors of job performance are compensatory (Rothstein, Paunonen, Rush, & King, 1994). For example, it is widely accepted in organizational research that performance is a function of the product of ability and effort (Chan, Schmitt, DeShon, Clause, & Delbridge, 1997). Thus, individuals with high levels of ability are typically successful as are individuals who put forth high levels of effort. However, single linear models that find ability to be most predictive of success may erroneously categorize individuals who have lower levels of ability as unsuccessful. In contrast, decision tree analyses theoretically have the capacity to simultaneously identify successful individuals with high levels of ability and individuals with lower levels of ability who achieve success through high levels of effort or other compensatory variables (e.g., personality).

Decision tree analysis grows decision trees via binary recursive partitioning, which recursively splits a data set one predictor variable at a time (Loh, 2011). This process begins with a root “node” that includes all cases within the training data set. Following the introduction

of a predictor variable, a split is made and two binary "child" nodes are created. The recursive process continues, subsequently converting child nodes into "parent" nodes as additional splits are made in the data. Recursive partitioning of data is often demonstrated geometrically by splitting values within a data set into rectangles, with each split resulting in more homogenous groups of data (see Figure 4; Breiman et al., 1993).

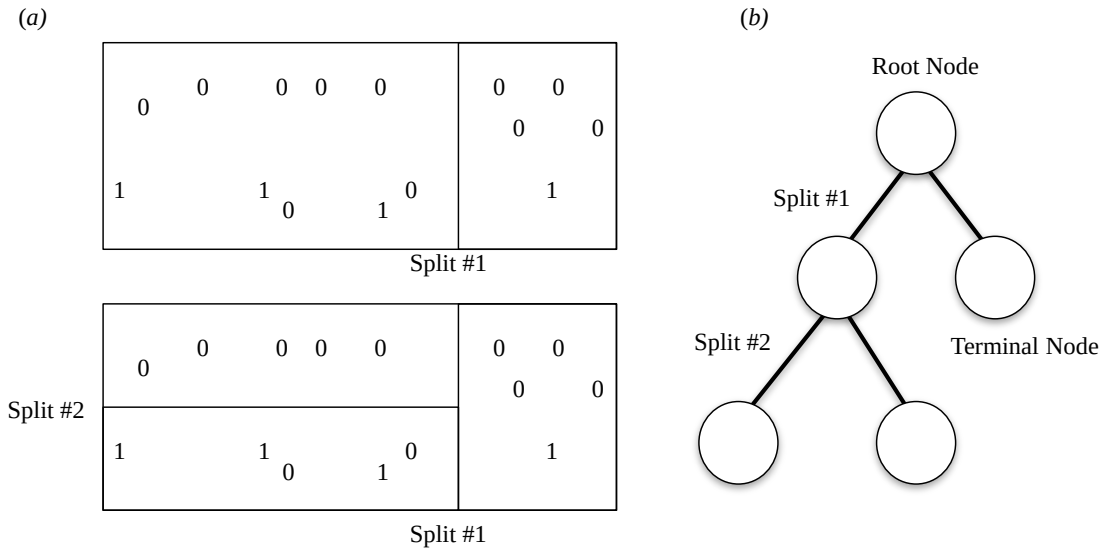


Figure 4. Rectangular partitions (a) and associated decision tree (b).

The creation of a decision tree primarily consists of three components: (1) the selection of points at which data can be split into nodes, (2) a decision regarding whether to terminate or continue splitting a node, and (3) the assignment of each node that is terminated (i.e., no longer split) to a particular outcome class (Breiman et al., 1993; Loh, 2011). Within the first component, the primary goal is to select a split that produces subsets of data that are "purer" than the data prior to the split. That is, individuals within each new node are better categorized regarding their status on a particular outcome. At each node, predictor variables are exhaustively evaluated on the basis of their potential to improve model purity. If a predictor variable is identified that improves the purity of the data, a split will be made. Within the second component, in order to determine whether to terminate or continue splitting a node, various

decision tree analyses use differing techniques such as stopping criteria and pruning. Pruning consists of constructing an overly complex tree and then removing splits to minimize the estimate of misclassification error, whereas, stopping criteria prevents a tree from over-developing if a pre-specified standard of model improvement is not achieved. Finally, the third component is easily accomplished by assigning a terminal node the outcome class that has the fewest misclassifications within the node (Loh, 2008).

Automatic Interaction Detection (AID). Designed by economists to predict economic events using a series of indicators, Automatic Interaction Detection (AID; Sonquist, Baker, & Morgan, 1971) was the first decision tree program ever developed (DeVille & Neville, 2013). The program was designed around the notion that statistical relationships are dependent on the context in which they exist, including the influence of other relationships within a model. This notion inspired researchers to attempt to develop a systematic statistical analysis that could take into account preceding relationships that could influence the presence and specific qualities of subsequent relationships. Attempts to design such an analysis led to the formation of decision trees via successively partitioning data to explore the impact of various categorizations on the outcome variable. In turn, researchers discovered that decision trees tend to outperform regression models when dealing with interactions and multicollinearity, both of which are common issues with regression (DeVille & Neville, 2013).

Although initial AID techniques produced decision trees with greater validity than regression models and higher levels of interpretability regarding complex interactions, several problems plagued its use (Doyle, 1973; Kass, 1975). Specifically, exhaustive searches of potential splits across all predictor variables tended to identify anomalous relationships and be biased toward predictors that had a large range of values. The latter issue is due to an increased

likelihood of identifying a split that demonstrates differences between groups using a variable with a greater range of observations (DeVille & Neville, 2013). That is, for example, if differences in a binary outcome are not found between the values of a dichotomized predictor variable (i.e., 0 and 1) the search will be terminated, whereas if differences are not found between the first two values of a continuous predictor variable, the search will continue with that variable until all values are investigated for differences (e.g., 1 and 2+, 1-2 and 3+, 1-3 and 4+, etc.).

In response to the shortcomings of AID, Kass (1980) designed Chi-Square Automatic Interaction Detection (CHAID). CHAID improved upon the procedures used in AID by incorporating significance testing to ensure that only groups of values that were significantly different from one another and not due to random chance were included as splits. Furthermore, CHAID conducted Bonferroni adjustments to avoid biased selection of variables with large numbers of values. Finally, CHAID terminated the decision tree growth process if significant differences could not be found between groups, thus eliminating the classification procedure's tendency to identify anomalous relationships (DeVille & Neville, 2013).

Classification and Regression Trees (CART). Following the development of AID and CHAID, Classification and Regression Trees (CART; Breiman et al., 1984) were developed. Similar to AID procedures, CART conducts an exhaustive search of all possible splits across all predictor variables. Moreover, CART produces fully-grown decision trees with all possible splits using the Gini index of diversity, which attempts to create maximum *homogeneity within* nodes and *heterogeneity between* nodes. They then undergo a pruning process that balances misclassification estimates and model complexity (Chien & Chen, 2008). Similar to the critiques of the AID procedure, CART has been also been accused of producing biased splits due to its

exhaustive search of all possible splits (Loh, 2009). In response to this limitation, CART employs cross-validation techniques to ensure model stability (DeVille & Neville, 2013). This feature, which verifies the accuracy and reproducibility of derived solutions, has made CART one of the most popular decision tree techniques.

Additional advantages of this technique include a sophisticated process of addressing missing values. Specifically, CART employs the use of surrogate variables, which have high correlations with the target predictor variable, when an individual has a missing value on the target predictor variable and it cannot be used in the model. In contrast, techniques such as CHAID model missing values as a node and reduce the validity and interpretability of the model. Finally, CART is a non-parametric procedure that makes no assumptions regarding the structure or normality of the data and can be conducted using widely available statistical software packages (e.g., SAS).

Decision-Making Criteria

Despite the widespread adoption of regression methods within the social sciences and emerging interest in decision tree analysis within other fields of study (e.g., medicine, fraud detection), large gaps exist within personnel decision-making regarding the use of one technique over the other. In order to assess the advantages and disadvantages of traditional and novel decision-making methods, each method's results must be evaluated in the context of several important decision-making criteria. Such criteria fall within three general categories: (1) validity of classification solutions, (2) presence of adverse impact, and (3) ease of interpretability. As extensive literature exists regarding the status of traditional regression techniques on the aforementioned criteria, the following section reviews each evaluative criterion and the expected advantages offered by decision tree analysis.

Validity

Schmidt and Hunter (1998) state that "from the point of view of practical value, the most important property of a personnel assessment method is predictive validity: the ability to predict future job performance, job-related learning, and other criteria" (p. 262). Furthermore, the authors cite the benefits of increased predictive validity, including increases in employee performance, organizational outputs, and financial gains (Schmidt & Hunter, 1998). Although such statements were made nearly two decades ago, validity still remains the most important component to any personnel decision-making system. Not only is validity the determining factor in the accuracy of a model, it is also the foundation for utility calculations, which assess the economic value of a system (Schmidt & Hunter, 1998).

Validity is critical to personnel decision-making due to the variability observed in organizational outcomes. Specifically, if variance did not exist in an outcome, such as job performance, then all job candidates would be evaluated as equal and expected to exhibit identical levels of subsequent job performance. Under such circumstances, organizations could make personnel decisions randomly and without the aid of selection systems. However, if extreme amounts of variance were present in the outcome, as is true for most modern jobs, the ability for an organization to determine the best candidates becomes paramount (Schmidt & Hunter, 1998). It is a system's validity that enables organizations to accurately describe and predict the variable relationship between candidate attributes and job performance. That is, the higher a system's validity, the more likely that organizational decisions are to be accurate.

When using a categorical outcome, binary logistic regression aims to produce the most valid solution possible by using maximum likelihood estimation to find the set of parameters for which the probability of the observed data is greatest (Czepiel, 2002). In contrast, decision tree

analysis attempts to produce the most valid solution by best parsing the observed data to create maximum homogeneity within groups and heterogeneity between groups. In both cases, validity can be assessed using classification accuracy, which demonstrates the percentage of correct classifications out of all possible predictions. Another method for evaluating the validity of classification solutions is the use of a receiver operating characteristic (ROC) curve (Cook, 2007). In general, ROC curve analysis evaluates the sensitivity and specificity of a model. A model's sensitivity and specificity is a result of its ability to balance correct predictions with false positive and false negative predictions. Classification accuracy and ROC curve analysis were used in the current study and several advantages of decision tree analysis were expected.

One characteristic of decision tree analysis that may lead to increased predictive accuracy is the non-parametric technique's lack of statistical assumptions. Personnel data can frequently violate the statistical assumptions of traditional techniques due to data characteristics such as heteroscedasticity. For example, in a review comparing decision tree analysis with traditional statistical techniques (i.e., linear discriminant analysis), Worth and Cronin (2003) found that the use of traditional techniques was not appropriate for modeling their study data on health classification. Specifically, they noted that decision tree analysis produces more accurate classification models when certain data characteristics are present, such as residuals that are not normally distributed (e.g., bimodal; Worth & Cronin, 2003). Furthermore, although regression techniques are often considered to be compensatory (Flynn & Peterson, 1972), especially when compared to other non-compensatory techniques such as multiple hurdle systems, this assumption may not always be accurate. In contrast, decision tree analysis is truly compensatory, which may result in increased classification accuracy and, as described in greater detail below, reduced adverse impact.

Adverse Impact

Adverse impact is a critically important topic for organizations to consider when employing a selection system or making personnel decisions. Adverse impact, as originally established by the Civil Rights Act of 1964, is an indicator of discriminatory organizational decisions. Specifically, Title VII of this act prohibits discriminatory employment practices in regard to race, sex, religion, age, and national origin (Civil Rights Act, 1964). Subsequently, organizations must be cautious when using selection instruments that are associated with meaningful subgroup differences. For example, cognitive ability tests result in significant mean group differences on the basis of race, such that minority respondents (e.g., Black/African-American) score considerably lower than non-minority respondents (e.g., White/Caucasian; Outtz & Newman, 2010). In order to determine if adverse impact is present in a selection system, practitioners perform a calculation known as the "four-fifths rule." Four-fifths rule calculations include dividing the selection rate of a protected minority class (e.g., African-Americans) by the selection rate of a majority group (e.g., Caucasians). Any selection ratio below .80, or a minority selection rate that is less than four-fifths of the non-minority selection rate, is indicative of adverse impact (Bobko & Roth, 2010). Although the four-fifths rule is descriptive in nature and does not imply statistical significance, the calculation is widely accepted as a standard for evaluating personnel decisions by legal and academic groups (Bobko & Roth, 2010).

The compensatory nature of decision tree analysis and its detection of multiple pathways to organizational outcomes hold a great deal of potential in regard to adverse impact. Specifically, decision tree analysis most likely provides the greatest benefit when used in a single hurdle system. The use of a multiple selection measures at a single time point allows candidates

to display any compensatory traits that may be present. This technique alone has the ability to decrease the presence of adverse impact. For example, Hunter and Hunter (1984) urged practitioners to use alternative predictors of job performance, such as personality, in addition to measures of general mental ability, in order to reduce the level of adverse impact present in such cognitive measures. However, decision tree analysis holds potential beyond simply using a single hurdle system.

Specifically, it is suggested that by identifying multiple pathways to outcomes, decision tree analysis is truly compensatory in that the relative importance of certain predictors may be different or entirely absent for some candidates in comparison to others. For example, after splitting candidates by cognitive ability, it is possible that although high cognitive performers may be further delineated by variables such as situational judgment, which contains aspects of general mental ability, low cognitive performers may be delineated by variables that are relatively independent from cognitive ability, such as integrity. Nevertheless, it is possible for both groups to include a large portion of individuals who are expected to be successful performers. In contrast, although regression equations in theory allow for compensation between variables, the relative importance of the variables included in the solution is fixed. Thus, regression solutions may present less of a compensatory opportunity than currently believed. In sum, both techniques must be evaluated regarding adverse impact in order to provide practitioners with the most useful and fair method.

Interpretability

The general interpretability of a statistical solution is particularly important when used by practitioners. Specifically, although a certain technique may provide robust technical solutions to classification problems, if such solutions are difficult to calculate or interpret, the results may

be misused. Consequently, it is important to evaluate the ease by which a technique is performed, the visual interpretability of its solution, and the way in which results can be communicated to others. The use of decision tree analysis potentially provides practitioners with greater levels of interpretability when making personnel decisions. This is primarily due to three aspects that characterize decision tree analyses: (1) visually appealing output, (2) precise cutoff points, and (3) simple execution of analytical procedures.

First, it is important that practitioners not only be capable of producing valid decision-making models, but also have the ability to clearly understand the derived results and share them with others (Farrington & Loeber, 2000). The use of decision trees satisfies both requirements by producing graphical statistical results, such as those exhibited by an example decision tree in Figure 5. More importantly, at each split in the decision tree, precise cutoff points are provided to assist organizations in the decision-making process. This facilitates quick categorization of individuals with high and low probabilities of falling into a desired outcome class (e.g., high performer). In contrast, information provided by linear regression models primarily consists of beta weights and semi-partial correlation coefficients, which may not be appropriately understood by organizations and result in misinterpretation. Moreover, clearer interpretation often requires additional dominance and/or relative importance analyses (e.g., Tonidandel & LeBreton, 2011). Third, with roots in data mining, the entire decision tree analysis process is relatively easy to implement. Although the user has several important considerations to take into account, once a proper decision tree analysis technique is selected that best fits a set of research questions and training data, modern statistical packages complete the analysis with very little researcher interaction and minimal computing time.

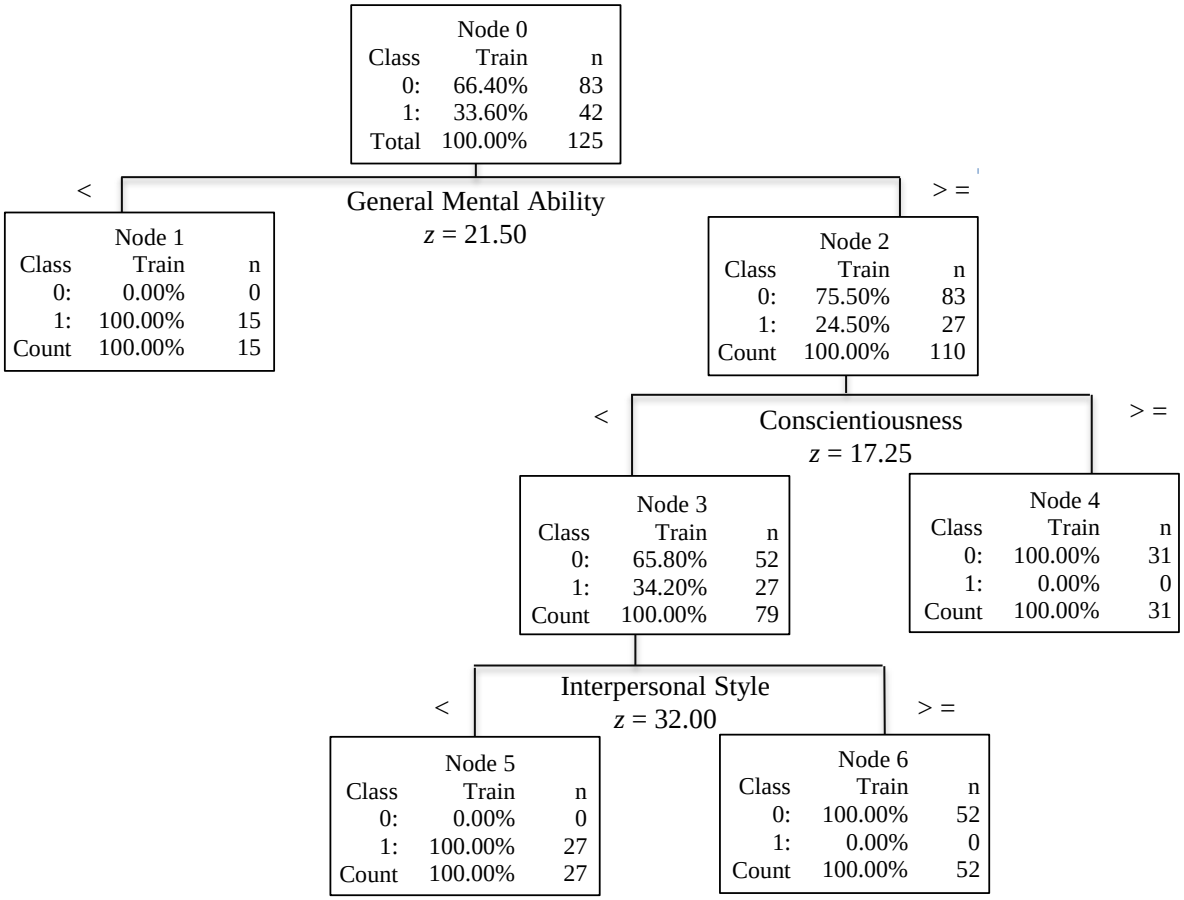


Figure 5. Sample decision tree analysis output.

Finally, decision tree analysis benefits organizations by clearly addressing complex data issues such as interactions and multicollinearity. Specifically, the use of recursive partitioning and multiple splits eliminates the effect of multicollinearity and identifies complex interactions by allowing groups to have differing slopes at each split (Gordon, 2013). Both of these factors make the use of traditional regression analyses substantially more difficult and result in diminished interpretability (Piper, Loh, Smith, Japuntich, & Baker, 2011). Decision tree analysis also allows organizations to identify non-linear data issues without any additional effort. Specifically, to successfully address this issue via traditional regression, intervention must occur with the data (e.g., fitting a quadratic term) and the true curvilinear nature of the relationship would likely not be ascertained (Connor, Symons, Feeney, Young, & Wiles, 2007; McDonald,

Poertner, & Harris, 2001). In contrast, utilizing a decision tree analysis that allows for multiple splits per node would address this issue automatically without any additional intervention (Gordon, 2013; see Figure 6).

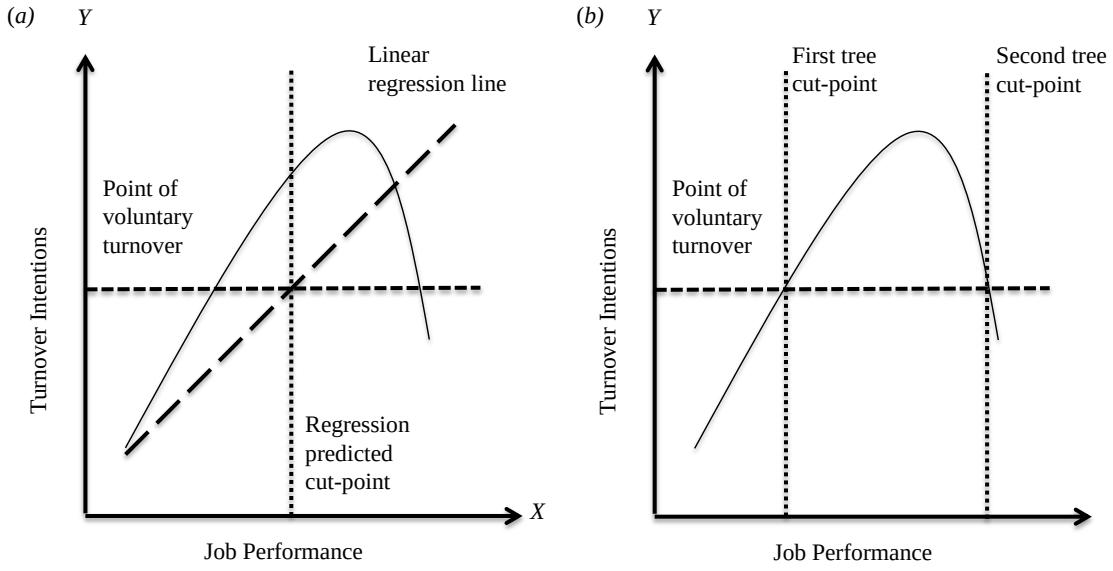


Figure 6. Sample regression solution that fails to model the depicted curvilinear relationship (a) and sample decision tree solution that models the depicted curvilinear relationship without additional specification (e.g., quadratic term; b).

Overall, decision tree analysis offers a range wide of potential benefits to practitioners. Furthermore, those who employ the procedure are not the only individuals to reap such benefits; rather, an entire field of study stands to benefit as researchers develop a greater understanding, produce best practices within their field, and provide generalizable results through the use of cross-validation procedures. Unfortunately, the successes experienced by other fields of study related to the use of decision tree analysis have yet to be explored within personnel decision-making. The current study consisted of a variety of important research questions that must be addressed prior to the widespread adoption of this technique within the field.

CHAPTER III: METHODOLOGY

Study 1

The relative advantages and disadvantages of decision tree analysis within the practice of employee selection were examined in Study 1. Specifically, binary logistic regression and decision tree analysis were used to model Monte Carlo simulated data on job performance. Job performance is one of the most studied criteria within the practice of employee selection and is, without exception, an extremely important outcome for all organizations (Campbell, 2012). Furthermore, job performance is utilized in selection systems as the leading dependent variable for making decisions to (1) initially hire a job candidate, (2) offer training to high performing incumbents, and (3) provide career guidance and interventions to low performing incumbents (Campbell, 2012). Consequently, it is vital to first investigate the use of decision tree analysis in modeling this important outcome. Additionally, by selecting empirically supported and commonly used predictors of job performance, the current study aimed to provide insight that is both useful and generalizable. Overall, Study 1 sought to answer the current research questions by modeling job performance via logistic regression and decision tree analysis, and subsequently comparing each method's model in regard to validity, adverse impact, and interpretability.

Procedure

Population parameters. In order to address the current research questions in the context of job performance, the following methodology was employed. First, a literature review was conducted to identify relationships between a common set of predictor variables and job performance. From this review, meta-analytical correlations (i.e., effect sizes) and demographic group mean standard deviation (*SD*) differences were acquired. The aim of such data was not to perfectly represent the strength of each relationship included in the study, as such precision is

beyond the scope of the current investigation, but rather to represent an acceptable approximation of the strength of each relationship provided by reputable sources in the literature.

Next, a correlation matrix featuring the collected effect sizes was constructed to serve as a population matrix for a series of Monte Carlo simulations. This correlation matrix was subsequently used to create a Cholesky decomposition from which a series of simulated sample data sets representing a range of organizational contexts were created. A single data set was created for each combination of parameters established in the current study. Parameters specified in the current study consisted of variations in sample size, job performance, and minority population. Each of the three study parameters consisted of two levels, resulting in eight total data sets that were created for the analysis (see Table 1).

Table 1. Data Set Parameters (Study 1)

	B_1	B_2
A_1C_1	Data Set 1 Sample Size: 1200 Minority Population: 80/20 Performance Cutoff: 0*z*	Data Set 2 Sample Size: 1200 Minority Population: 80/20 Performance Cutoff: -0.5*z*
A_1C_2	Data Set 3 Sample Size: 1200 Minority Population: 60/40 Performance Cutoff: 0*z*	Data Set 4 Sample Size: 1200 Minority Population: 60/40 Performance Cutoff: -0.5*z*
A_2C_1	Data Set 5 Sample Size: 400 Minority Population: 80/20 Performance Cutoff: 0*z*	Data Set 6 Sample Size: 400 Minority Population: 80/20 Performance Cutoff: -0.5*z*
A_2C_2	Data Set 7 Sample Size: 400 Minority Population: 60/40 Performance Cutoff: 0*z*	Data Set 8 Sample Size: 400 Minority Population: 60/40 Performance Cutoff: -0.5*z*

Notes.
A = Sample Size;
B = Job Performance Cutoff Point;
C = Minority Population.

Model variations. Specifically, data sets consisted of either a large or moderate sample size, a cutoff point for average or low job performance, and a minority population that is approximately representative of the current distribution in the United States or a minority population that would be expected from high minority recruitment. Model parameters were selected to represent useful values that may be experienced by organizations. Specifically, in the current study, large sample size data sets consisted of 1200 cases, whereas moderate sample size data sets consisted of 400 cases. For job performance, the cutoff point for average performance was set at a *z* score of 0, whereas the cutoff point for low performance was set at a *z* score of -0.5. Finally, data sets representing the U.S. minority distribution consisted of 80% White/Caucasian and 20% Black/African-American candidates, whereas data sets representing a high minority recruitment distribution consisted of 60% White/Caucasian and 40% Black/African-American candidates.

Each data set was created by first specifying the appropriate sample size and then generating a normally distributed data set via Monte Carlo simulation, which included a job performance outcome variable and eight predictor variables. In order to ensure high data quality, all simulated data sets' correlation matrices were investigated for discrepancies in comparison to the originally imputed correlation matrix. Comparisons between simulated data sets were conducted by calculating a mean discrepancy score for each simulated data set, which averaged all discrepancies observed between the simulated data set's correlation matrix and the originally imputed correlation matrix.

After iteratively creating and comparing data sets, the data set that deviated the least from the originally imputed correlation matrix was selected for each sample size. Next, a binary outcome variable was created within each data set by dichotomizing job performance based on

the parameter's two specified levels. Finally, race was randomly designated to all cases within each data set via a random number generator that assisted with specifying race in the appropriate proportions. Correlation matrices for each race class (e.g., Black/African-American White/Caucasian) were also investigated for discrepancies. Randomly generated assignments of race were compared and the assignment that produced correlation matrices that deviated the least from the simulated data set's correlation matrix was selected.

Model Variables

Big Five Personality. Since the early 1990s, researchers have dedicated a great deal of attention to the subject of personality testing within the context of personnel decision-making and selection (Barrick & Mount, 1991; Hurtz & Donovan, 2000). Personality is an important psychological construct to consider in such contexts due to its contribution in predicting job performance and its relative independence from other common predictors, such as cognitive ability (Barrick & Mount, 1991; McCrae & Costa, 1987). Over time, the measurement of personality has undergone many iterations and, at times, has included variables such as dependability, compliance, and gregariousness. Nevertheless, current conceptualizations of personality most commonly consist of five personality characteristics. These five personality characteristics, known as the "Big Five," include emotional stability, extraversion, openness to experience, agreeableness, and conscientiousness (Barrick & Mount, 1991). Big Five personality was included in the current study, as each of the dimensions are consistently shown to have a reliable relationship with job performance and are relatively independent from each other and additional predictors of job performance.

Cognitive ability. Cognitive ability, also known as general mental ability, is considered to be the most valid and reliable predictor of job performance across a wide range of jobs

(Hunter & Hunter, 1984). In addition to its high validity, cognitive ability is also known to be a relatively inexpensive predictor measure, as the construct can be reliably measured using paper-and-pencil instruments (Schmidt & Hunter, 1998). One explanation for the high value of cognitive ability measures in personnel decision-making notes that general mental ability is capable of measuring a job candidate's ability to have and use job knowledge, regardless of previous experience in the particular job of interest. Furthermore, general mental ability can be used to predict a candidate's ability to acquire new skills and information delivered during job-related training or learning. Finally, decades of research have helped to define and test the meaning of cognitive ability. In turn, organizations may employ measures of cognitive ability and interpret their results more confidently than more modern personnel decision-making methods, such as assessment centers (Schmidt & Hunter, 1998). As a result of the aforementioned strengths, cognitive ability was included in the current study.

Situational judgment. The use of situational judgment measures for personnel decision-making has dramatically increased in recent time. Situational judgment tests (SJTs) present job candidates with simulated situations that may be experienced on the job and a variety of possible responses to the situations. SJTs are commonly delivered in paper-and-pencil format and may use multiple response instructions. Knowledge SJTs ask candidates to select the correct or best possible response, whereas behavioral SJTs ask candidates to select the response that they would most likely select given they were actually presented with the situation (McDaniel, Hartman, Whetzel, & Grubb, 2007). Research on STJs suggest that the measures reliably predict job performance by assessing a construct that incorporates a variety of traits such as job knowledge, cognitive ability, and personality, as well as the traits' interactions with one another (McDaniel

et al., 2007). In turn, situation judgment was included in the current study as a promising, emergent predictor in the field of personnel decision-making.

Integrity. Over the last several decades, employers have become exceedingly interested in the honesty, integrity, and trustworthiness of their employees. As a result, the use of integrity tests has increased exponentially and there is now a substantial amount of both scientific and practical knowledge about the construct. Integrity tests are inexpensive paper-and-pencil measures that allow organization to assess a job candidate's level of integrity or trustworthiness, as well as his or her predisposition towards disruptive or counterproductive behavior such as theft, absenteeism, and disciplinary problems (Ones, 1993). Not only are integrity tests shown to reduce an organization's losses attributed to such problematic behaviors, but the construct is also predictive of job performance. Furthermore, the validity of integrity tests is demonstrated over a variety of situations, settings, and jobs (Ones, 1993). Consequently, integrity tests were included in the current study for their unique relationship to job performance, relatively inexpensive administration, and increasing popularity within personnel decision-making.

Data Analysis

Following data generation, a series of analyses were conducted on each of the simulated data sets. All analyses were conducted using SAS Enterprise Miner (v.14.1; SAS Institute Inc., 2013) and SPSS (v.22) statistical software packages. First, each data set was imported into the SAS statistical software package. Next, a data partition was used to randomly split each data set into a training and validation sample. Training samples consisted of 80% of each data set, whereas validation samples consisted of 20% of each data set. Next, a binary logistic regression and decision tree analysis were conducted simultaneously. No unique settings were selected for the logistic regression and the default classification cutoff (i.e., .5) was utilized for both analyses.

The default classification cutoff was determined to be sufficient for the current study, as falsely classified positive and negative cases were considered equal. However, in practice, organizational researchers may determine that there is a greater cost to incorrectly selecting or rejecting an individual. Accordingly, it may be determined that a certain classification cutoff should be specified, in order to maximize the model's true positive or true negative rate.

Classification and Regression Tree (CART; Breiman et al., 1984) was selected as the method for producing decision tree models in the current study. Consequently, the Gini index of diversity was specified as the splitting method prior to conducting each analysis. Following the specification of settings and completion of the analysis, each decision tree model underwent a pruning process that attempted to balance misclassification estimates, accuracy, and model complexity (Chien & Chen, 2008). Although this pruning process is a subjective procedure, feedback provided by the statistical software (e.g., change in model accuracy) helped to guide the process, similar to selecting variables in a regression model.

Following the development of both traditional (i.e., binary logistic regression) and novel (i.e., decision tree analysis) statistical solutions, model comparisons were made in regard to their (1) validity, (2) adverse impact, and (3) interpretability. Specifically, validity was assessed in the current study by comparing each model's overall classification accuracy and receiver operating characteristic (ROC) curve (Cook, 2007). Classification accuracy was calculated by dividing the sum of correctly classified positive (TP) and negative (TN) cases by the total number of cases in each model. ROC curve analysis provides a graphical plot that represents the portion of correctly classified positive cases (e.g., true positives; sensitivity), plotted on the Y-axis, and incorrectly classified positive cases (e.g., false positives; 1-specificity), plotted on the X-axis. Following the plotting of these two statistics, a line is generated that represents the discriminant

capacity of the model (Hajian-Tilaki, 2013). A diagonal line reflects that performance is no better than chance and contains 50% of the area within the plot beneath the designated line. In contrast, a concave line, or curve, depicts a greater level of discriminant capacity and possesses larger amounts of area under the curve. In sum, the total area under the ROC curve (AUC; Hajian-Tilaki, 2013) is an indicator of model classification accuracy. Consequently, each model's AUC was also used in order to assess and compare the validity of each model.

Next, adverse impact was compared using a four-fifths rule calculation (EEOC, 1978). Specifically, using each model, a predicted outcome value was created and assigned to each case in the sample data set. Such values were saved and each data set was imported into SPSS statistical software for further analysis. In particular, selection ratios (i.e., frequency of predicted class) were calculated for non-minority (i.e., White/Caucasian) and minority (i.e., Black/African-American) demographic groups in the sample. Finally, the selection ratio for the minority group was divided by the selection ratio for the non-minority group in order to determine if adverse impact was present in each model. Adverse impact was signified via an adjusted selection ratio below .80, or a minority selection rate that is less than four-fifths of the non-minority selection rate (Bobko & Roth, 2010). Model comparisons were based on a yes/no classification that describes the presence or absence of adverse impact, as well as a specific selection ratio for each model. Lastly, an informal evaluation of each solution's interpretability was conducted in order to best provide recommendations to the field regarding the use of each analysis.

Study 2

The relative advantages and disadvantages of decision tree analysis within the context of employee health were examined in Study 2. Specifically, binary logistic regression and decision tree analysis were also used in Study 2 to model Monte Carlo simulated data; however, the outcome of interest in the current study was subjective well-being. In turn, the aim of Study 2 was to add to the results derived from Study 1 by investigating the potential benefits of decision tree analysis in areas outside the context of performance, such as the emerging area of employee health (e.g., subjective well-being).

Subjective well-being refers to an individual's global assessment of his or her life and consists of both cognitive and affective dimensions. Specifically, subjective well-being encompasses life satisfaction, high positive affect, and low negative affect (Bowling, Eshleman, & Wang, 2010). Moreover, an individual's subjective appraisal of the aforementioned dimensions is relatively independent of any objective life conditions, such as wealth (Bowling et al., 2010). The construct of subjective well-being is an important area of study to both behavioral and organizational researchers for a variety of reasons (Bowling et al., 2010; Diener, 2009). For example, researchers consider happiness to be one of the strongest motivators of human action (Diener, 2009). Specifically, it is suggested that overall feelings of well-being encourage individuals to pursue goals and meet future challenges with the necessary resources to be successful (De Neve, Diener, Tay, & Xuereb, 2013). As a result, subjective well-being is associated with a range of important organizational outcomes. Such outcomes include: (1) employee health and longevity, (2) increased productivity, organizational citizenship, and income, and (3) improved individual and social behavior (De Neve et al., 2013). Consequently, it is vital to investigate the use of decision tree analysis in modeling this key outcome.

Procedure

Population parameters. In order to address the current research questions in the context of employee health, the following methodology was employed. First, a literature review was conducted to identify relationships between a common set of predictor variables and subjective well-being. Meta-analytical correlations were acquired and a correlation matrix was constructed to serve as a population matrix for Monte Carlo simulations. Similar to Study 1, the aim of such correlations was not to perfectly represent the strength of each variable's relationship with subjective well-being, but rather to represent each relationship with an acceptable degree of precision as is available in the literature. Next, a Cholesky decomposition was created and a series of simulated sample data sets were generated. In contrast to the three study parameters used in Study 1, Study 2 only consisted of one parameter—variations in subjective well-being.

Specifically, in comparison to Study 1, which examined variations in sample size, Study 2 used a single sample size to investigate employee health. Although the examination of various sample sizes is an important parameter within the context of personnel selection, as organizations may conduct selection procedures on both small and large candidate pools, when predicting employee health within an organization, all employees would likely be included in analyses. Consequently, a single large sample size was used in Study 2 to reflect the employee population within a moderately-sized organization. Next, although minority population differences are a key characteristic to personnel selection decisions, such considerations are of decreased importance to employee health models that aim to be inclusive and promote better health among all employees. Accordingly, minority population was not specified in Study 2. In turn, Study 2 only included variations in the cutoff points for subjective well-being. Such variations consisted of three levels and resulted in three total data sets that were created for the analysis (see Table 2).

Table 2. Data Set Parameters (Study 2)

	B_1	B_2	B_3
A_1	Data Set 1 Sample Size: 5000 Performance Cutoff: 0*z*	Data Set 2 Sample Size: 5000 Performance Cutoff: +0.5*z*	Data Set 3 Sample Size: 5000 Performance Cutoff: -0.5*z*

Notes.
A = Sample Size;
B = Subjective Well-Being Cutoff Point.

Model variations. In a similar fashion to Study 1, model parameters in Study 2 were selected to represent useful values that may be of interest to organizations. Specifically, in the current study, the sample size consisted of 5000 cases in order to represent the entire employee population of a moderate to large organization. For subjective well-being, the cutoff point for average employee health was set at a *z* score of 0, the cutoff point for high employee health was set at a *z* score of +0.5, and the cutoff point for low employee health was set at a *z* score of -0.5.

A single, normally distributed data set was generated via Monte Carlo simulation, which included 5000 cases, a subjective well-being outcome variable, and nine predictor variables. To ensure high data quality, the simulated data set's correlation matrix was investigated for discrepancies in comparison to the originally imputed correlation matrix by calculating both a mean and absolute value mean discrepancy score, as performed in Study 1. After the simulated data set was created, the subjective well-being outcome variable was dichotomized three times based on the levels specified by the model parameter. In sum, data set 1 consisted of 5000 cases and a subjective well-being cutoff point of 0*z*, data set 2 consisted of 5000 cases and a cutoff point of +0.5*z*, and data set 3 consisted of 5000 cases and a cutoff point of -0.5*z*.

Model Variables

Big Five Personality. Although many studies of subjective well-being have focused on biosocial indicators of the construct, such as age and sex, few strong relationships have been identified. Rather, researchers have found that personality, in most cases, is the strongest

determinant of an individual's subjective well-being (DeNeve & Cooper, 1998). Evidence supporting this notion suggests that personality demonstrates both the stability and lack of specificity required to affect an individual's global assessment of satisfaction across a variety of life domains (DeNeve & Cooper, 1998; Diener, 1984). The personality characteristics known as the "Big Five" (i.e., emotional stability, extraversion, openness to experience, agreeableness, and conscientiousness; Barrick & Mount, 1991) are instances of traits known to have a relationship with subjective well-being. For example, extraversion is associated with experiences of positive affect and cheerfulness; agreeableness is related to enhanced relationship quality; and conscientiousness promotes the achievement of goals and tasks (DeNeve & Cooper, 1998; McCrae & Costa, 1991). As a result, the Big Five personality was included in the current study, as each trait is shown to have a consistent relationship with subjective well-being.

Self-esteem. Self-esteem is included within a group of personality traits that characterize an individual's "core self-evaluation," or fundamental conclusions about oneself (Judge, Locke, & Durham, 1997). Specifically, self-esteem refers to the overall value that one places on oneself as a person, and is considered a primary manifestation of core self-evaluation (Judge & Bono, 2001). Moreover, self-esteem represents the degree to which individuals evaluate themselves as competent and personally satisfying (Pierce & Gardner, 2004). In turn, self-esteem is related to various important outcomes such as life satisfaction, job satisfaction, organizational commitment and citizenship, and motivation (Judge, Erez, Bono, & Thorensen, 2002; Pierce & Gardner, 2004). As a result, self-esteem was included in the current study.

Generalized self-efficacy. Generalized self-efficacy is also included within the group of traits that define an individual's core self-evaluation. Generalized self-efficacy refers to an individual's estimate of his or her overall ability to cope, perform, and be successful (Judge &

Bono, 2001). Furthermore, this assessment of oneself applies to a wide variety of achievement situations, task demands, and environmental contexts (Chen, Gully, & Eden, 2001). Generalized self-efficacy is also posited to influence behavior such as the type of activities an individual chooses to engage in, the amount of effort expended, and the level of perseverance displayed amidst challenges (Tipton & Worthington, 1984). In turn, the construct is related to broad outcomes including health, physical activity, life and job satisfaction, training proficiency, and job performance (Chen et al., 2001; Judge et al., 2002; Maibach & Murphy, 1995). Consequently, generalized self-efficacy was included in the current study.

Internal locus of control. A third variable included in an individual's core self-evaluation is internal locus of control. Internal locus of control refers to people's beliefs that they can control a broad range of factors that may occur in their lives (Judge & Bono, 2001). As a result, an internal control orientation is related to positive outcomes in cognitive processing, motivation, autonomy, and self-reliance (Duttweiler, 1984). In addition to broad outcomes, internal locus of control is also known to have a particularly positive effect on health outcomes. Specifically, the construct is related to smoking cessation, preventative health behaviors, and physical activity (Wallston & Wallston, 1978). In turn, internal locus of control is an important variable to consider in regard to subjective well-being both within and outside of the organizational context. Consequently, the variable was included in the current study.

Job satisfaction. Several studies have demonstrated a consistent relationship between job satisfaction and subjective well-being. Job satisfaction is defined as "a pleasurable or positive emotional state resulting from the appraisal of one's job or job experiences" (Locke, 1976, p. 1304). Research suggests that this positive emotional state and appraisal is subsequently related to positive relationships with happiness, positive affect, and subjective well-

being as a whole (Bowling et al., 2010). One explanation for such findings is provided by the spillover hypothesis, which posits that positive experiences at work may influence other non-work areas that may be considered when assessing subjective well-being, such as marital and leisure domains (Bowling et al., 2010). Job satisfaction is also known to improve other important outcomes such as employee health and job performance (Faragher, Cass, & Cooper, 2005; Judge, Thoresen, Bono, & Patton, 2001). As a result of the benefits of job satisfaction and its relationship with subjective well-being, this construct was included in the current study.

Data Analysis

Following data generation, a series of analyses were performed on each of the three simulated data sets. First, each data set was imported into SAS Enterprise Miner and a data partition was applied to randomly split the data set into a training and validation sample. Training and validation samples consisted of 80% and 20% of each data set, respectively. Next, a binary logistic regression and decision tree analysis were conducted simultaneously. Study 2 utilized identical procedure settings as were used in Study 1. In particular, the binary logistic regression included a default classification cutoff of .5 and the CART method was selected for producing decision tree models. Following the specification of procedure settings and completion of the analysis, each decision tree model underwent a pruning process guided by feedback provided by the statistical software (e.g., change in model accuracy). Lastly, the binary logistic regression and decision tree analysis statistical solution for each data set was compared in regard to their validity and interpretability. Adverse impact was not assessed for Study 2, which aimed to inclusively model employee health and not inform selection decisions that must consider the potential for inadvertent discrimination against certain demographic groups.

CHAPTER IV: RESULTS

Study 1

Correlations and Mean Group Differences

Correlations (i.e., effect sizes) and demographic group mean standard deviation (*SD*) differences for the current study are shown in Tables 3 and 4, respectively. Variables included in the current study consisted of both common and emerging predictors of job performance. Furthermore, each variable's importance to personnel selection is well established by meta-analyses and all variables' correlation with job performance was obtained from such sources. Variables included in the current study consisted of 1) Job Performance, 2) Emotional Stability, 3) Extraversion, 4) Openness to Experience, 5) Agreeableness, 6) Conscientiousness, 7) Cognitive Ability, 8) Situational Judgment, and 9) Integrity.

As each variable's correlation with job performance was obtained via meta-analytic sources, no single instrument is specified as the source of the effect size. Rather, most correlations are the aggregate of many effect sizes derived from a variety of measures. For example, correlations between variables included in the Big Five (Barrick & Mount, 1991) and job performance include effect sizes derived from the following measures: the NEO Personality Inventory (NEO-PI), the NEO Personality Inventory-Revised (NEO-PI-R), and the Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992), as well as Goldberg's Big Five (Goldberg, 1992), the Hogan Personality Inventory (HPI; Hogan & Hogan, 1995), and the Personal Characteristics Inventory (PCI; Barrick & Mount, 1993). Lastly, demographic group mean *SD* differences on the basis of race (e.g., Black/African-American, White/Caucasian) were also collected from meta-analysis sources in the literature.

Review of meta-analysis sources identified a positive correlation between Big Five personality and Job Performance (Hurtz & Donovan, 2000). Specifically, correlations between

Big Five personality and Job Performance included, Emotional Stability (.14), Extraversion (.10), Openness to Experience (.07), Agreeableness (.13), and Conscientiousness (.22). Additionally, all Big Five personality variables were found to have a positive correlation with one another (Mount, Barrick, Scullen, & Rounds, 2005; see Table 3). Next, a positive correlation was identified between Cognitive Ability and Job Performance (.51; Hunter, 1980; Schmidt & Hunter, 1998). Furthermore, Cognitive Ability had a positive correlation with all Big Five personality variables, with the exception of Conscientiousness, which had a small negative correlation with Cognitive Ability (-.04; Judge, Jackson, Shaw, Scott, & Rich, 2007; see Table 3). Situational Judgment was also found to have a positive correlation with Job Performance (.26), as well as all Big Five personality variables and Cognitive Ability (McDaniel et al., 2007; see Table 3). Finally, Integrity (i.e., Integrity Tests) was found to have a positive correlation with Job Performance (.34) and all other study variables, with the exception of Extraversion, which had a small negative correlation with Integrity (-.08; Meijer, Born, Zielst, & Molen, 2010; Ones, 1993; Ones, Viswesvaran, & Schmidt, 1993; see Table 3).

Table 3. Meta-Analytic Correlations (Study 1)

Variable	1	2	3	4	5	6	7	8	9
1. Job Performance	1.00								
2. Emotional Stability	.14	1.00							
3. Extraversion	.10	.24	1.00						
4. Openness to Experience	.07	.19	.45	1.00					
5. Agreeableness	.13	.42	.26	.17	1.00				
6. Conscientiousness	.22	.52	.17	.09	.39	1.00			
7. Cognitive Ability	.51	.09	.02	.22	.00	-.04	1.00		
8. Situational Judgment	.26	.22	.14	.13	.25	.27	.32	1.00	
9. Integrity Tests	.34	.33	-.08	.12	.40	.42	.02	.30	1.00

Demographic group mean *SD* differences were also obtained via meta-analytic sources and group mean *SD* differences on the basis of race were identified for nearly all study variables (see Table 4). In regard to Big Five personality, mean *SD* differences were found between

Black/African-American and White/Caucasian individuals for all variables (Foldes, Duehr, & Ones, 2008). Specifically, White/Caucasian individuals demonstrate higher mean scores than Black/African-American individuals on Emotional Stability (*SD* = +.09), Extraversion (*SD* = +.16), Openness to Experience (*SD* = +.10), and Agreeableness (*SD* = +.03), whereas Black/African-American individuals demonstrate higher mean scores than White/Caucasian individuals on Conscientiousness (*SD* = +.07). Next, mean *SD* differences were also found between Black/African-American and White/Caucasian individuals on Cognitive Ability (Hunter & Hunter, 1984) and Situational Judgment (Whetzel, McDaniel, & Nguyen, 2008). Specifically, White/Caucasian individuals demonstrate higher mean scores than Black/African-American individuals on both Cognitive Ability (*SD* = +1.00) and Situational Judgment (*SD* = +.38). No mean *SD* differences were found for Integrity Tests (Berry, Sackett, & Wiemann, 2007).

Table 4. Demographic Group Mean Standard Deviation (*SD*) Differences

Variable	Black/African-American	White/Caucasian
1. Job Performance	-	-
2. Emotional Stability	-	+.09
3. Extraversion	-	+.16
4. Openness to Experience	-	+.10
5. Agreeableness	-	+.03
6. Conscientiousness	+.07	-
7. Cognitive Ability	-	+1.00
8. Situational Judgment	-	+.38
9. Integrity Tests	-	-

Data Generation and Assessment

Following the generation of data via Monte Carlo simulation, each simulated data set's correlation matrix was assessed for discrepancies in comparison to the originally imputed correlation matrix. The current study consists of a total of eight data sets; however, the specification of a job performance cutoff point within each data set did not influence correlations between the other study variables. Consequently, only four data sets, consisting of variations in

sample size (i.e., 1200 cases; 400 cases) and minority distribution (i.e., 80% White/Caucasian, 20% Black/African-American; 60% White/Caucasian, 40% Black/African-American) were assessed. Furthermore, each data set was assessed at three levels, which included 1) the overall data set, 2) the Black/African-American portion of the overall data set, and 3) the White/Caucasian portion of the overall data set. Each data set's discrepancies are displayed as both a mean discrepancy score and an absolute value mean discrepancy score in Table 5.

Overall, a data set's mean discrepancy score represents how balanced discrepancies are between the simulated data set's correlation matrix and the originally imputed correlation matrix, as discrepancies can be either greater than or less than the originally imputed value. In comparison, a data set's absolute value mean discrepancy score represents a total deviation from the originally imputed matrix, regardless of direction (i.e., positive or negative). When comparing the overall data sets to their originally imputed correlation matrices, mean discrepancy scores ranged from -.004 to .000 and absolute value mean discrepancy scores ranged from -.015 to .022 (see Table 5). When comparing the data sets' racial distributions, mean discrepancy scores ranged from -.003 to .002 for Black/African-American cases and from -.001 to .002 for White/Caucasian cases (see Table 5). Additionally, absolute value mean discrepancy scores ranged from .019 to .046 for Black/African-American cases and from .010 to .023 for White/Caucasian cases (see Table 5).

Table 5. Data Set Correlation Discrepancies

Level of Comparison	Discrepancy Type	Discrepancy			
		1	2	3	4
Overall	Mean	-.004	-.004	.000	.000
	Absolute Value Mean	.015	.015	.022	.022
B/A-A	Mean	-.001	-.003	.002	-.003
	Absolute Value Mean	.039	.019	.046	.034
W/C	Mean	.000	.002	-.001	.002
	Absolute Value Mean	.010	.013	.011	.023

Notes.
B/A-A = Black/African-American; W/C = White/Caucasian;
1: *N* = 1200, 80% White/Caucasian, 20% Black/African-American;
2: *N* = 1200, 60% White/Caucasian, 40% Black/African-American;
3: *N* = 400, 80% White/Caucasian, 20% Black/African-American;
4: *N* = 400, 60% White/Caucasian, 40% Black/African-American.

Test of Research Questions

Following the creation of a binary logistic regression and decision tree analysis solution, each model's validity, adverse impact, and interpretability was assessed. Furthermore, each model's validity was assessed using both a training and validation sample to evaluate model stability. Validity was demonstrated via classification accuracy, which was calculated as a total percentage of correct classifications derived from the model, as well as area under the ROC curve (AUC), which was exhibited as a value between 0 and 1 and represents the discriminant capacity of the model. Higher values on both indicators of validity were indicative of a more accurate model (see Table 6). Next, adverse impact was demonstrated via a four-fifths adjusted selection ratio, which compared the selection rate of the two race classes. A selection ratio was calculated by dividing the selection rate (i.e., percentage of individuals selected) of the minority class by the selection rate of the non-minority class. Any selection ratio below .80 was indicative of adverse impact (see Table 7). Lastly, interpretability was assessed via an explanation of each model and included an analysis of model structure, variable importance, and

clarity of cutoff points. Specifically, odds ratios were used as an indicator of variable importance for the logistic regression model and the order in which variables appeared in the decision tree was used as an indicator of variable importance for the decision tree model. Validity indicators and adverse impact for models within each data set are shown in Tables 6 and 7, respectively, and complete decision tree models can be found in Appendix A.

Table 6. Model Classification Accuracy and AUC Index (Study 1)

Data Set	Model	Classification Accuracy		AUC Index	
		Training	Validation	Training	Validation
1	Binary Logistic Regression	67.5%	70.7%	.75	.78
	Decision Tree Analysis	67.4%	70.2%	.73	.76
2	Binary Logistic Regression	74.5%	77.1%	.77	.77
	Decision Tree Analysis	74.7%	70.1%	.75	.71
3	Binary Logistic Regression	66.1%	72.3%	.74	.80
	Decision Tree Analysis	69.0%	66.9%	.74	.73
4	Binary Logistic Regression	74.8%	73.9%	.76	.79
	Decision Tree Analysis	74.3%	73.9%	.72	.71
5	Binary Logistic Regression	70.4%	65.9%	.77	.73
	Decision Tree Analysis	67.3%	62.2%	.72	.66
6	Binary Logistic Regression	76.7%	74.4%	.78	.75
	Decision Tree Analysis	72.6%	70.7%	.73	.66
7	Binary Logistic Regression	65.7%	69.5%	.74	.75
	Decision Tree Analysis	66.7%	68.3%	.72	.69
8	Binary Logistic Regression	75.8%	70.7%	.77	.71
	Decision Tree Analysis	74.8%	73.2%	.73	.70

Notes.
1: *N* = 1200, 80% White/Caucasian, 20% Black/African-American, 0*z*;
2: *N* = 1200, 80% White/Caucasian, 20% Black/African-American, -0.5*z*;
3: *N* = 1200, 60% White/Caucasian, 40% Black/African-American, 0*z*;
4: *N* = 1200, 60% White/Caucasian, 40% Black/African-American, -0.5*z*;
5: *N* = 400, 80% White/Caucasian, 20% Black/African-American, 0*z*;
6: *N* = 400, 80% White/Caucasian, 20% Black/African-American, -0.5*z*;
7: *N* = 400, 60% White/Caucasian, 40% Black/African-American, 0*z*;
8: *N* = 400, 60% White/Caucasian, 40% Black/African-American, -0.5*z*.

Table 7. Model Adverse Impact

Data Set	Model	Non-Minority Selection Rate	Minority Selection Rate	Four-Fifths Adjusted Selection Ratio	AI
1	Binary Logistic Regression	55.4%	26.7%	.48	Yes
	Decision Tree Analysis	54.1%	32.9%	.61	Yes
2	Binary Logistic Regression	85.2%	65.8%	.77	Yes
	Decision Tree Analysis	87.1%	78.8%	.91	No
3	Binary Logistic Regression	61.9%	34.2%	.55	Yes
	Decision Tree Analysis	47.9%	31.7%	.66	Yes
4	Binary Logistic Regression	89.4%	71.9%	.80	No
	Decision Tree Analysis	86.0%	75.0%	.87	No
5	Binary Logistic Regression	44.4%	26.3%	.59	Yes
	Decision Tree Analysis	37.2%	33.8%	.91	No
6	Binary Logistic Regression	87.2%	68.8%	.79	Yes
	Decision Tree Analysis	89.4%	90.0%	1.00	No
7	Binary Logistic Regression	50.4%	25.0%	.50	Yes
	Decision Tree Analysis	24.6%	23.8%	.97	No
8	Binary Logistic Regression	89.2%	75.0%	.84	No
	Decision Tree Analysis	84.6%	73.8%	.87	No

Notes.
AI = Adverse Impact;
1: *N* = 1200, 80% White/Caucasian, 20% Black/African-American, 0*z*;
2: *N* = 1200, 80% White/Caucasian, 20% Black/African-American, -0.5*z*;
3: *N* = 1200, 60% White/Caucasian, 40% Black/African-American, 0*z*;
4: *N* = 1200, 60% White/Caucasian, 40% Black/African-American, -0.5*z*;
5: *N* = 400, 80% White/Caucasian, 20% Black/African-American, 0*z*;
6: *N* = 400, 80% White/Caucasian, 20% Black/African-American, -0.5*z*;
7: *N* = 400, 60% White/Caucasian, 40% Black/African-American, 0*z*;
8: *N* = 400, 60% White/Caucasian, 40% Black/African-American, -0.5*z*.

Data set 1. Data set 1 consisted of a sample of 1200 cases. Additionally, the data set was characterized by a minority distribution of 80% White/Caucasian and 20% Black/African-American individuals and a job performance cutoff point of 0*z*. Data set 1 aimed to represent a relatively large sample that featured a minority distribution similar to that of the U.S. and an average job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 67.5% and 70.7% for training and validation samples, respectively. In comparison,

the decision tree model had a classification accuracy of 67.4% and 70.2% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .75 and .78 for training and validation samples, respectively, and the decision tree model had an AUC index of .73 and .76 for training and validation samples, respectively. Overall, logistic regression demonstrated slightly higher classification accuracy and discriminant capacity than decision tree analysis in data set 1.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 55.4% and a minority (i.e., Black/African-American) selection rate of 26.7%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .48 and thus indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 54.1% and a minority selection rate of 32.9%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .61 and thus indicative of adverse impact. Overall, both models were associated with adverse impact; however, decision tree analysis had a higher (i.e., less discriminatory) adjusted selection ratio than logistic regression.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 1200) = 209.26, p < .001. Analyses revealed that the logistic regression model consisted of seven significant predictor variables: Emotional Stability, B = -.25, $\chi^2(1)$ = 7.28, p = .007, *Exp(B)* = .78, 95% CI [-.43, -.07]; Extraversion, B = .41, $\chi^2(1)$ = 20.25, p < .001, *Exp(B)* = 1.50, 95% CI [.23, .58]; Openness to Experience, B = -.27, $\chi^2(1)$ = 9.90, p = .002, *Exp(B)* = .76, 95% CI [-.44, -.10]; Agreeableness, B = -.19, $\chi^2(1)$ = 4.33, p = .04, *Exp(B)* = .83, 95% CI [-.36, -.01]; Conscientiousness, B = .30, $\chi^2(1)$ = 10.43, p = .001, *Exp(B)* = 1.35, 95% CI [.12, .48]; Cognitive Ability, B = .87, $\chi^2(1)$ = 104.55, p < .001, *Exp(B)* = 2.38, 95%

CI [.70, 1.03]; and Integrity, $B = .72$, $\chi^2(1) = 53.44$, $p < .001$, *Exp(B)* = 2.04, 95% CI [.52, .91]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Extraversion, Conscientiousness, Openness to Experience, Emotional Stability, and Agreeableness. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Emotional Stability, Openness to Experience, and Agreeableness were associated with an increased likelihood of demonstrating an acceptable level of job performance. Furthermore, the logistic regression model identified no specific cutoff points in any analyses.

In comparison, the decision tree model consisted of five predictor variables, including Extraversion, Conscientiousness, Cognitive Ability, Integrity, and Situational Judgment. The first variable included in the decision tree model was Cognitive Ability, followed by Integrity and Conscientiousness (appeared simultaneously), Situational Judgment, and Extraversion. Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, eight specific cutoff points were identified, resulting in nine branches and nine terminal nodes (see Appendix A). Finally, the decision tree model consisted of six terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 1 provided key insight into the current study's research questions (RQs) and, in general, suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1). Specifically, although logistic regression demonstrated slightly higher classification accuracy and discriminant capacity, the margin by which it out-performed decision tree analysis was extremely small. In turn, the use of decision tree analysis is suggested to benefit decision-makers by performing as

well as logistic regression under data conditions similar to data set 1, and resulting in equally accurate personnel decisions (RQ_2). Moreover, although both models were associated with adverse impact, decision tree analysis resulted in a less-discriminatory model than logistic regression – demonstrating potential benefits to non-discriminatory hiring (RQ_3). Decision tree analysis also produced a more parsimonious model than logistic regression by using fewer predictor variables to model the outcome with the same degree of accuracy. Lastly, decision tree analysis identified discrete cutoff points, thus offering benefits to interpretability not demonstrated by logistic regression (RQ_4).

Data set 2. Data set 2 consisted of a sample of 1200 cases. Additionally, the data set was characterized by a minority distribution of 80% White/Caucasian and 20% Black/African-American individuals and a job performance cutoff point of -0.5*z*. Data set 2 aimed to represent a relatively large sample that featured a minority distribution similar to that of the U.S. and a low job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 74.5% and 77.1% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 74.7% and 70.1% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .77 and .77 for training and validation samples, respectively, and the decision tree model had an AUC index of .75 and .71 for training and validation samples, respectively. Overall, decision tree analysis demonstrated slightly higher classification accuracy in the training sample of data set 2, whereas logistic regression demonstrated higher classification accuracy in the smaller, validation sample. In contrast, logistic regression demonstrated greater discriminant capacity than decision tree analysis in both samples.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 85.2% and a minority (i.e., Black/African-American) selection rate of 65.8%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .77 and thus indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 87.1% and a minority selection rate of 78.8%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .91 and thus not indicative of adverse impact. Overall, in comparison to logistic regression in data set 2, decision tree analysis demonstrated superior performance regarding adverse impact.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 1200) = 201.38, $p < .001$. Analyses revealed that the logistic regression model consisted of six significant predictor variables: Extraversion, $B = .35$, $\chi^2(1) = 12.69$, $p < .001$, *Exp(B)* = 1.42, 95% CI [.16, .55]; Openness to Experience, $B = -.20$, $\chi^2(1) = 4.52$, $p = .03$, *Exp(B)* = .82, 95% CI [-.38, -.02]; Agreeableness, $B = -.28$, $\chi^2(1) = 8.59$, $p = .003$, *Exp(B)* = .75, 95% CI [-.47, -.09]; Conscientiousness, $B = .38$, $\chi^2(1) = 13.96$, $p < .001$, *Exp(B)* = 1.47, 95% CI [.18, .58]; Cognitive Ability, $B = .81$, $\chi^2(1) = 81.27$, $p < .001$, *Exp(B)* = 2.25, 95% CI [.64, .99]; and Integrity, $B = .76$, $\chi^2(1) = 51.09$, $p < .001$, *Exp(B)* = 2.14, 95% CI [.55, .97]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Conscientiousness, Extraversion, Agreeableness, and Openness to Experience. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Agreeableness and Openness to Experience were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of five predictor variables, including Extraversion, Conscientiousness, Cognitive Ability, Integrity, and Situational Judgment. The first variable included in the decision tree model was Cognitive Ability, followed by Integrity and Conscientiousness, which appeared simultaneously, Situational Judgment, and Extraversion. Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, eight specific cutoff points were identified, resulting in nine branches and nine terminal nodes (see Appendix A). Finally, the decision tree model consisted of six terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 2 suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1). Specifically, decision tree analysis demonstrated the potential to provide benefits to classification accuracy by producing a slightly more accurate model than logistic regression in the training sample (RQ_2). However, such benefits were not extended to the smaller, validation sample or the discriminant capacity of the model (i.e., a model's ability to correctly classify both positive and negative cases). In contrast, decision tree analysis demonstrated clear benefits to non-discriminatory hiring, as decision tree analysis produced a model void of adverse impact, whereas logistic regression did not (RQ_3). Finally, decision tree analysis again produced a more parsimonious model and identified discrete cutoff points, thus offering benefits to interpretability not demonstrated by logistic regression (RQ_4).

Data set 3. Data set 3 consisted of a sample of 1200 cases. Additionally, the data set was characterized by a minority distribution of 60% White/Caucasian and 40% Black/African-American individuals and a job performance cutoff point of 0*z*. Data set 3 aimed to represent a

relatively large sample that featured a minority distribution that would be expected from high minority recruitment and an average job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 66.1% and 72.3% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 69.0% and 66.9% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .74 and .80 for training and validation samples, respectively, and the decision tree model had an AUC index of .74 and .73 for training and validation samples, respectively. Overall, decision tree analysis demonstrated higher classification accuracy in the training sample of data set 3, whereas logistic regression demonstrated higher classification accuracy in the smaller, validation sample. Moreover, the two models demonstrated the same degree of discriminant capacity in the training sample; however, logistic regression demonstrated greater discriminant capacity in the validation sample.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 61.9% and a minority (i.e., Black/African-American) selection rate of 34.2%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .55 and thus indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 47.9% and a minority selection rate of 31.7%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .66 and thus indicative of adverse impact. Overall, both models were associated with adverse impact; however, decision tree analysis had a higher (i.e., less discriminatory) adjusted selection ratio than logistic regression.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 1200) = 188.78, p < .001. Analyses revealed that the

logistic regression model consisted of seven significant predictor variables: Emotional Stability, B = -.26, $\chi^2(1)$ = 8.01, p = .005, *Exp(B)* = .77, 95% CI [-.44, -.08]; Extraversion, B = .38, $\chi^2(1)$ = 18.16, p < .001, *Exp(B)* = 1.46, 95% CI [.20, .55]; Openness to Experience, B = -.24, $\chi^2(1)$ = 8.39, p = .004, *Exp(B)* = .79, 95% CI [-.40, -.08]; Agreeableness, B = -.18, $\chi^2(1)$ = 4.09, p = .04, *Exp(B)* = .84, 95% CI [-.35, -.01]; Conscientiousness, B = .37, $\chi^2(1)$ = 15.18, p < .001, *Exp(B)* = 1.44, 95% CI [.18, .55]; Cognitive Ability, B = .81, $\chi^2(1)$ = 97.14, p < .001, *Exp(B)* = 2.24, 95% CI [.65, .97]; and Integrity, B = .64, $\chi^2(1)$ = 42.13, p < .001, *Exp(B)* = 1.89, 95% CI [.44, .83]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Extraversion, Conscientiousness, Agreeableness, Emotional Stability, and Openness to Experience. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Emotional Stability, Openness to Experience, and Agreeableness were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of four predictor variables, including Extraversion, Conscientiousness, Cognitive Ability, and Integrity. The first variable included in the decision tree model was Cognitive Ability, followed by Integrity, Conscientiousness, and Extraversion. Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, eight specific cutoff points were identified, resulting in nine branches and nine terminal nodes (see Appendix A). Finally, the decision tree model consisted of four terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 3 suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1). Specifically, decision tree analysis demonstrated the potential to provide benefits to classification accuracy by producing an equally discriminant and slightly more accurate model than logistic regression in the training sample (RQ_2). Furthermore, although both models were associated with adverse impact, decision tree analysis resulted in a less-discriminatory model than logistic regression, thus offering potential benefits to non-discriminatory hiring (RQ_3). Finally, decision tree analysis demonstrated potential benefits to interpretability by producing a more parsimonious model than logistic regression and identifying discrete cutoff points (RQ_4).

Data set 4. Data set 4 consisted of a sample of 1200 cases. Additionally, the data set was characterized by a minority distribution of 60% White/Caucasian and 40% Black/African-American individuals and a job performance cutoff point of -0.5*z*. Data set 4 aimed to represent a relatively large sample that featured a minority distribution that would be expected from high minority recruitment and a low job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 74.8% and 73.9% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 74.3% and 73.9% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .76 and .79 for training and validation samples, respectively, and the decision tree model had an AUC index of .72 and .71 for training and validation samples, respectively. Overall, logistic regression demonstrated slightly higher classification accuracy in the training sample of data set 4; however, the two models demonstrated the same level of classification accuracy in the validation sample. In contrast,

logistic regression demonstrated greater discriminant capacity than decision tree analysis in both samples.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 89.4% and a minority (i.e., Black/African-American) selection rate of 71.9%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .80 and thus not indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 86.0% and a minority selection rate of 75.0%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was.87 and thus not indicative of adverse impact. Overall, both models were not associated with adverse impact; however, it is noteworthy that decision tree analysis had a higher adjusted selection ratio than logistic regression.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 1200) = 178.63, $p < .001$. Analyses revealed that the logistic regression model consisted of five significant predictor variables: Extraversion, $B = .36$, $\chi^2(1) = 14.18$, $p < .001$, *Exp(B)* = 1.44, 95% CI [.17, .55]; Agreeableness, $B = -.33$, $\chi^2(1) = 11.53$, $p < .001$, *Exp(B)* = .72, 95% CI [-.52, -.14]; Conscientiousness, $B = .40$, $\chi^2(1) = 14.98$, $p < .001$, *Exp(B)* = 1.49, 95% CI [.20, .60]; Cognitive Ability, $B = .78$, $\chi^2(1) = 78.73$, $p < .001$, *Exp(B)* = 2.17, 95% CI [.61, .95]; and Integrity, $B = .71$, $\chi^2(1) = 44.17$, $p < .001$, *Exp(B)* = 2.03, 95% CI [.50, .92]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Conscientiousness, Extraversion, and Agreeableness. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Agreeableness were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of seven predictor variables, including Emotional Stability, Extraversion, Agreeableness, Conscientiousness, Cognitive Ability, Situational Judgment, and Integrity. The first variable included in the decision tree model was Cognitive Ability, followed by Integrity and Agreeableness (appeared simultaneously), Emotional Stability and Situational Judgment (appeared simultaneously), and Conscientiousness and Extraversion (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, eight specific cutoff points were identified, resulting in nine branches and nine terminal nodes (see Appendix A). Finally, the decision tree model consisted of five terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 4 suggested that the use of both methods could provide relatively equivalent benefits to the practice of personnel decision-making (RQ_1). Specifically, both models demonstrated high levels of classification accuracy and discriminant capacity (RQ_2) without producing adverse impact (RQ_3). Furthermore, logistic regression produced a more parsimonious model than decision tree analysis; however, the decision tree analysis identified discrete cutoff points, whereas the logistic regression did not (RQ_4).

Data set 5. Data set 5 consisted of a sample of 400 cases. Additionally, the data set was characterized by a minority distribution of 80% White/Caucasian and 20% Black/African-American individuals and a job performance cutoff point of 0*z*. Data set 5 aimed to represent a moderately sized sample that featured a minority distribution similar to that of the U.S. and an average job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 70.4% and 65.9% for training and validation samples, respectively. In comparison,

the decision tree model had a classification accuracy of 67.3% and 62.2% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .77 and .73 for training and validation samples, respectively, and the decision tree model had an AUC index of .72 and .66 for training and validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy and discriminant capacity than decision tree analysis in data set 5.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 44.4% and a minority (i.e., Black/African-American) selection rate of 26.3%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .59 and thus indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 37.2% and a minority selection rate of 33.8%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .91 and thus not indicative of adverse impact. Overall, in comparison to logistic regression in data set 5, decision tree analysis demonstrated superior performance regarding adverse impact.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 400) = 74.66, p < .001. Analyses revealed that the logistic regression model consisted of five significant predictor variables: Extraversion, B = .37, $\chi^2(1)$ = 4.76, p = .03, *Exp(B)* = 1.44, 95% CI [.04, .70]; Openness to Experience, B = -.31, $\chi^2(1)$ = 4.03, p = .05, *Exp(B)* = .73, 95% CI [-.62, -.01]; Conscientiousness, B = .38, $\chi^2(1)$ = 5.69, p = .02, *Exp(B)* = 1.46, 95% CI [.07, .69]; Cognitive Ability, B = .88, $\chi^2(1)$ = 31.76, p < .001, *Exp(B)* = 2.42, 95% CI [.58, 1.19]; and Integrity, B = .85, $\chi^2(1)$ = 22.66, p < .001, *Exp(B)* = 2.34, 95% CI [.50, 1.20]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Conscientiousness, Extraversion, and Openness to

Experience. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Openness to Experience were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of five predictor variables, including Extraversion, Conscientiousness, Cognitive Ability, Situational Judgment, and Integrity. The first variable included in the decision tree model was Integrity, followed by Conscientiousness and Situational Judgment (appeared simultaneously), and Cognitive Ability and Extraversion (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, seven specific cutoff points were identified, resulting in eight branches and eight terminal nodes (see Appendix A). Finally, the decision tree model consisted of four terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 5 suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1); however, such benefits may be limited under data conditions similar to the current data set. Specifically, logistic regression out-performed decision tree analysis in classification accuracy (RQ_2); however, decision tree analysis out-performed logistic regression in regard to adverse impact and demonstrated clear benefits to non-discriminatory hiring by selection nearly equal portions of minority and non-minority cases (RQ_3). Lastly, the two models used an equal number of predictor variables to predict job performance, but decision tree analysis provided more useful output (e.g., discrete cutoff points, visual output), which may be more interpretable to decision-makers (RQ_4).

Data set 6. Data set 6 consisted of a sample of 400 cases. Additionally, the data set was characterized by a minority distribution of 80% White/Caucasian and 20% Black/African-American individuals and a job performance cutoff point of -0.5*z*. Data set 6 aimed to represent a moderately sized sample that featured a minority distribution similar to that of the U.S. and a low job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 76.7% and 74.4% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 72.6% and 70.7% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .78 and .75 for training and validation samples, respectively, and the decision tree model had an AUC index of .73 and .66 for training and validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy and discriminant capacity than decision tree analysis in data set 6.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 87.2% and a minority (i.e., Black/African-American) selection rate of 68.8%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .79 and thus indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 89.4% and a minority selection rate of 90.0%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was 1.00 and thus not indicative of adverse impact. Rather, the decision tree model resulted in a selection rate of minorities that was approximately equal to the selection rate of non-minorities. Overall, in comparison to logistic regression in data set 6, decision tree analysis demonstrated superior performance regarding adverse impact.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 400) = 67.22, p < .001. Analyses revealed that the logistic regression model consisted of two significant predictor variables: Cognitive Ability, B = .95, $\chi^2(1)$ = 30.91, p < .001, *Exp(B)* = 2.59, 95% CI [.62, 1.29]; and Integrity, B = .83, $\chi^2(1)$ = 18.51, p < .001, *Exp(B)* = 2.30, 95% CI [.45, 1.21]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity. Overall, the logistic regression model found that increases in all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of six predictor variables, including Extraversion, Openness to Experience, Conscientiousness, Cognitive Ability, Situational Judgment, and Integrity. The first variable included in the decision tree model was Integrity, followed by Openness to Experience and Situational Judgment (appeared simultaneously), Cognitive Ability, and Extraversion and Conscientiousness (appeared simultaneously). Overall, the decision tree model also found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, seven specific cutoff points were identified, resulting in eight branches and eight terminal nodes (see Appendix A). Finally, the decision tree model consisted of seven terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 6 suggested that the use of both methods could provide relatively equivalent benefits to the practice of personnel decision-making; however, such benefits may be limited in various ways (RQ_1). Specifically, logistic regression out-performed decision tree analysis in classification accuracy (RQ_2) and decision tree analysis out-performed logistic regression in non-discriminatory hiring (RQ_3). However, the degree to which

decision tree analysis out-performed logistic regression in non-discriminatory hiring suggested that the benefits of decision tree analysis might outweigh its limitations. Finally, logistic regression produced a more parsimonious model than decision tree analysis, but decision tree analysis provided more useful output for decision-making (RQ_4).

Data set 7. Data set 7 consisted of a sample of 400 cases. Additionally, the data set was characterized by a minority distribution of 60% White/Caucasian and 40% Black/African-American individuals and a job performance cutoff point of 0*z*. Data set 7 aimed to represent a moderately sized sample that featured a minority distribution that would be expected from high minority recruitment and an average job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 65.7% and 69.5% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 66.7% and 68.3% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .74 and .75 for training and validation samples, respectively, and the decision tree model had an AUC index of .72 and .69 for training and validation samples, respectively. Overall, decision tree analysis demonstrated higher classification accuracy in the training sample of data set 7, whereas logistic regression demonstrated higher classification accuracy in the smaller, validation sample. In contrast, logistic regression demonstrated greater discriminant capacity than decision tree analysis in both samples.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 50.4% and a minority (i.e., Black/African-American) selection rate of 25.0%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .50 and thus indicative of adverse impact. In comparison, the decision tree model had

a non-minority selection rate of 24.6% and a minority selection rate of 23.8%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .97 and thus not indicative of adverse impact. Overall, in comparison to logistic regression in data set 7, decision tree analysis demonstrated superior performance regarding adverse impact.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 400) = 61.57, p < .001. Analyses revealed that the logistic regression model consisted of five significant predictor variables: Extraversion, B = .36, $\chi^2(1)$ = 4.97, p = .03, *Exp(B)* = 1.44, 95% CI [.04, .68]; Openness to Experience, B = -.32, $\chi^2(1)$ = 4.17, p = .04, *Exp(B)* = .73, 95% CI [-.63, -.01]; Conscientiousness, B = .38, $\chi^2(1)$ = 6.25, p = .01, *Exp(B)* = 1.46, 95% CI [.08, .68]; Cognitive Ability, B = .78, $\chi^2(1)$ = 31.08, p < .001, *Exp(B)* = 2.19, 95% CI [.51, 1.06]; and Integrity, B = .62, $\chi^2(1)$ = 13.24, p < .001, *Exp(B)* = 1.86, 95% CI [.29, .95]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Conscientiousness, Extraversion, and Openness to Experience. Overall, the logistic regression model found that increases in Cognitive Ability, Integrity, Extraversion, and Conscientiousness, as well as decreases in Openness to Experience were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of five predictor variables, including Extraversion, Conscientiousness, Cognitive Ability, Situational Judgment, and Integrity. The first variable included in the decision tree model was Integrity, followed by Conscientiousness and Situational Judgment (appeared simultaneously), and Extraversion and Cognitive Ability (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable

level of job performance. Additionally, five specific cutoff points were identified, resulting in six branches and six terminal nodes (see Appendix A). Finally, the decision tree model consisted of one terminal node that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 7 suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1). Specifically, decision tree analysis demonstrated the potential to provide benefits to classification accuracy by producing a slightly more accurate model than logistic regression in the training sample (RQ_2). However, such benefits were not extended to the smaller, validation sample or the discriminant capacity of the model. In contrast, decision tree analysis demonstrated clear benefits to non-discriminatory hiring (RQ_3). Finally, decision tree analysis produced an equally parsimonious model as logistic regression and identified discrete cutoff points, thus offering benefits to interpretability above and beyond logistic regression (RQ_4).

Data set 8. Data set 8 consisted of a sample of 400 cases. Additionally, the data set was characterized by a minority distribution of 60% White/Caucasian and 40% Black/African-American individuals and a job performance cutoff point of -0.5*z*. Data set 8 aimed to represent a moderately sized sample that featured a minority distribution that would be expected from high minority recruitment and a low job performance cutoff.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 75.8% and 70.1% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 74.8% and 73.2% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .77 and .71 for training and validation samples, respectively, and the decision tree model had an AUC index of .73 and .70 for training and

validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy in the training sample of data set 8, whereas decision tree analysis demonstrated higher classification accuracy in the smaller, validation sample. In contrast, logistic regression demonstrated greater discriminant capacity than decision tree analysis in both samples.

Adverse impact. Analyses revealed that the logistic regression model had a non-minority (i.e., White/Caucasian) selection rate of 89.2% and a minority (i.e., Black/African-American) selection rate of 75.0%. In turn, the four-fifths adjusted selection ratio for the logistic regression model was .84 and thus not indicative of adverse impact. In comparison, the decision tree model had a non-minority selection rate of 84.6% and a minority selection rate of 73.8%. Consequently, the four-fifths adjusted selection ratio for the decision tree model was .87 and thus not indicative of adverse impact. Overall, both models were not associated with adverse impact; however, it is noteworthy that decision tree analysis had a slightly higher adjusted selection ratio than logistic regression.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (8, N = 400) = 66.69, p < .001. Analyses revealed that the logistic regression model consisted of four significant predictor variables: Agreeableness, B = -.36, $\chi^2(1)$ = 5.01, p = .03, *Exp(B)* = .70, 95% CI [-.67, -.04]; Conscientiousness, B = .45, $\chi^2(1)$ = 7.07, p = .008, *Exp(B)* = 1.57, 95% CI [.12, .78]; Cognitive Ability, B = .90, $\chi^2(1)$ = 31.90, p < .001, *Exp(B)* = 2.46, 95% CI [.59, 1.21]; and Integrity, B = .78, $\chi^2(1)$ = 16.91, p < .001, *Exp(B)* = 2.19, 95% CI [.41, 1.16]. As shown above, logistic regression model assigned the largest odds ratio to Cognitive Ability, followed by Integrity, Conscientiousness, and Agreeableness. Overall, the logistic regression model found that increases in Conscientiousness, Cognitive Ability, and Integrity, as well as decreases in Agreeableness were associated with an increased likelihood of demonstrating an acceptable level of job performance.

In comparison, the decision tree model consisted of six predictor variables, including Emotional Stability, Extraversion, Conscientiousness, Cognitive Ability, Situational Judgment, and Integrity. The first variable included in the decision tree model was Integrity, followed by Openness to Experience and Situational Judgment (appeared simultaneously), Emotional Stability and Cognitive Ability (appeared simultaneously), and Extraversion. Overall, the decision tree model also found that higher values on all model variables were associated with an increased likelihood of demonstrating an acceptable level of job performance. Additionally, six specific cutoff points were identified, resulting in seven branches and seven terminal nodes (see Appendix A). Finally, the decision tree model consisted of five terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an acceptable level of job performance.

In sum, the results derived from data set 8 suggested that the use of both methods could provide relatively equivalent benefits to the practice of personnel decision-making (RQ_1). Specifically, although logistic regression largely out-performed decision tree analysis in classification accuracy and discriminant capacity, decision tree analysis did produce a more accurate model in the smaller, validation sample (RQ_2). Furthermore, although both methods produced models that were void of adverse impact, decision tree analysis resulted in a slightly higher four-fifths adjusted selection ratio, thus offering potentially greater benefits to non-discriminatory hiring (RQ_3). Finally, the logistic regression model utilized fewer predictor variables to predict job performance; however, decision tree analysis provided the method's characteristic output (e.g., discrete cutoff points, visual output), which may be more useful and interpretable than output provided by logistic regression.

Study 2

Correlations

Correlations (i.e., effect sizes) for the current study are shown in Table 8. Variables included in the current study consisted of constructs that demonstrate consistent relationships with subjective well-being. Each variable's relationship with subjective well-being has been studied by meta-analyses and all correlations used in Study 2 were obtained from such sources. Variables included in the current study consisted of 1) Subjective Well-Being, 2) Emotional Stability, 3) Extraversion, 4) Openness to Experience, 5) Agreeableness, 6) Conscientiousness, 7) Self-Esteem, 8) Generalized Self-Efficacy, 9) Internal Locus of Control, and 10) Job Satisfaction.

Review of meta-analysis sources identified a positive correlation between Big Five personality and Subjective Well-Being (DeNeve & Cooper, 1998). Specifically, correlations between Big Five personality and Subjective Well-Being included, Emotional Stability (.22), Extraversion (.17), Openness to Experience (.11), Agreeableness (.17), and Conscientiousness (.21). Additionally, all Big Five personality variables were found to have a positive correlation with one another (Mount et al., 2005; see Table 8). Next, a positive correlation between Self-Esteem and Subjective Well-Being was identified (.35; Judge et al., 2002). Moreover, Self-Esteem had a positive correlation with all Big Five personality variables (Judge & Bono, 2001; Judge et al., 2002; see Table 8). Generalized Self-Efficacy was also found to have a positive correlation with Subjective Well-Being (.22), as well as all Big Five personality variables and Self-Esteem (Judge & Bono, 2001; Judge et al., 2002; see Table 8). Internal Locus of Control was found to have similar positive correlations with all other study variables, including Subjective Well-Being (.17; Judge et al., 2002; see Table 8). Finally, Job Satisfaction was found

to have a positive correlation with Subjective Well-Being (.44) and all other study variables (Judge & Bono, 2001; Judge, Heller, & Mount, 2002; Tait, Padgett, & Baldwin, 1989).

Table 8. Meta-Analytic Correlations (Study 2)

Variable	1	2	3	4	5	6	7	8	9	10
1. Subjective Well-Being	1.00									
2. Emotional Stability	.22	1.00								
3. Extraversion	.17	.24	1.00							
4. Openness to Experience	.11	.19	.45	1.00						
5. Agreeableness	.17	.42	.26	.17	1.00					
6. Conscientiousness	.21	.52	.17	.09	.39	1.00				
7. Self-Esteem	.35	.66	.36	.14	.22	.39	1.00			
8. Generalized Self-Efficacy	.22	.59	.39	.33	.23	.43	.85	1.00		
9. Internal Locus of Control	.17	.51	.26	.24	.19	.31	.59	.63	1.00	
10. Job Satisfaction	.44	.29	.25	.02	.17	.26	.26	.45	.32	1.00

Data Generation and Assessment

Following Monte Carlo simulation, the final data set's correlation matrix was assessed for discrepancies in comparison to the originally imputed correlation matrix. Although Study 2 consists of three total data sets, the specification of a subjective well-being cutoff point did not influence correlations between other study variables. As a result, only the single, final data set was assessed for deviation from the population parameters. Overall, the final data set demonstrated minimal discrepancies in comparison to the originally imputed correlation matrix. Specifically, the data set used in the current study had a mean discrepancy of -.004 and an absolute value mean discrepancy score of -.015.

Test of Research Questions

Following the creation of a binary logistic regression and decision tree analysis solution, each model's validity and interpretability was assessed. Furthermore, validity was assessed

using both a training and validation sample, in order to evaluate model stability. Validity was demonstrated via classification accuracy, or the total percentage of correct classifications derived from the model, and area under the ROC curve (AUC), exhibited as a value between 0 and 1. Higher values on both indicators of validity represented a more accurate model overall (see Table 9). In comparison, interpretability was assessed via an explanation of each model's structure, included variables, and clarity of cutoff points. Odds ratios were used as an indicator of variable importance for logistic regression models and the order in which variables appeared in the decision tree was used as an indicator of variable importance for decision tree models. Validity indicators for models within each data set are shown in Table 9 and complete decision tree models can be found in Appendix A.

Table 9. Model Classification Accuracy and AUC Index (Study 2)

Data Set	Model	Classification Accuracy		AUC Index	
		Training	Validation	Training	Validation
1	Binary Logistic Regression	74.7%	72.9%	.84	.81
	Decision Tree Analysis	67.1%	65.7%	.72	.70
2	Binary Logistic Regression	79.1%	76.8%	.84	.81
	Decision Tree Analysis	71.6%	70.7%	.71	.69
3	Binary Logistic Regression	79.2%	77.9%	.84	.83
	Decision Tree Analysis	73.3%	71.7%	.70	.69

Notes.
1: N = 5000, 0z;
2: N = 5000, +0.5z;
3: N = 5000, -0.5z.

Data set 1. Data set 1 consisted of a sample of 5000 cases. Additionally, the data set was characterized by a subjective well-being cutoff point of 0z. Data set 1 aimed to represent a relatively large sample, which may be similar in size to the employee population of a moderate to large organization, and an average level of subjective well-being.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 74.7% and 72.9% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 67.1% and 65.7% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .84 and .81 for training and validation samples, respectively, and the decision tree model had an AUC index of .72 and .70 for training and validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy and discriminant capacity than decision tree analysis in data set 1 of Study 2.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (9, N = 5000) = 1580.42, $p < .001$. Analyses revealed that the logistic regression model consisted of nine significant predictor variables: Emotional Stability, B = -.68, $\chi^2(1)$ = 115.93, $p < .001$, *Exp(B)* = .51, 95% CI [-.81, -.56]; Extraversion, B = -.47, $\chi^2(1)$ = 93.80, $p < .001$, *Exp(B)* = .62, 95% CI [-.57, -.38]; Openness to Experience, B = .91, $\chi^2(1)$ = 275.27, $p < .001$, *Exp(B)* = 2.49, 95% CI [.81, 1.02]; Agreeableness, B = .24, $\chi^2(1)$ = 27.14, $p < .001$, *Exp(B)* = 1.27, 95% CI [.15, .33]; Conscientiousness, B = .38, $\chi^2(1)$ = 59.25, $p < .001$, *Exp(B)* = 1.46, 95% CI [.28, .48]; Self-Esteem, B = 3.17, $\chi^2(1)$ = 683.32, $p < .001$, *Exp(B)* = 23.85, 95% CI [2.93, 3.41]; Generalized Self-Efficacy, B = -2.76, $\chi^2(1)$ = 542.26, $p < .001$, *Exp(B)* = .06, 95% CI [-3.00, -2.53]; Internal Locus of Control, B = -.22, $\chi^2(1)$ = 17.52, $p < .001$, *Exp(B)* = .81, 95% CI [-.32, -.12]; and Job Satisfaction, B = 1.72, $\chi^2(1)$ = 701.98, $p < .001$, *Exp(B)* = 5.56, 95% CI [1.59, 1.84]. As shown above, logistic regression model assigned the largest odds ratio to Job Satisfaction, followed by Self-Esteem, Generalized Self-Efficacy, Openness to Experience, Emotional Stability, Extraversion, Conscientiousness, Agreeableness, and Internal Locus of Control. Overall, the logistic regression model found that increases in Job

Satisfaction, Self-Esteem, Openness to Experience, Conscientiousness, and Agreeableness, as well as decreases in Generalized Self-Efficacy, Emotional Stability, Extraversion, and Internal Locus of Control were associated with an increased likelihood of demonstrating an average level of Subjective Well-Being. Furthermore, the logistic regression model identified no specific cutoff points in any analyses.

In comparison, the decision tree model consisted of four predictor variables, including Agreeableness, Conscientiousness, Self-Esteem, and Job Satisfaction. The first variable included in the decision tree model was Job Satisfaction, followed by Self-Esteem, and Agreeableness and Conscientiousness (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an average level of Subjective Well-Being, with the exception of Self-Esteem, which demonstrated a curvilinear relationship with Subjective Well-Being. Specifically, although higher values of Self-Esteem were originally associated with an increased likelihood of demonstrating an average level of Subjective Well-Being, a point of diminishing returns was reached, at which point a higher value of Self-Esteem was associated with a decreased likelihood of demonstrating an average level of the outcome. Additionally, seven specific cutoff points were identified, resulting in eight branches and eight terminal nodes (see Appendix A). Finally, the decision tree model consisted of five terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an average level of Subjective Well-Being. This is indicative of multicollinearity in the regression model. As multicollinearity cannot occur in DTA models, we see two differing results across methods with DTA providing more accurate and interpretable results.

In sum, the results derived from data set 1 in Study 2 provided key insight into the current study's research questions and, in general, suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1). However, the current

study's results suggested fewer benefits when predicting employee health (e.g., subjective well-being than when predicting performance. Furthermore, adverse impact was not calculated for the models in Study 2, thus limiting the potential benefits of decision tree analysis. Nevertheless, decision tree analysis produced models with relatively high levels of classification accuracy and discriminant capacity; however, logistic regression clearly out-performed decision tree analysis in data set 1 (RQ_2). In contrast, decision tree analysis produced a more parsimonious model when predicting employee health and provided highly interpretable output (RQ_4).

Data set 2. Data set 2 consisted of a sample of 5000 cases. Additionally, the data set was characterized by a subjective well-being cutoff point of +0.5*z*. Data set 2 aimed to represent a relatively large sample and an above-average level of subjective well-being.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 79.1% and 76.8% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 71.6% and 70.7% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .84 and .81 for training and validation samples, respectively, and the decision tree model had an AUC index of .71 and .69 for training and validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy and discriminant capacity than decision tree analysis in data set 2 of Study 2.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (9, N = 5000) = 1433.61, $p < .001$. Analyses revealed that the logistic regression model consisted of nine significant predictor variables: Emotional Stability, B = -.61, $\chi^2(1)$ = 81.57, $p < .001$, *Exp(B)* = .54, 95% CI [-.75, -.48]; Extraversion, B = -.55, $\chi^2(1)$ = 106.50, $p < .001$, *Exp(B)* = .58, 95% CI [-.66, -.45]; Openness to Experience, B = 1.02, $\chi^2(1)$ = 290.21, $p < .001$, *Exp(B)* = 2.77, 95% CI [.90, 1.14]; Agreeableness, B = .16, $\chi^2(1)$

= 10.22, p = .001, *Exp(B)* = 1.17, 95% CI [.06, .26]; Conscientiousness, B = .39, $\chi^2(1)$ = 52.16, p < .001, *Exp(B)* = 1.47, 95% CI [.28, .49]; Self-Esteem, B = 3.29, $\chi^2(1)$ = 642.33, p < .001, *Exp(B)* = 26.87, 95% CI [3.04, 3.55]; Generalized Self-Efficacy, B = -2.92, $\chi^2(1)$ = 517.74, p < .001, *Exp(B)* = .05, 95% CI [-3.17, -2.67]; Internal Locus of Control, B = -.19, $\chi^2(1)$ = 11.89, p < .001, *Exp(B)* = .83, 95% CI [-.30, -.08]; and Job Satisfaction, B = 1.72, $\chi^2(1)$ = 628.52, p < .001, *Exp(B)* = 5.59, 95% CI [1.59, 1.85]. The logistic regression model assigned the largest odds ratio to Self-Esteem, followed by Job Satisfaction, Generalized Self-Efficacy, Openness to Experience, Extraversion, Emotional Stability, Conscientiousness, Internal Locus of Control, and Agreeableness. Overall, the logistic regression model found that increases in Job Satisfaction, Self-Esteem, Openness to Experience, Conscientiousness, and Agreeableness, as well as decreases in Generalized Self-Efficacy, Emotional Stability, Extraversion, and Internal Locus of Control were associated with an increased likelihood of demonstrating an above-average level of Subjective Well-Being.

In comparison, the decision tree model consisted of five predictor variables, including Emotional Stability, Agreeableness, Conscientiousness, Self-Esteem, and Job Satisfaction. The first variable included in the decision tree model was Job Satisfaction, followed by Self-Esteem and Emotional Stability (appeared simultaneously), and Agreeableness and Conscientiousness (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating an above-average level of Subjective Well-Being. Additionally, seven specific cutoff points were identified, resulting in eight branches and eight terminal nodes (see Appendix A). Finally, the decision tree model consisted of three terminal nodes that consisted of more than 50% of individuals predicted to demonstrate an above-average level of Subjective Well-Being.

In sum, the results derived from data set 2 in Study 2 suggested that the use of decision tree analysis could provide benefits to the practice of personnel decision-making (RQ_1), although such benefits are limited. Specifically, when modeling employee health, the majority of benefits provided by decision tree analysis are associated with the output produced by the analysis. That is, although logistic regression outperformed decision tree analysis in classification accuracy (RQ_2), decision tree analysis still produced a more parsimonious model and provided useful and interpretable output (RQ_4). Consequently, decision-makers may find an advantage to using decision tree analysis instead of logistic regression.

Data set 3. Data set 3 consisted of a sample of 5000 cases. Additionally, the data set was characterized by a subjective well-being cutoff point of -0.5*z*. Data set 3 aimed to represent a relatively large sample and a below-average level of subjective well-being.

Validity. Analyses revealed that the logistic regression model had a classification accuracy of 79.2% and 77.9% for training and validation samples, respectively. In comparison, the decision tree model had a classification accuracy of 73.3% and 71.7% for training and validation samples, respectively. Additionally, ROC curve analysis revealed that the logistic regression model had an AUC index of .84 and .83 for training and validation samples, respectively, and the decision tree model had an AUC index of .70 and .69 for training and validation samples, respectively. Overall, logistic regression demonstrated higher classification accuracy and discriminant capacity than decision tree analysis in data set 3 of Study 2.

Interpretability. A test of the full logistic regression model versus a model with intercept only was statistically significant, χ^2 (9, N = 5000) = 1336.82, $p < .001$. Analyses revealed that the logistic regression model consisted of nine significant predictor variables: Emotional Stability, B = -.57, $\chi^2(1) = 73.42$, $p < .001$, *Exp(B)* = .57, 95% CI [-.70, -.44]; Extraversion, B = -.45, $\chi^2(1) = 73.26$, $p < .001$, *Exp(B)* = .64, 95% CI [-.55, -.35]; Openness to Experience, B =

1.03, $\chi^2(1) = 293.10$, $p < .001$, *Exp(B)* = 2.81, 95% CI [.91, 1.15]; Agreeableness, $B = .21$, $\chi^2(1) = 16.87$, $p < .001$, *Exp(B)* = 1.23, 95% CI [.11, .30]; Conscientiousness, $B = .32$, $\chi^2(1) = 36.74$, $p < .001$, *Exp(B)* = 1.37, 95% CI [.21, .42]; Self-Esteem, $B = 3.01$, $\chi^2(1) = 569.76$, $p < .001$, *Exp(B)* = 20.35, 95% CI [2.77, 3.26]; Generalized Self-Efficacy, $B = -2.74$, $\chi^2(1) = 479.23$, $p < .001$, *Exp(B)* = .07, 95% CI [-2.99, -2.50]; Internal Locus of Control, $B = -.17$, $\chi^2(1) = 9.47$, $p = .002$, *Exp(B)* = .84, 95% CI [-.28, -.06]; and Job Satisfaction, $B = 1.71$, $\chi^2(1) = 624.11$, $p < .001$, *Exp(B)* = 5.52, 95% CI [1.58, 1.84]. The logistic regression model assigned the largest odds ratio to Job Satisfaction, followed by Self-Esteem, Generalized Self-Efficacy, Openness to Experience, Emotional Stability, Extraversion, Conscientiousness, Agreeableness, and Internal Locus of Control. Overall, the logistic regression model found that increases in Job Satisfaction, Self-Esteem, Openness to Experience, Conscientiousness, and Agreeableness, as well as decreases in Generalized Self-Efficacy, Emotional Stability, Extraversion, and Internal Locus of Control were associated with an increased likelihood of demonstrating a below-average level of Subjective Well-Being.

In comparison, the decision tree model consisted of six predictor variables, including Openness to Experience, Conscientiousness, Self-Esteem, Generalized Self-Efficacy, Internal Locus of Control, and Job Satisfaction. The first variable included in the decision tree model was Generalized Self-Efficacy, followed by Job Satisfaction, and Openness to Experience, Conscientiousness, Self-Esteem, and Internal Locus of Control (appeared simultaneously). Overall, the decision tree model found that higher values on all model variables were associated with an increased likelihood of demonstrating a below-average level of Subjective Well-Being. Additionally, nine specific cutoff points were identified, resulting in ten branches and ten terminal nodes (see Appendix A). Finally, the decision tree model consisted of seven terminal

nodes that consisted of more than 50% of individuals predicted to demonstrate a below-average level of Subjective Well-Being.

In sum, the results derived from data set 3 in Study 2 suggested similarly limited benefits, as were specified by Data Sets 1 and 2 (RQ_1). Specifically, when modeling employee health, decision tree analysis failed to produce a model with greater classification accuracy than models produced by logistic regression (RQ_3). However, in data set 3, decision tree analysis offered some benefit by producing a more parsimonious and interpretable model than logistic regression, featuring discrete cutoff points and visual output (RQ_4).

CHAPTER V: DISCUSSION

In today's highly competitive global economy, the strategic management of talented employees is of vital importance to organizational success (Vaiman et al., 2012). Furthermore, organizational decision-makers who recognize that employees play a central role in top performing companies, now dedicate significant portions of their time to personnel decision-making (Scullion et al., 2010). Unfortunately, although there is increasing emphasis on personnel decision-making in practice, academic research on the topic has been slow to develop (Scullion et al., 2010). Moreover, many of the individuals who desire to practice strategic personnel decision-making, lack expertise in the area (Vaiman et al., 2012). As a result, it is critical that research be aimed at the ways in which organizations can make more accurate and useful personnel decisions.

One field of study that is particularly well suited to address such a topic is the field of Occupational Health Psychology (OHP). OHP is an emerging interdisciplinary specialty within the broader field of psychology that focuses on organizational improvements that benefit both the organization and its employees (Raymond, Wood, & Patrick, 1990). Moreover, research within OHP focuses on multifaceted goals, such as improving organizational productivity, while also enhancing the quality of life and well-being of employees (Quick, Tetrick, Adkins, & Klunder, 2003). It is through this framework that the topic of personnel decision-making is addressed in the current study. Specifically, both employee performance within the context of personnel selection, and employee health (i.e., subjective well-being) were examined.

In particular, the use of a novel statistical technique for decision-making—decision tree analysis—was examined in the current study. Decision tree analysis uses binary recursive partitioning to construct multi-pathway predictive models (Lewis, 2000). In contrast to traditional statistical techniques such as multiple linear regression and binary logistic regression,

decision tree analysis is non-parametric and does not carry statistical assumptions that limit the usefulness of other techniques. As a result, it is used within the fields of statistics, medical diagnosis, and fraud detection (Lavanya & Rani, 2011) and favored by researchers who use the analysis to produce accurate and interpretable statistical solutions despite challenging data characteristics (Karels et al., 2004; Lewis, 2000).

The use of decision tree analysis was examined in the current study, in order to investigate the analysis' ability to play a role in accomplishing the goals of improving employee performance and health. Such an examination was conducted in the current study by constructing a binary logistic regression and decision tree analysis solution for a variety of data sets on employee performance and health, and comparing such solutions on the basis of validity, adverse impact, and interpretability. Overall, results sought to inform organizations on best practices when making personnel decisions not only at the onset of employment, but also throughout the employee life-cycle, to ensure long-term satisfaction and retention of talented individuals.

Discussion of Results and Research Questions

Following the creation of a logistic regression and decision tree analysis solution for each data set in the current study, each model's validity, adverse impact, and interpretability were assessed. Overall, there were a total of 11 data sets included in the current study. Specifically, eight data sets were included within Study 1, which addressed the topic of employee performance, and three data sets were included within Study 2, which addressed the topic of employee health (i.e., subjective well-being). Due to its relevance within the context of personnel selection, Study 1 data sets included information regarding the demographic characteristics (i.e., race) of each case. Consequently, adverse impact was calculated for each statistical solution in Study 1 and model comparisons were made. In contrast, demographic

information was not included in Study 2 data sets and such model comparisons were not incorporated. As a result, Research Question 3 only includes information derived from Study 1 results; however, all other research questions (RQs) include information from both Study 1 and Study 2 and results are discussed collectively.

Research Question 1

Research Question 1 asked, "In comparison to traditional personnel decision-making methods (i.e., binary logistic regression), can the use of decision tree analysis provide benefits to the field?" RQ1 was answered in the current study by investigating the validity, adverse impact, and interpretability of logistic regression and decision tree analysis solutions for a variety of data conditions. In regard to validity, both statistical techniques produced models with relatively high classification accuracy. For example, in Study 1, the highest level of classification accuracy across data sets for logistic regression and decision tree analysis solutions was 76.7% and 74.8%, respectively. Furthermore, within each data set, differences in validity between the two techniques ranged from 0.1% to 4.1%, with both logistic regression and decision tree analysis producing the highest level of validity in some cases. In Study 2, the highest level of classification accuracy across data sets for logistic regression and decision tree analysis solutions was 79.2% and 73.3%, respectively. Within each data set in Study 2, differences in validity between the two techniques ranged from 5.9% to 7.6%. Overall, both statistical techniques performed well in regard to producing accurate predictive models and support was found for the use of decision tree analysis within OHP and other related fields.

Regarding adverse impact, the two statistical techniques diverged in their ability to produce a statistical solution that was non-discriminatory. For example, in Study 1, the use of logistic regression resulted in a model that contained adverse impact in six out of eight data sets. In contrast, the use of decision tree analysis resulted in a model that contained adverse impact in

only two out of eight data sets. Furthermore, within the two data sets that the decision tree analysis produced models with adverse impact, the use of logistic regression also resulted in discriminatory models. In sum, the current study resulted in strong evidence that the use of decision tree analysis is associated with the development of both accurate and non-discriminatory predictive models.

Lastly, in regard to interpretability, the two models also demonstrated notable differences. For example, across all 11 data sets included in both studies, decision tree analysis produced a more parsimonious model, containing fewer predictor variables than logistic regression models, in six data sets. Moreover, the two techniques produced models with an equal number of variables in two data sets. Additionally, decision tree analysis almost exclusively produced splits that suggested that higher levels of predictor variables were associated with an increased likelihood of being classified in the target class. The only exception to this result was found in Study 2, which consisted of a single curvilinear relationship characterized by a certain point at which higher values of a predictor variable were no longer associated with an increased likelihood of being classified in the target class. However, this demonstrates a potential strength of decision tree analysis, rather than a weakness. In contrast, logistic regression models frequently included beta-weights that suggested lower levels of some predictor variables were associated with an increased likelihood of being classified in the target class. As all variables included in the current study have a positive correlation with their respective study's outcome, such results produced from logistic regression are counterintuitive and conflict with meta-analytic findings. Finally, decision tree analysis produced specific cutoff points and visual output for all data sets, whereas logistic regression did not. In conclusion, decision tree analysis demonstrated results that encourage its use. As a result, RQ1 is affirmative; the use of decision tree analysis can provide benefits to the field (see Table 10).

Table 10. Research Question and Results Summary

RQ	Key Results
RQ_1	• DTA produced models with high levels of classification accuracy and, in some cases, greater accuracy than LR models (Study 1 Max = 74.8%; Study 2 Max = 73.3%) • DTA produced models with little or no adverse impact across a wide range of data conditions • All DTA models consisted of discrete cutoff points and visual output; thus, increasing each model's interpretability
RQ_2	• DTA produced models with nearly equivalent levels of classification accuracy as LR and, under certain conditions, produced models with greater accuracy (2/11) or accuracy within 1.0% of LR models (2/11) • Similar levels of classification accuracy were observed across both training and validation samples, demonstrating the ability for DTA results to be generalized to future samples • Stronger results were demonstrated when predicting employee performance than when predicting employee health
RQ_3	• The use of DTA resulted in adverse impact in only two of the eight data sets in Study 1 • Under no circumstances did LR ever produce a model with a higher four-fifths adjusted selection ratio than DTA • DTA demonstrated the ability to produce non-discriminatory models across a wide range of data conditions
RQ_4	• All DTA models consisted of discrete cutoff points, visual output, and multiple pathways to employee performance and employee health • The majority of models produced by DTA were more parsimonious (i.e., utilized fewer predictor variables) than LR; thus, suggesting the potential for greater utility • DTA demonstrated the potential to identify curvilinear relationships without additional effort and through unique abilities such as 3-way splits

Notes.
RQ = Research Question;
DTA = Decision Tree Analysis;
LR = Logistic Regression.

Research Question 2

Next, Research Question 2 asked, "In comparison to binary logistic regression, what potential benefits are provided by decision tree analysis in regard to classification accuracy when modeling key organizational outcomes (i.e., employee performance and health)?" Several interesting findings were derived from the current study regarding the specific benefits of decision tree analysis. First, it is important to note that both statistical techniques produced models with similar levels of classification accuracy. For example, in Study 1, logistic

regression and decision tree analysis produced models with classification accuracy within 1.0% of one another in two of the eight data sets. Moreover, both models performed well across all data conditions and in both training and validation samples. This finding suggests that not only can decision tree analysis produce models with similar levels of accuracy as more commonly used regression techniques, but it can also do so in smaller validation samples thereby demonstrating the generalizability of its solutions.

One area in which decision tree analysis was out-performed by logistic regression was in producing models that had high levels of area under the ROC curve (AUC). Specifically, with only one exception, logistic regression resulted in higher levels of AUC than decision tree analysis across all data conditions. This result indicates that although decision tree analysis was capable of achieving equal levels of classification accuracy, such accuracy was at the cost of precise discrimination. In particular, AUC represents a model's ability to correctly identify both positive and negative cases. Given the high levels of classification accuracy, but lower levels of AUC, it is suspected that decision tree analysis classified some cases better than others. Taken together with results that revealed little adverse impact in decision tree analysis models, it is likely that, in the current study, decision tree analysis was generally more inclusive and correctly classified more positive cases than negative cases. Although in some cases it may be more beneficial to correctly classify negative cases (e.g., medical diagnosis), within the field of OHP, inclusive models that ensure the identification of positive target cases (e.g., successful and healthy employees) are extremely important.

In sum, the current study's results consisted of several key insights regarding Research Question 2 (see Table 10). First, decision tree analysis produced models that performed as well as models produced by logistic regression. This finding establishes the groundwork for the use of decision tree analysis in the field of OHP and exposes the topic of personnel decision-making

to benefits, above and beyond validity, which may be provided by decision tree analysis. Next, decision tree analysis demonstrated the ability to produce accurate classification models using small validation samples. Not only does this result suggest that decision tree analysis solutions are generalizable, but also that it can be used in organizational settings, which may consist of smaller groups of employees. Finally, although decision tree analysis produced models with slightly less AUC than models produced by logistic regression, it is suspected that it favored the correct classification of positive target cases, which can be useful in organizational settings where the cost of not including a successful or healthy employee may be greater than not including an unsuccessful or unhealthy employee.

Research Question 3

Research Question 3 asked, "In comparison to binary logistic regression, what potential benefits are provided by decision tree analysis in regard to adverse impact when modeling key organizational outcomes (i.e., employee performance and health)?" Several strong results were derived from the current study concerning benefits associated with the use of decision tree analysis in regard to adverse impact. Overall, decision tree analysis demonstrated the ability to model employee performance with little or no adverse impact. Specifically, in Study 1, the use of decision tree analysis resulted in adverse impact in only two of the eight data sets, whereas the use of logistic regression resulted in adverse impact in six data sets. Moreover, within the two data sets for which decision tree analysis resulted in adverse impact, logistic regression also produced a discriminatory model; however, the four-fifths adjusted selection ratio was higher (i.e., less discriminatory) in the decision tree analysis model. In particular, these two data sets consisted of 1200 cases and a higher performance cutoff point (i.e., 0*z*), which may have resulted in a large number of successful non-minority cases that could be selected.

Next, decision tree analysis demonstrated the ability to produce non-discriminatory models of employee performance in a variety of data conditions. For example, decision tree analysis produced models without adverse impact in data sets that included a large sample (i.e., 1200 cases), 80%/20% non-minority/minority distribution, and lower performance cutoff point (i.e., -0.5*z*), as well as data sets that included a small sample (i.e., 400 cases), 60%/40% non-minority/minority distribution, and higher performance cutoff point (i.e., 0*z*). In contrast, logistic regression was only able to produce non-discriminatory models in data conditions that were most supportive of high minority selection (e.g., 60%/40% non-minority/minority distribution and lower performance cutoff point).

Lastly, within select conditions, decision tree analysis produced models that resulted in identical selection ratios for minority and non-minority cases. This was a capability that was not demonstrated by logistic regression, which at its best produced a model that resulted in a four-fifths adjusted selection ratio of .84. In contrast, decision tree analysis resulted in four-fifths adjusted selection ratios of 1.00 and .97 for data conditions that included different minority distributions and performance cutoff points. This result suggests that decision tree analysis models are truly compensatory and facilitate certain employees to be selected at equal rates as others, despite standard deviation group differences that exist on multiple selection variables. Moreover, this result demonstrates the ability for decision tree analysis to provide benefits that would not be possible with traditional non-compensatory selection techniques such as multiple cutoffs or multiple hurdle systems. In sum, decision tree analysis demonstrated several strong benefits in regard to adverse impact in employee selection (see Table 10). In particular, it demonstrated the ability to, (1) produce non-discriminatory models of employee performance when logistic regression is unable to produce such models, (2) produce non-discriminatory models across a variety of data conditions including differences in sample size, minority

distribution, and performance cutoff point, and (3) produce models that result in equal selection rates for both minority and non-minority employees.

Research Question 4

Finally, Research Question 4 asked, "In comparison to binary logistic regression, what potential benefits are provided by decision tree analysis in regard to interpretability when modeling key organizational outcomes (i.e., employee performance and health)?" Results that suggest that decision tree analysis is associated with a variety of benefits regarding the interpretability of predictive models were derived from the current study. First, all models produced by decision tree analysis consisted of discrete cutoff points from which employees can be classified into key target groups. Furthermore, in the current study, such cutoff points were characterized by a high degree of precision, as each value was specified to the forth-decimal place. This degree of precision is beneficial to organizations that must clearly classify all individuals in one of two groups at each split point in the decision tree. In contrast, logistic regression models did not include cutoff points. Rather, only final classification decisions were made based on the regression equation derived from the analysis. As a result, logistic regression models provided less clarity in regard to the important variables, interactions, and points within the model that most influenced how each case was eventually classified.

In addition to the presence of discrete cutoff points, all decision tree analysis models were also characterized by visual output that depicted the statistical solution (see Appendix A). Such output is critically important to organizations that may lack expertise in interpreting statistical models and using such models to make important personnel decisions. Through the use of visual output, organizations can trace each cutoff point and follow the branches illustrated in the decision tree model to classify a current or future employee. Additionally, visual output can be used to share statistical results with wider audiences within the organization and foster a

deeper understanding of key employee insights. For example, visual output from data set 1 in Study 2 can be used to clearly identify a potential curvilinear relationship between a predictor variable and the outcome. Specifically, within one portion of the decision tree, Self-Esteem is split three consecutive times. Moreover, closer evaluation of each split reveals that at each *z* score there is a different portion of individuals classified in the target class. In particular, at the lowest cutoff point ($z < -.48$), 20.5% of individuals in the training data set are classified in the target class (i.e., average Subjective Well-being). This portion of individuals then increases to 39.1% and 62.2% at the next two ranges specified by the cutoff points (i.e., $z = -.48 – 1.07$, $z = 1.08 – 1.38$). However, at the final cutoff point ($z >= 1.39$) the portion of individuals classified in the target class drops to 50.0%, demonstrating a point of diminishing returns and a curvilinear relationship between Self-Esteem and Subjective Well-Being. This relationship can also be demonstrated by permitting the decision tree to produce a single, three-way split on Self-Esteem (see Appendix A). Moreover, although two-way splits were utilized in the current study to maintain simplicity for demonstrative purposes, if curvilinear relationships are expected to be present within a data set, the use of such multi-way splits can be useful and provide even greater interpretability.

Additionally, although the curvilinear relationship was only demonstrated in the training sample of data set 1 in Study 2 and not replicated in the smaller validation sample, the result demonstrates the potential for decision tree analysis to provide organizations with such insights. That is, regardless of the strength of the result in the current example, decision tree analysis demonstrated a unique ability to convey complex findings that may be otherwise overlooked without additional effort that may surpass the expertise of organizational decision-makers (e.g., testing polynomial regression equations). Furthermore, it is critical that non-linear relationships be incorporated into predictive models to reduce the presence of error and that such relationships

are presented in a way that can be easily understood by decision-makers, such as via visual output.

In contrast, logistic regression is not associated with visual output, but rather provides organizations with beta-weights that describe each variable's relationship with the outcome. Although such information can be informative, it can also be easily misused or misinterpreted by organizations that lack the necessary competence to understand model statistics. As a result, organizations may make costly errors in personnel decision-making. Furthermore, OHP practitioners may experience difficulties when sharing regression model results with organizational leadership and fostering support for the use of statistical models, despite having the necessary level of expertise to understand statistical models.

Third, decision tree analysis models utilized, in general, fewer predictor variables to model each outcome. Moreover, not only did decision tree analysis models consist of fewer variables, but they also produced nearly equal levels of classification accuracy as less parsimonious models. This is critically important for a variety of reasons. For example, models that utilize fewer variables may be more easily explained to organizational leadership and are overall less cumbersome to incorporate into practice. Next, perhaps most importantly, models that consist of fewer predictor variables are potentially less expensive to implement. As a result, decision tree analysis models are likely to result in higher levels of utility, or practical and financial value to an organization (Schmidt & Hunter, 1998). In sum, decision tree analysis models demonstrated benefits to the practice of personnel decision-making by, (1) specifying precise cutoff points from which personnel decision can be made, (2) providing visual output that facilitates more efficient and effective comprehension and distribution of results, and (3) producing highly accurate models that consist of fewer variables than models produced by other methods, thereby leading to higher levels of overall utility (see Table 10).

Study Implications

Over the last half of a century, dramatic changes have taken place in the world of work (Sparks, Faragher, & Cooper, 2001). For example, in the 1970s, organizations were faced with emerging technologies such as the use of computers. Later, in the 1980s and 1990s, organizations adapted to the globalization of business and an increasingly competitive economy. Currently, organizations must manage a variety of issues, such as the popularity of teamwork, flexible work schedules, and employee autonomy (Cox, Griffiths, & Rial-Gonzalez, 2000). Furthermore, such issues must now be balanced in the context of organizational goals that include both organizational effectiveness and employee well-being (Sparks et al., 2001). In turn, it is imperative that the field of OHP provides organizations with the necessary guidance required to successfully manage the workforce of today.

Personnel decision-making is a key area in which OHP can provide such guidance. Strategic decision-making is critical at all stages of personnel management, including the selection, development, and retention of talented employees (Vaiman et al., 2012). Furthermore, it is likely that decisions made at each of the aforementioned stages have a significant impact on the long-term productivity and health of employees. Fortunately, the study of OHP not only emphasizes the expertise required to practice behavioral science (e.g., knowledge of statistics), it also shares similar goals with current organizations, such as reducing job-related illness and increasing productivity through continuous improvement of the workplace (Sauter, Hurrel, Fox, Tetrick, & Barling, 1999; Quick et al., 2003). These assets are unique to OHP and allow the field to provide guidance that cannot be provided by other disciplines that lack a similar multidimensional focus, such as occupational safety, clinical psychology, and organizational behavior (Sauter et al., 2003). As a result, the intent of the current study was to make advancements in the field of OHP that could be used to advise decision-makers and improve

organizational success. Subsequently, the study results consisted of a variety of important findings that are associated with a host of implications.

First, results demonstrated that decision tree analysis is capable of providing benefits in regard to validity and adverse impact. Specifically, in comparison to logistic regression models, decision tree analysis models were characterized by nearly identical levels of validity across a variety of data conditions, as well as little to no adverse impact. These findings suggest that by using a procedure that is less rigid in its selection and treatment of variables, organizations have the ability to generate predictive models that both maintain validity and reduce the probability of adverse impact. Furthermore, such results indicate that decision tree analysis is capable of identifying the compensatory nature of certain predictor variables, which allow individuals to utilize strengths in some traits to achieve a particular outcome, despite weaknesses in other traits. Additionally, in both studies, decision tree analysis consistently utilized a unique composition of variables to produce similarly accurate models as those produced by logistic regression. This result suggests that decision tree analysis takes advantage of compensatory variables when modeling both employee performance and health. Moreover, such findings also suggest that, in general, employee performance and health are outcomes that can be achieved through multiple means.

Next, the current research consisted of a pair of studies, which addressed two separate, key organizational topics. Additionally, both studies included the use of training and validation samples. Although this is intrinsically a strength of the current study, taken together with the results, it also provides a critical implication for practice. Specifically, results from the current study suggest that decision tree analysis models are generalizable to smaller, future data samples and capable of addressing multiple organizational topics. In turn, organizations can utilize it at

all stages of personnel management and use the results to make accurate decisions regarding current and prospective employees.

Lastly, the benefits provided by decision tree analysis are associated with broader implications regarding the use of science in organizations. Specifically, through characteristics such as visual output and the specification of precise cutoff points, decision tree analysis potentially makes organizational science more accessible to organizations that lack expertise in statistics and personnel decision-making. Furthermore, the use of decision tree analysis may connect organizations with practitioners who can help guide organizations through the analysis and decision-making process. Consequently, it is possible that by improving the process by which organizations analyze, interpret, and share data that supports organizational decisions, organizational leaders will better understand their workplace and be capable of facilitating processes that improve the work and life of employees.

Study Limitations and Future Research

Although the current study is characterized by a variety of strengths, it is also associated with a set of limitations. In turn, the results must be considered within the context of these confines. One such limitation is the use of Monte Carlo data simulation in a limited capacity. Specifically, in the current study, Monte Carlo data simulation was utilized to generate all data sets included within Study 1 and Study 2. A total of 11 data sets were included across both studies. Although this methodology was sufficient to investigate the research questions, full Monte Carlo analyses would use a differing methodology. In particular, Monte Carlo data simulation is typically characterized by the generation of a very large number of samples, usually consisting of hundreds or thousands of data sets (Enders, 2010). Following the generation of simulated data samples, analyses are conducted on all samples in order to estimate the effect of a statistical model on a population with a particular set of parameters (Enders, 2010).

Unfortunately, this methodology could be extremely challenging in the context of decision tree analysis. Specifically, although other analyses may be applied to simulated data sets repeatedly with little to no user interaction, decision tree analysis requires a moderate amount of user interaction. In particular, decision tree analysis models gain much of their value through the interactive process of growing and pruning model branches, which would be lost in a traditional Monte Carlo simulation study. As a result, the focus of the current study was to generate a limited number of samples, which were as representative of a population as possible, and provide an initial overview of the benefits of decision tree analysis. Nevertheless, although the complexities of a traditional Monte Carlo simulation study were beyond the demonstrative purposes of the current study, future researchers should test the usefulness of decision tree analysis within that context.

Next, the current study consisted of variables that had well-established relationships with the targeted outcome and known meta-analytical effect sizes. However, in practice, organizations may not have access to data on such variables. In contrast, organizations may be required to model key outcomes using predictor variables that are merely available to the organization, regardless of their theoretical relationship with the outcome. As a result, it is useful to investigate the benefits of decision tree analysis when using a variety of predictor variables, which may or may not be correlated with the outcome. Unfortunately, the current study is limited by the fact that simulated data were used, requiring that meta-analytic effect sizes be available. In turn, only included variables with well-known relationships with the outcome were used in the current study. Future researchers should investigate the benefits of decision tree analysis using organizational data collected for the purpose of the study. Furthermore, such data should include only variables that are currently available to the

organization and likely to be used outside of the context of the study, in order to gain a more representative understanding of the benefits of decision tree analysis in practice.

Lastly, several methodological decisions were made in the current study that facilitated a level of simplicity that was important to maintain. However, such decisions may also play a role in limiting the results. For example, minority and non-minority cases included only two race classes (e.g., Black/African-American and White/Caucasian); however, the proportions of the two race classes added to 100% of cases. This is, of course, not representative of the actual demographic profile of most organizations, which consists of employees who represent a wide variety of races. In turn, future researchers should consider incorporating samples that represent other race classes that can be compared to the majority group for adverse impact evaluations. Additional methodological limitations include factors such as the subjective selection of sample size and the number of predictor variables to include in each study. That said, all such decisions were made in an effort to provide an educated approximation of the reality faced by the type of organizations that are engaged in the prediction of employee outcomes and large-scale organizational decision-making. Nevertheless, future research that investigates the current study's research questions on varying levels of size and scale would be beneficial to the field.

Conclusions

Overall, strong support for the use of decision tree analysis in the practice of personnel decision-making was derived from the current study. Such support includes results that demonstrate the ability for decision tree analysis to produce predictive models that have: (1) nearly equal levels of accuracy as models produced by traditional regression techniques, (2) little to no adverse impact, and (3) precise cutoff points, visual output, and overall model simplicity. Additionally, it is suggested that decision tree analysis could be used to predict multiple organizational outcomes and produce accurate models in both training and validation samples.

As a result of the aforementioned results, all research questions are answered affirmatively and denote specific benefits provided by decision tree analysis. However, such results also suggest implications, which extend above and beyond the benefits addressed by the current study's research questions. For example, it is possible that by incorporating the use of decision tree analysis in organizations, that organizational leaders gain a greater appreciation for both employee data and the employees themselves. Specifically, by improving the process by which organizations use employee data, organizations are likely to better understand their workforce and make improved personnel decisions that positively affect employees. Overall, this process could lead to improved organizational and employee outcomes (e.g., productivity).

Additionally, study results are associated with major implications for the practice of personnel decision-making as a whole. Specifically, aspects of personnel decision-making (e.g., personnel selection) have played a role in the lives of organizations and employees for nearly 100 years (Hunter, 1983). Traditionally, organizations have relied on traditional regression techniques to assist them in developing models for making employee decisions (Raju, Steinhaus, Edwards, & DeLessio, 1991; Stauffer & Ree, 1996). Moreover, regression is one of the most frequently employed statistical techniques in all of the social sciences (Sanders & Brynin, 1998). However, results from the current study can be used to challenge this convention and suggest that the novel statistical technique of decision tree analysis can provide organizational sciences with benefits not currently experienced. In sum, results not only support a shift in organizational practice, but also in academic research, and can be used to suggest that the use of decision tree analysis is not only groundbreaking to organizations, but also to entire fields of study.

REFERENCES

Abdi, H. (2007). Part (semi partial) and partial regression coefficients. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 1-9). Thousand Oaks, CA: Sage.

Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975).

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models*. Newbury Park, CA: Sage.

Anderson, N. (2005). Relationship between practice and research in personnel selection: Does the left hand know what the right is doing?. In A. Evers, N. Anderson, & O. Voskuijl (Eds.), *The Blackwell handbook of personnel selection* (pp. 1-24). Malden, MA: Blackwell.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 508-600). Washington, DC: American Council on Education.

Axelrod, B., Handfield-Jones, H., & Welsh, T. (2001). War for talent, part two. *McKinsey Quarterly, 2*, 9-11.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111-118.

Beechler, S., & Woodward, I. C. (2009). The global "war for talent". *Journal of International Management, 15*, 273-285.

Behrman, D. N., & Perreault, W. D. (1982). Measuring the performance of industrial salespersons. *Journal of Business Research, 10*, 355-370.

Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research, 56*, 137-172.

Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. *Personnel Psychology, 60*, 271-301.

Board of Regents of the University of the State of New York v. Tomanio, 446 U.S. 478 (1980).

Bobko, P., & Roth, P. L. (2010). An analysis of two methods for assessing and indexing adverse impact: A disconnect between the academic literature and some practice. In J. L. Outtz (Ed.), *Adverse Impact: Implications for Organizational Staffing and High Stakes Selection* (pp. 29-49). New York, NY: Routledge.

Borman, W. C., Hanson, M. A., & Hedge, J. W. (1997). Personnel selection. *Annual Review of Psychology, 48*, 299-337.

Bowling, N. A., Eschleman, K. J., & Wang, Q. (2010). A meta-analytic examination of the relationship between job satisfaction and subjective well-being. *Journal of Occupational and Organizational Psychology, 83*, 915-934.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, P. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, P. J. (1993). *Classification and regression trees* (2nd ed.). Belmont, CA: Wadsworth International Group.

Campbell, J. P. (2012). Behavior, performance, and effectiveness in the twenty-first century. In S. W. J. Kozlowski (Ed.), *The Oxford handbook of organizational psychology* (pp. 159-196). New York, NY: Oxford University Press.

Carless, S. A. (2009). Psychological testing for selection purposes: A guide to evidence-based practice for human resource professionals. *The International Journal of Human Resource Management, 20*, 2517-2532.

Cascio, W. F., Alexander, R. A, & Barrett, G. V. (1988). Setting cutoff scores: Legal, psychometric, and professional issues and guidelines. *Personnel Psychology, 41*, 1-24.

Cascio, W. F., Outtz, J., Zedeck, S., & Goldstein, I. L. (1991). Statistical implications of six methods of test score use in personnel selection. *Human Performance, 4*, 233-264.

Cascio, W. F., Outtz, J., Zedeck, S., & Goldstein, I. L. (1995). Statistical implications of six methods of test score use in personnel selection. *Human Performance, 8*, 133-164.

Chambers, E. G., Foulon, M., Handfield-Jones, H., Hankin, S. M., & Michaels, E. G. (1998). The war for talent. *The McKinsey Quarterly, 3*, 44-57.

Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82*, 300-310.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods, 4*, 62-83.

Chien, C., & Chen, L. (2008). Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Systems with Applications, 34*, 280-290.

Civil Rights Act of 1964, 7, 42 U.S.C., 2000e *et seq* (1964).

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Connor, J. P., Symons, M., Feeney, G. F. X., Young, R., & Wiles, J. (2007). The application of machine learning techniques as an adjunct to clinical decision making in alcohol dependence treatment. *Substance Use and Misuse, 42*, 2193-2206.

Cook, N. R. (2007). Use and misuse of receiver operating characteristic curve in risk prediction. *Ciruclation, 115*, 928-935.

Corte, W. D. (1995). Cutoff scores that optimize the utility of a fixed quota of successful employees selection decision. *International Journal of Selection and Assessment, 3*, 1-10.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment, 4*, 5-13.

Cox, T., Griffiths, A., & Rial-González, E. (2000). *Research on work-related stress*. European Agency for Safety & Health at Work. Luxembourg, DE: Office for Official Publications of the European Communities.

Czepiel, S. A. (2002). Maximum likelihood estimation of logistic regression models: Theory and implementation. Retrieved from: http://czep.net/stat/mlelr.pdf.

De Neve, J. E., Diener, E., Tay, L., & Xuereb, C. (2013). The objective benefits of subjective well-being. In J. F. Helliwell, R. Layard, & J. Sachs (Eds.), *World happiness report 2013* (Vol. 2, pp. 54-79). New York, NY: UN Sustainable Network Development Solutions Network.

DeMaris, A. (2013). Logistic regression: Basic foundations and new directions. In J. A. Schinka, W. F. Velicer, & I. B. Weiner (Eds.), *Handbook of psychology, Vol. 2: Research methods in psychology (2nd ed.)* (pp. 543-570). Hoboken, NJ: Wiley.

DeNeve K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124*, 197-229.

DeVille, B., & Neville, P. (2013). *Decision trees for analytics using SAS enterprise miner*. Cary, NC: SAS Institute Inc.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*, 542-575.

Diener, E. (2009). *The science of well-being: The collected works of Ed Diener* (Social Indicators Research Series, Vol. 37). Amsterdam, NL: Springer.

Doyle, P. (1973). The use of automatic interaction detector and similar search procedures. *Operational Research Quarterly, 24*, 465-467.

Dunnette, M. D., & Borman, W. C. (1979). Personnel selection and classification systems. *Annual Review of Psychology, 30*, 477-525.

Duttweiler, P. C. (1984). The internal control index: A newly developed measure of locus of control. *Educational and Psychological Measurement, 44*, 209-221.

Economist Intelligence Unit. (2006). *The CEO's role in talent management: How top executives from ten countries are nurturing the leaders of tomorrow.* New York, NY: The Economist.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guildford Press.

Equal Employment Opportunity Commission (EEOC; 1978). *Uniform guidelines on employee selection procedures.*

Erickson, P. B. (2004, May 16). Employer hiring tests grow sophisticated in quest for insight about applicants. *Knight Ridder Tribune Business News,* 1.

Faragher, E., Cass, M., & Cooper, C. (2005). The relationship between job satisfaction and health: A meta-analysis. *Occupational and Environmental Medicine, 62*, 105-112.

Farndale, E., Scullion, H., & Sparrow, P. (2010). The role of the corporate HR function in global talent management. *Journal of World Business, 45*, 161-168.

Farrington, D. P., & Loeber, R. (2000). Some benefits of dichotomization in psychiatric and criminological research. *Criminal Behaviour and Mental Health, 10*, 100-122.

Finch, D. M., Edwards, B. D., & Wallace, J. C. (2009). Multistage selection strategies: Simulating the effects on adverse impact and expected performance for various predictor combinations. *Journal of Applied Psychology, 94*, 318-340.

Flynn, J. T., & Peterson, M. (1972). The use of regression analysis in police patrolman selection. *Journal of Criminal Law, Criminology, and Police Science, 63*, 564-569.

Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five U.S. racial groups. *Personnel Psychology, 61*, 579-616.

Freyd, M. (1923). Measurement in vocational selection: An outline of research procedure. *Journal of Personnel Research, 2*, 215-249.

Giffin, M. E. (1989). Personnel research on testing, selection, and performance appraisal. *Public Personnel Management, 18*, 127-137.

Goldberg, L. R. (1992). The development of markers of the big-five factor structure. *Psychological Assessment, 4*, 26-42.

Gordon, L. (Ed.). (2013). Proceedings from SAS Global Forum 2013: *Using Classification and Regression Trees (CART) in SAS Enterprise Miner for Applications in Public Health*. Cary, NC: SAS.

Gorman, C. A., Thibodeaux, C. N., Eisinger, S. E., & Overstreet, B. L. (2012). Selection for training: The forgotten employment decision?. In N. P. Reilly, M. J. Sirgy, & C. A. Gorman (Eds.), *Work and quality of life: Ethical practices in organizations* (pp. 95-105). New York, NY: Springer.

Gowan, M. A., & Gatewood, R. D. (2013). Personnel selection. In N. Brewer, & C. Wilson (Eds.), *Psychology and policing* (pp. 177-204). Hillsdale, NJ: Lawrence Erlbaum Associates.

Griggs v. Duke Power Company, 401 U.S. 424 (1971).

Guion, R. M., & Gibson, W. M. (1988). Personnel selection and placement. *Annual Review of Psychology, 39*, 349-374.

Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine, 4*, 627-635.

Halinski, R. S., & Feldt, L. S. (1970). The selection of variables in multiple regression analysis. *Journal of Educational Measurement, 7*, 151-157.

Harrison, P. L., Pope, J. E., Coberley, C. R., & Rula, E. Y. (2012). Evaluation of the relationship between individual well-being and future health care utilization and cost. *Population Health Management, 15*, 325-330.

Heller, M. (2005). Court ruling that employer's integrity test violated ADA could open door to litigation. *Workforce Management, 84*, 74-77.

Herrero, C., & Villar, A. (2013). On the comparison of group performance with categorical data. *PLoS One, 8*, 1-7.

Hoffman, C. C., & Thornton, G. C. (1997). Examining selection utility where competing predictors differ in adverse impact. *Personnel Psychology, 50*, 455-470.

Hoffmeister, K., Gibbons, A., Schwatka, N., & Rosecrance, J. (2015). Ergonomics climate assessment: A measure of operational performance and employee well-being. *Applied Ergonomics, 50*, 160-169.

Hogan, J., & Brinkmeyer, K. (1997). Bridging the gap between overt and personality-based integrity tests. *Personnel Psychology, 50*, 587-599.

Hogan, R., & Hogan, J. (1995). *Hogan personality inventory manual.* Tulsa, OK: Hogan Assessment Systems.

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment, 15*, 446-455.

Hunter, J. E. (1980). *Test validation for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB).* Washington, DC: U.S. Employment Service.

Hunter, J. E. (1983). The economic benefits of personnel selection using ability tests: A state of the art review including a detailed analysis of the dollar benefit of U.S. Employment Service placements and a critique of the low-cutoff method of test use (USDOL USES Test Research Report No. 47). Washington, DC: U.S. Employment Service.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.

Hurley, R. F., & Estelami, H. (2007). An exploratory study of employee turnover indicators as predictors of customer satisfaction. *Journal of Services Marketing, 21*, 186-199.

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*, 869-879.

Jaccard, J. J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Thousand Oaks, CA: Sage.

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits – self-esteem, generalized self-efficacy, locus of control, and emotional stability – with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*, 80-92.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?. *Journal of Personality and Social Psychology, 83*, 693-710.

Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*, 530-541.

Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., & Rich, B. L. (2007). Self-efficacy and work-related performance: The integral role of individual differences. *Journal of Applied Psychology, 92*, 107-127.

Judge, T. A., Locke, E. A., & Durham, C. C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior, 19*, 151-188.

Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin, 127*, 376-407.

Karels, T. J., Bryant, A. A., & Hik, D. S. (2004). Comparison of discriminant function and classification tree analyses for age classification of marmots. *Oikos, 105*, 575-587.

Kass, G. V. (1975). Significance testing in automatic interaction detection (A.I.D.). *Journal of the Royal Statistical Society, 24*, 178-189.

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics, 29*, 119–127.

Kriska, S. D., & Milligan, G. W. (1982). Multiple regression analysis for categorical data with an illustrative application in personnel selection. *Psychological Bulletin, 92*, 193-202.

Lavanya, D., & Rani, K. U. (2011). Performance evaluation of decision tree classifiers on medical datasets. *International Journal of Computer Applications, 26*, 1-4.

Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: Methodological review and comparison with logistic regression. *Annals of Behavioral Medicine, 26*, 172-181.

Leslie, D. S., Kohn, R., & Nott, D. J. (2007). A general approach to heteroscedastic linear regression. *Statistics and Computing, 17*, 131-146.

Lewis, R. J. (2000, May). *An introduction to classification and regression tree (CART) analysis.* Paper presented at the Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, California.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297-1343). Chicago, IL: Rand McNally.

Loh, W. Y. (2008). Classification and regression tree methods. In F. Ruggeri, R. Kenett, & F. Faltin (Eds.), *Encyclopedia of statistics in quality and reliability* (pp. 315-323). Hoboken, NJ: Wiley.

Loh, W. Y. (2009). Improving the precision of classification trees. *The Annals of Applied Statistics, 3,* 1710-1737.

Loh, W. Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1,* 14-23.

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician, 54,* 217-224.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health, 23,* 151-169.

Maiback, E., & Murphy, D. A. (1995). Self-efficacy in health promotion research and practice: Conceptualization and measurement. *Health Education Research, 10,* 37-50.

Maurer, T. J., Alexander, R. A., Callahan, C. M., Bailey, J. J., & Dambrot, F. H. (1991). Methodological and psychometric issues in setting cutoff scores using the Angoff method. *Personnel Psychology, 44,* 235-262.

McCornack, R. L. (1970). A comparison of three predictor selection techniques in multiple regression. *Psychometrika, 35,* 257-271.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 56*, 586-595.

McCrae, R. R., & Costa, P. T. (1991). Adding liebe und arbeit: The full five-factor model and well-being. *Personality and Social Psychology Bulletin, 17*, 227-232.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63-91.

McDonald, T., Poertner, J., & Harris, G. (2001). Predicting placement in foster care: A comparison of logistic regression and neural network analysis. *Journal of Social Science Research, 28*, 1-20.

Meijer, L. A. L., Born, M. P., Zielst, J., & Molen, H. T. (2010). Construct-driven development of a video-based situation judgment test for integrity. *European Psychologist, 15*, 229-236.

Mellahi, K., & Collings, D. G. (2010) The barrier to effective global talent management: The example of corporate elites in MNEs. *Journal of World Business, 45*, 143-149.

Mendoza, J. L., Bard, D. E., Mumford, M. D., & Ang, S. C. (2004). Criterion-related validity in multiple-hurdle designs: Estimation and bias. *Organizational Research Methods, 7*, 418-441.

Morris, S. B., Daisley, R. L., Wheeler, M., & Boyer, P. (2015). A meta-analysis of the relationship between individual assessments and job performance. *Journal of Applied Psychology, 100*, 5-20.

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the big five personality traits and the big six vocational interest types. *Personnel Psychology, 58*, 447-478.

Ngo, T. F., Govindu, R., & Agarwal, A. (2015). Assessing the predictive utility of logistic regression, classification and regression tree, chi-squared automatic interaction detection, and neural network models in predicting inmate misconduct. *American Journal of Criminal Justice, 40*, 47-74.

Nishii, L. H., & Wright, P. M. (2007). *Variability within organizations: Implications for strategic human resource management* (CAHRS Working Paper #07-02). Ithaca, NY: Cornell University.

Ones, D. S. (1993). *The construct validity of integrity tests* (Doctoral dissertation). Retrieved from ProQuest. (9404525)

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*, 679-703.

Outtz, J. L., & Newman, D. A. (2010). A theory of adverse impact. In J. L. Outtz (Ed.), *Adverse impact: Implications for organizational staffing and high stakes selection* (pp. 53-94). New York, NY: Routledge.

Peng, C. Y. J., Lee, K. L, & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*, 3-14.

Peterson, N. G., & Wing, H. (2001). The search for new measures: Sampling from a population of selection/classification predictor variables. In J. P. Campbell, & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 53-70). Mahwah, NJ: Lawrence Erlbaum Associates.

Pierce, J. L., & Gardner, D. G. (2004). Self-esteem within the work and organizational context: A review of the organization-based self-esteem literature. *Journal of Management, 30*, 591-622.

Piper, M. E., Loh, W. Y., Smith, S. S., Japuntich, S. J., & Baker, T. B. (2011). Using decision tree analysis to identify risk factors for relapse to smoking. *Substance Use and Misuse, 46*, 492-510.

Pohlman, J. T., & Leitner, D. W. (2003). A comparison of ordinary least squares and logistic regression. *Ohio Journal of Science, 103*, 118-125.

Quick, J. C., Tetrick, L. E., Adkins, J. A., & Klunder, C. S. (2003). Occupational health psychology. In I. Weiner (Ed.), *Comprehensive handbook of psychology* (pp. 569-590). New York, NY: Wiley.

Quinlan, J. R. (1986). Introduction of decision tree. *Machine Learning, 1*, 81-106.

Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo, CA: Morgan Kaufman.

Raju, N. S., Steinhaus, S. D., Edwards, J. E., & DeLessio, J. (1991). A logistic regression model for personnel selection. *Applied Psychological Measurement, 15*, 139-152.

Raymond, J. S., Wood, D. W., & Patrick, W. K. (1990). Psychology doctoral training in work and health. *American Psychologist, 45*, 1159-1161.

Robertson, I. T., & Smith, M. (2001). Personnel selection. *Journal of Occupational and Organizational Psychology, 74*, 441-472.

Rogers v. International Paper Co., 510 F.2d 1340 (8th Cir. 1975).

Rossi, R. J. (2010). *Applied biostatistics for the health sciences*. Hoboken, NJ: Wiley.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support?. *Human Resource Management Review, 16*, 155-180.

Rothstein, M. G., Paunonen, S. V., Rush, J. C., & King, G. A. (1994). Personality and cognitive ability predictors of performance in graduate business school. *Journal of Educational Psychology, 86*, 516-530.

Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419-450.

Sackett, P. R., & Roth, L. (1996). Multi-stage selection strategies: A Monte Carlo investigation of effects on performance and minority hiring. *Personnel Psychology, 49*, 549-572.

Salgado, J. F., Visewsvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods and techniques. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswevaran (Eds.), *Handbook of Industrial, Work & Organizational Psychology, Volume 1, Personnel Psychology* (pp. 165-199). Thousand Oaks, CA: Sage.

Sanders, D., & Brynin, M. (1998). Ordinary least squares and logistic regression analysis. In E. Scarbrough, & E. Tanenbaum (Eds.), *Research strategies in the social sciences: A guide to new approaches* (pp. 29-52). New York, NY: Oxford University Press.

SAS Institute, Inc. (2013). SAS Enterprise Miner [computer software]. Cary, NC: SAS Institute, Inc.

Sauter, S. L., Hurrel, J. J., Fox, H. R., Tetrick, L. E., & Barling, J. (1999). Occupational health psychology: An emerging discipline. *Industrial Health, 37*, 199-211.

Schaefer, R. L. (1986). Alternative estimators in logistic regression when the data are collinear. *Journal of Statistical Computation and Simulation, 25*, 75-91.

Schmidt, F. L. (1991). Why all banding procedures are logically flawed. *Human Performance, 4*, 265-277.

Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologists, 36*, 1128-1137.

Schmidt, F. L., & Hunter, J. E. (1995). The fatal internal contradiction in banding: Its statistical rationale is logically inconsistent with its operational procedures. *Human Performance, 8*, 203-214.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Schmidt, F. L., Mack, M. J., & Hunter, J. E. (1984). Selection utility in the occupation of U.S. park ranger for three modes of test use. *Journal of Applied Psychology, 69*, 490-497.

Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology, 43*, 627-670.

Schultz, A. B., & Edington, D. W. (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation, 17*, 547-579.

Schwartz, S. M., & Riedel, J. (2010). Productivity and health: Best practices for better measures of productivity. *Journal of Occupational and Environmental Medicine, 52*, 865-871.

Scullion, H., Collings, D. G., & Caligiuri, P. (2010). Global talent management. *Journal of World Business, 45*, 105-108.

Sears, L. E., Shi, Y., Coberley, C. R., & Pope, J. E. (2013). Overall well-being as a predictor of health care productivity, and retention outcomes in a large employer. *Population Health Management, 16*, 397-405.

Sharma, A. K. (2005). *Textbook of correlation and regression.* New Delhi, IN: Discovery.

Shepard, L. (1980). Standard setting issues and methods. *Applied Psychological Measurement, 4*, 447-467.

Singh, S., Darwish, T. K., Costa, A. C., & Anderson, N. (2012). Measuring HRM and organisational performance: Concepts, issues, and framework. *Management Decision, 50*, 651-667.

Sonquist, J. A., Baker, E. L., & Morgan, J. N. (1971). *Searching for structure*. Ann Arbor: Institute for Social Research.

Sparks, K., Faragher, B., & Cooper, C. L. (2001). Well-being and occupational health in the 21st century workplace. *Journal of Occupational and Organizational Psychology, 74*, 489-509.

Stauffer, J., & Ree, M. J. (1996). Predicting with logistic or linear regression: Will it make a difference in who is selected for pilot traning?. *The International Journal of Aviation Psychology, 6*, 233-240.

Stewart, W. F., Ricci, J. A., Chee, E., & Morganstein, D. (2003). Lost productive work time costs from health conditions in the United States: Results from the American Productivity Audit. *Journal of Occupational and Environmental Medicine, 45*, 1234-1246.

Stone, C. H., & Kendall, W. E. (1956). Setting test standards: Interpreting and reporting test results. In C. H. Stone, & W. E. Kendall (Eds.), *Effective Personnel Selection Procedures* (pp. 286-315). Englewood Cliffs, NJ: Prentice-Hall.

Sweet, S. A., & Grace-Martin, K. A. (2012). Modeling relationships of multiple variables with linear regression. In S. A. Sweet, & K. A. Grace-Martin (Eds.), *Data Analysis with SPSS: A First Course in Applied Statistics* (pp. 161-188). New York, NY: Pearson.

Tait, M., Padgett, M. Y., & Baldwin, T. T. (1989). Job and life satisfaction: A reevaluation of the strength of the relationship and gender effects as a function of the date of the study. *Journal of Applied Psychology, 74*, 502-507.

Temme, L. A., Still, D. L., & Fatcheric, A. J. (1995). Jet pilot, helicopter pilot, and college student: A comparison of central vision. *Aviation, Space, and Environmental Medicine, 66*, 297-302.

Terpstra, D. E., & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology, 46*, 27-48.

Thomas, S. L., & Scroggins, W. A. (2006). Psychological testing in personnel selection: Contemporary issues in cognitive ability and personality testing. *Journal of Business Inquiry, 5*, 28-38.

Thornton, G. C., & Kedharnath, U. (2013). Work sample tests. In K. F. Geisinger, & B. A. Bracken (Eds.), *APA handbook of testing and assessment in psychology* (pp. 533-550). Washington, DC: APA.

Tipton, R. M., & Worthington, E. L. (1984). The measurement of generalized self-efficacy: A study of construct validity. *Journal of Personality Assessment, 48*, 545-548.

Tonidandel, S. T., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business Psychology, 26*, 1-9.

Trevor, C. O., Gerhart, B., & Boudrea, J. W. (1997). Voluntary turnover and job performance: Curvilinearity and the moderating influences of salary growth and promotions. *Journal of Applied Psychology, 82*, 44-61.

Truxillo, D. M., Donahue, L. M., & Sulzer, J. L. (1996). Setting cutoff scores for personnel selection tests: Issues, illustrations, and recommendations. *Human Performance, 9*, 275-295.

Vaiman, V., Scullion, H., & Collings, D. (2012). Talent management decision making. *Management Decision, 50*, 925-941.

Van De Voorde, K., Paauwe, J., & Veldhoven, M. V. (2012). Employee well-being and the HRM-organizational performance relationship: A review of quantitative studies. *International Journal of Management Reviews, 14*, 391-407.

Van Iddekinge, C. H., Roth, P. L., Raymark, P. H., & Odle-Dusseau, H. N., (2012). The criterion-related validity of integrity tests: An updated meta-analysis. *Journal of Applied Psychology, 97*, 499-530.

Vaughan, T. S., & Berry, K. E. (2005). Using monte carlo techniques to demonstrate the meaning and implications of multicollinearity. *Journal of Statistics Education, 13*, 1-9.

Vickery, D., Golaszewski, T., Wright, E., & McPhee, L. (1986). Lifestyle and organizational health insurance costs. *Journal of Occupational Medicine, 28*, 1165-1168.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*, 557-574.

Vulcan Pioneers v. New Jersey Department of Civil Service, 625 F. Supp. 527 (D. N. J. 1985).

Wallston, B. S., & Wallston, K. A. (1978). Locus of control and health: A review of the literature. *Health Education Monographs, 6*, 107-117.

Weissmuller, J. J., & Damos, D. L. (2014). Improving the pilot selection system: Statistical approaches and selection processes. *The International Journal of Aviation Psychology, 24*, 99-118.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance, 21*, 291-309.

Williams, C. R., & Livingston, L. P. (1994). Another look at the relationship between performance and voluntary turnover. *The Academy of Management Journal, 37*, 269-298.

Worth, A. P., & Cronin, M. T. D. (2003). The use of discriminant analysis, logistic regression and classification tree analysis in the development of classification models for human health effects. *Journal of Molecular Structure: THEOCHEM, 622*, 97-111.

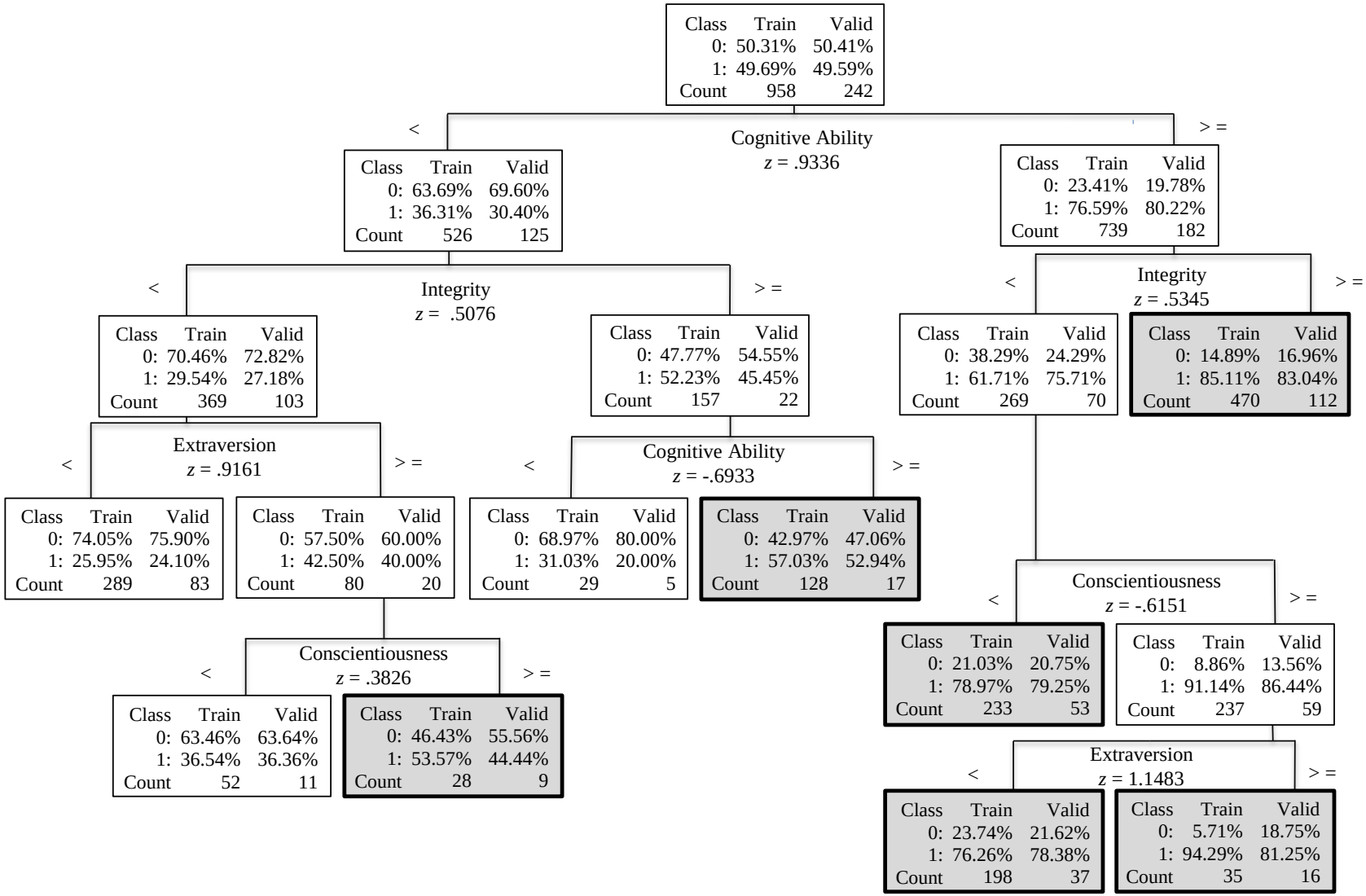
Yen, L. T., Edington, D. W., & Witting, P. (1991). Associations between employee health-related measures and prospective medical insurance costs in a manufacturing company. *American Journal of Health Promotion, 6*, 46-54.

Zhang, S. F., & Liu, S. Y. (2011). A GRA-based intuitionistic fuzzy multi-criteria group decision making method for personnel selection. *Expert Systems with Applications, 38*, 11401-11405.

Zhou, W., & Kapoor, G. (2011). Detecting evolutionary financial statement fraud. *Decision Support Systems, 50*, 570-575.

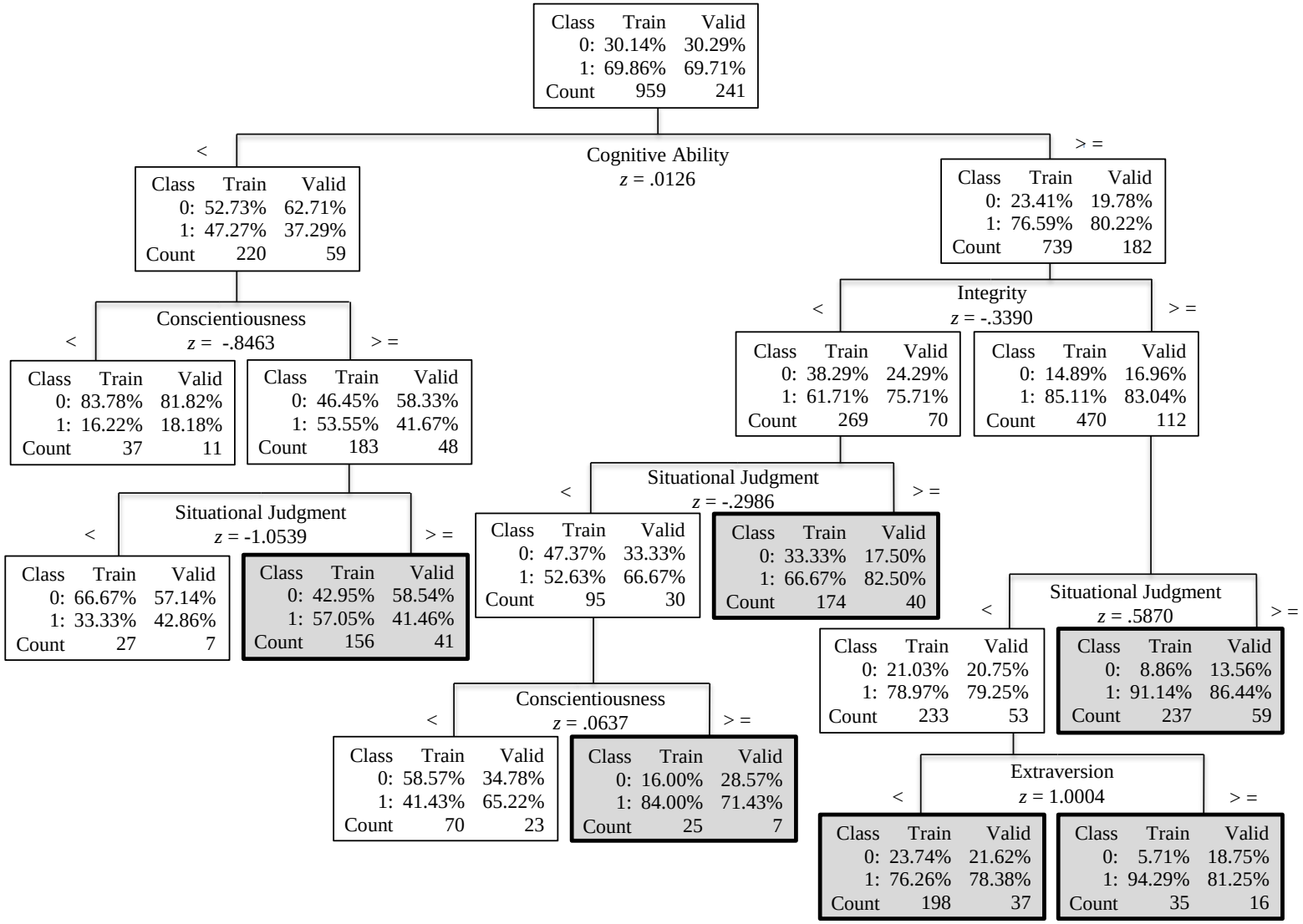
APPENDIX A: DECISION TREE MODELS

Study 1: Data Set 1



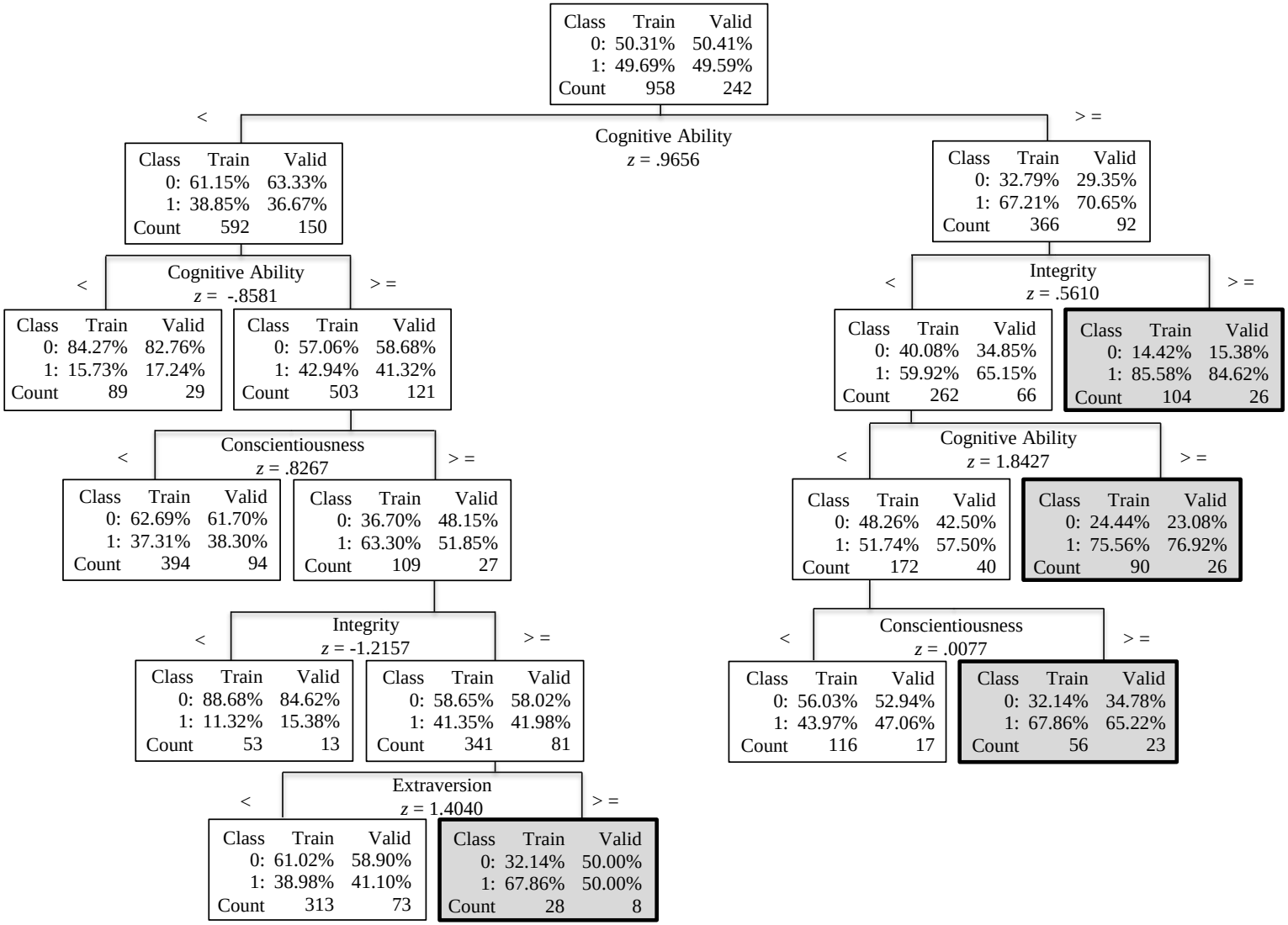
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 2



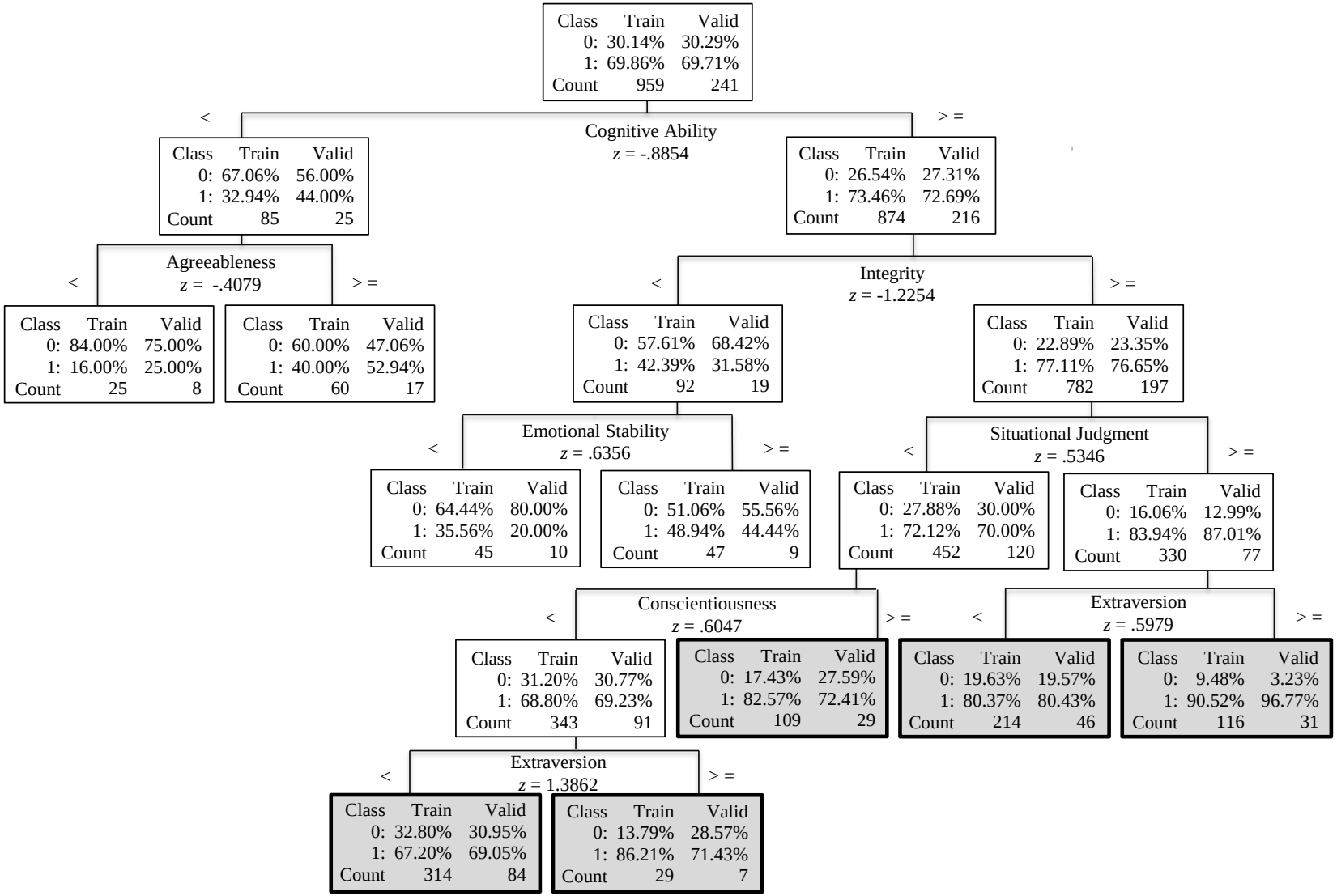
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 3



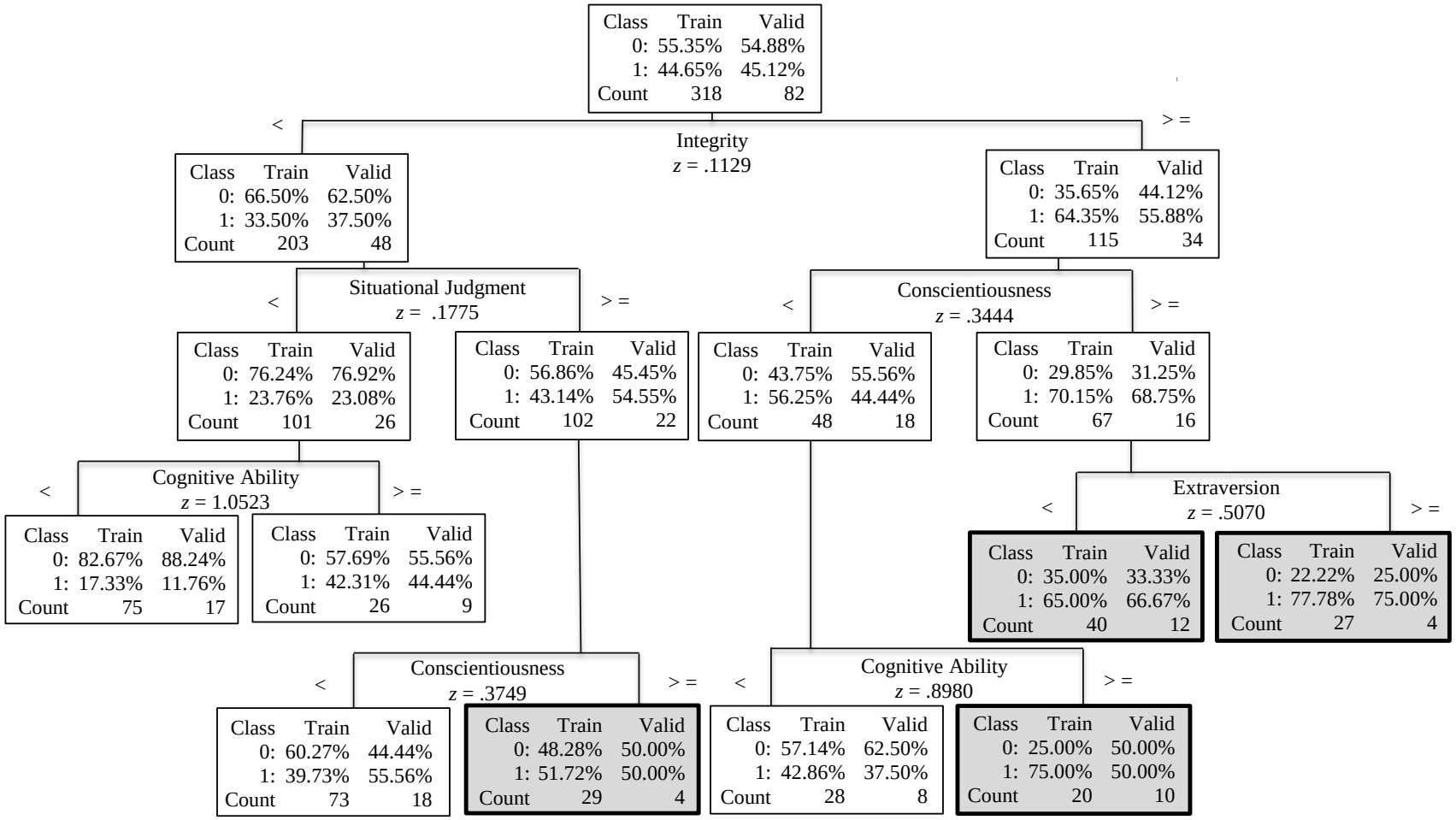
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 4



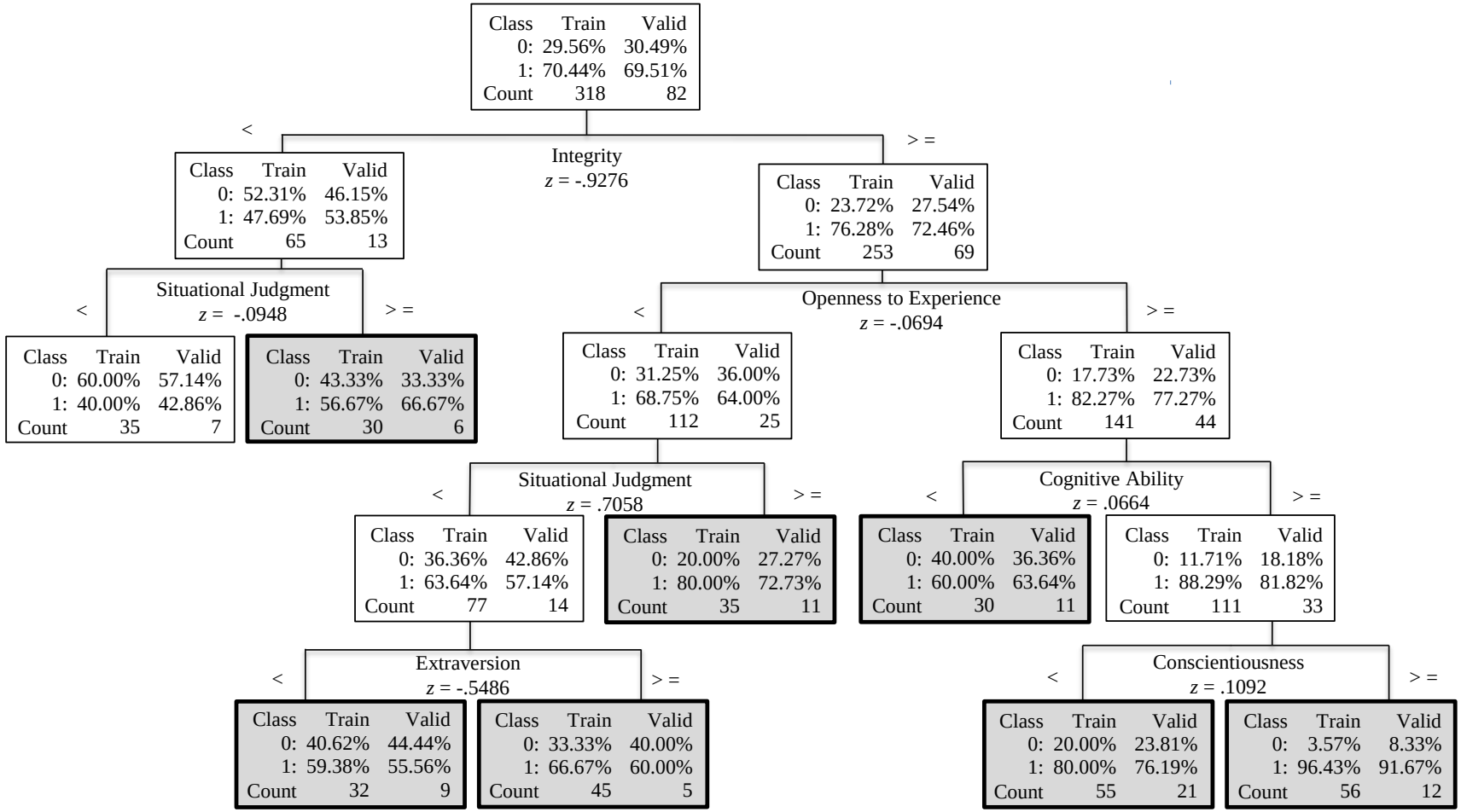
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 5



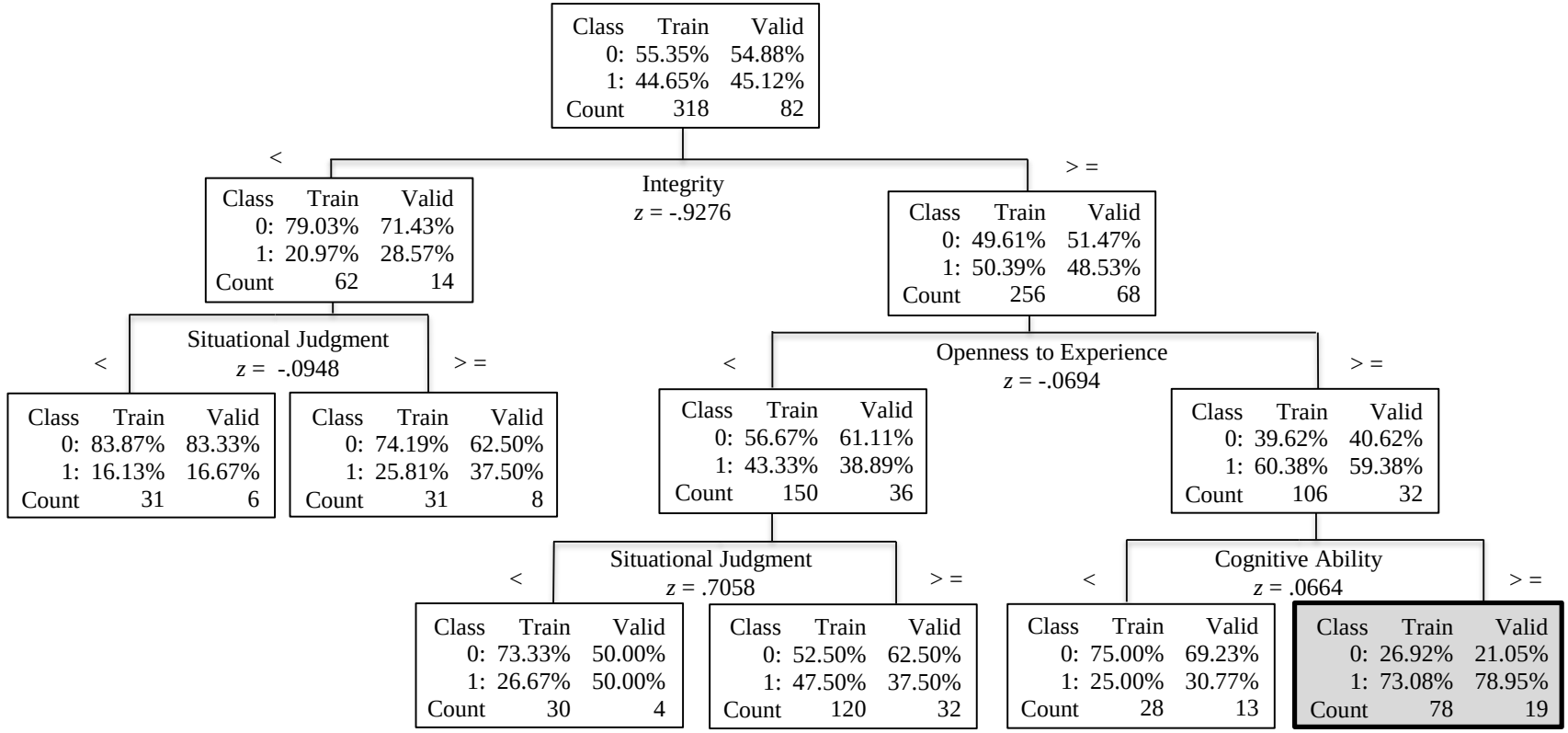
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 6



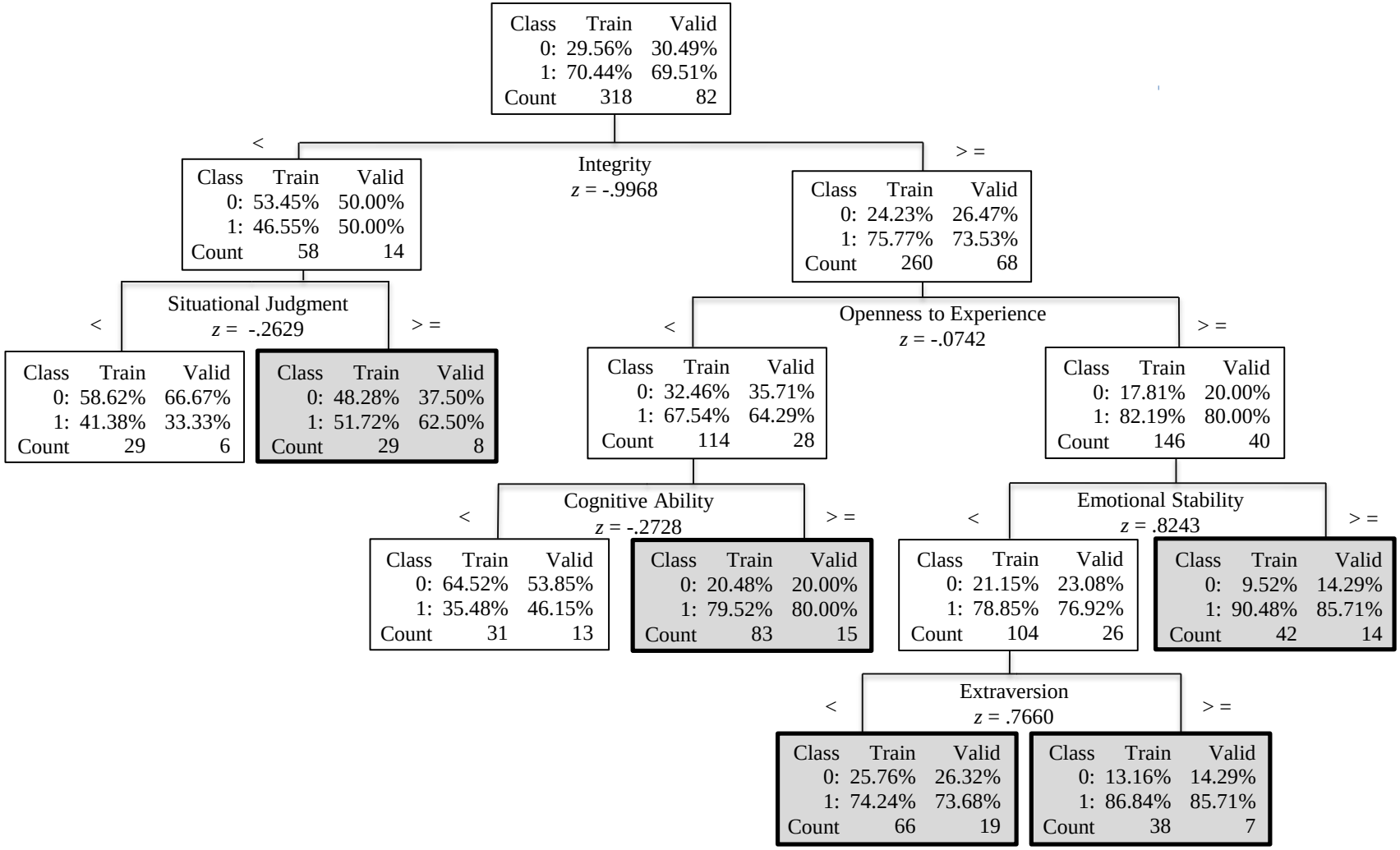
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 7



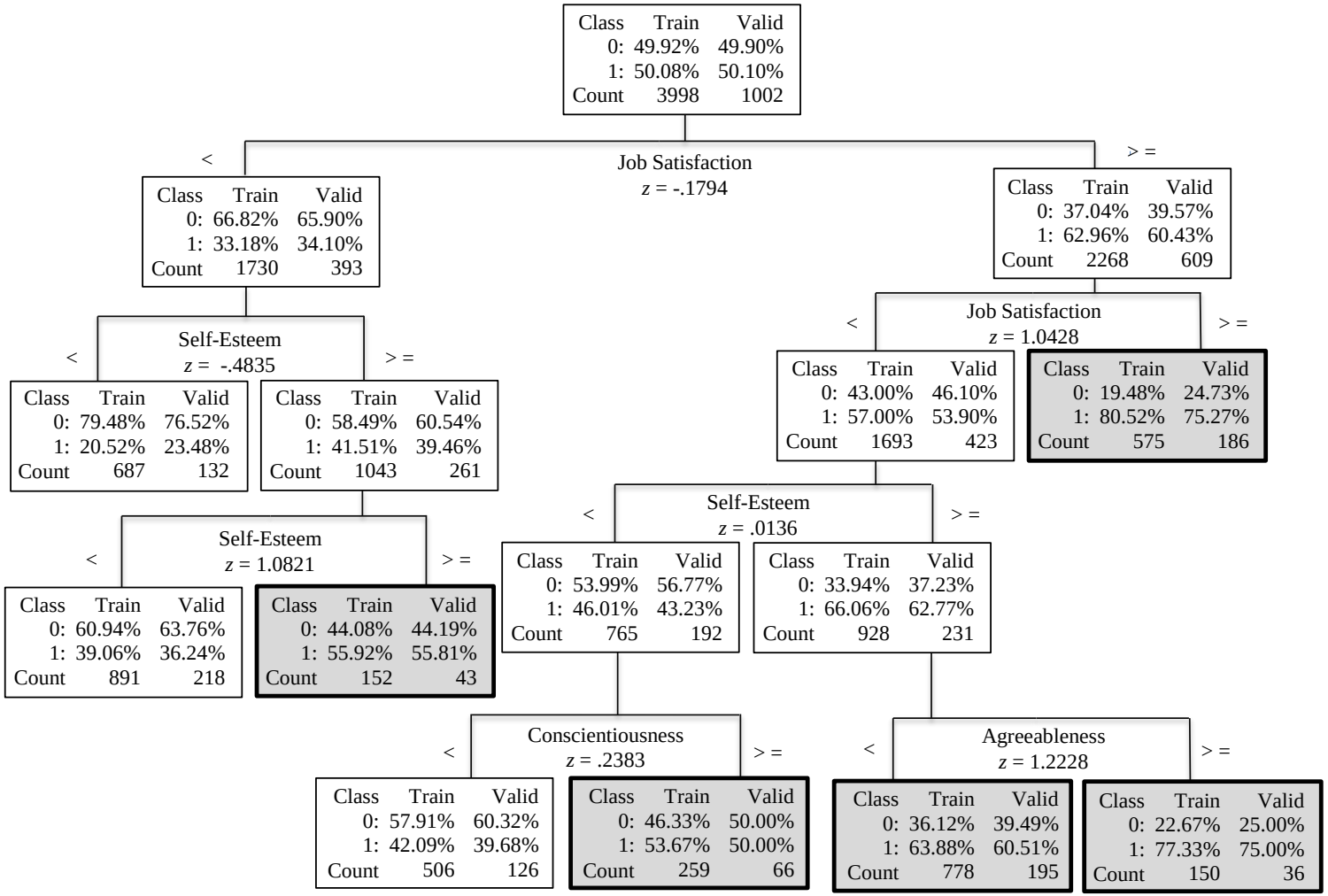
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 1: Data Set 8



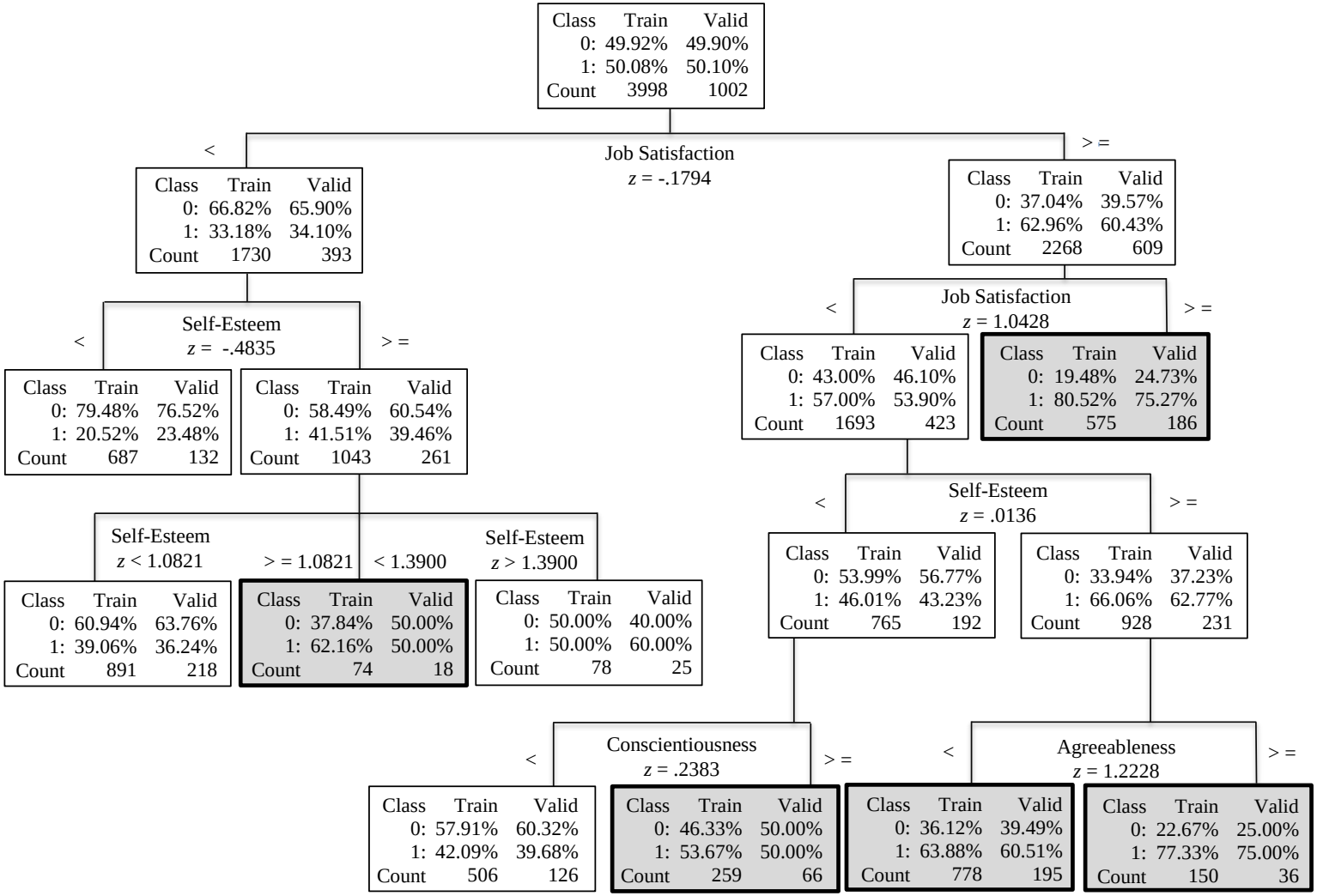
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 2: Data Set 1



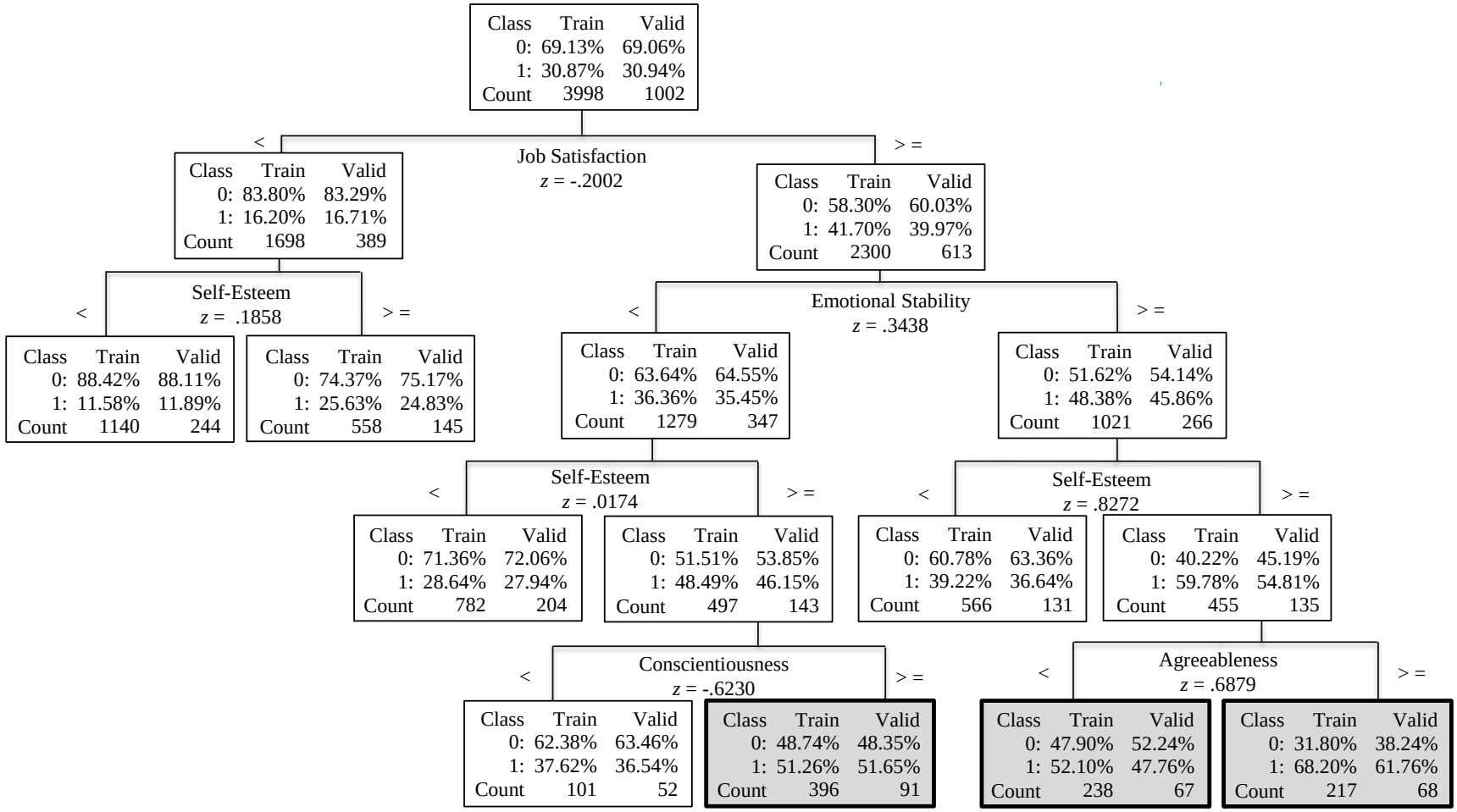
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 2: Data Set 1 (Three-Way Split)



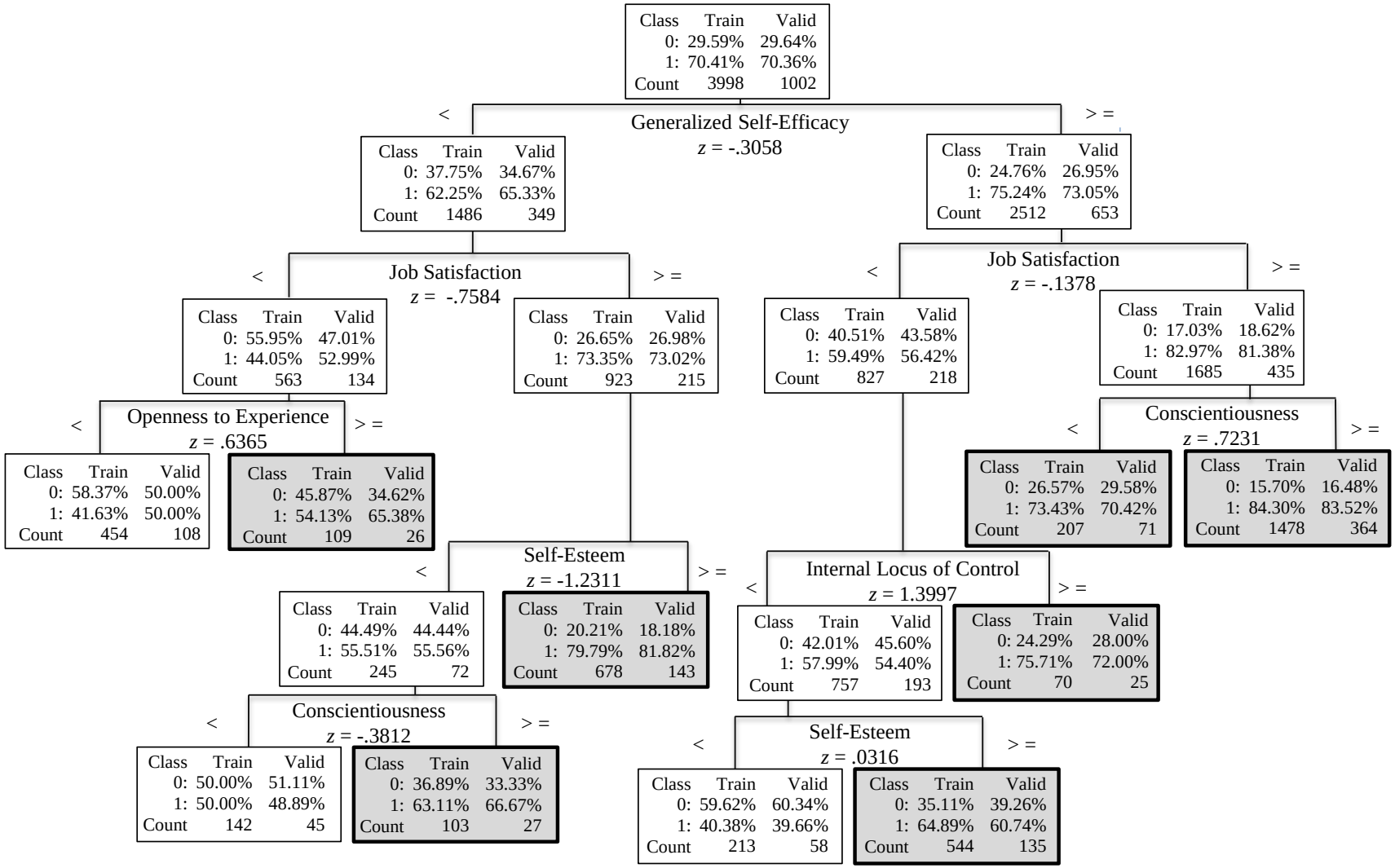
Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 2: Data Set 2



Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.

Study 2: Data Set 3



Notes. Shaded boxes indicate terminal nodes with a predicted target class (i.e., “1”) of greater than 50.00%.