1-1-2012

The development and validation of a conditional reasoning test of withdrawal

Levi Ryan gust Nieminen
Wayne State University,

THE DEVELOPMENT AND VALIDATION OF A CONDITIONAL REASONING TEST OF WITHDRAWAL

by

LEVI R. G. NIEMINEN

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2012

MAJOR: PSYCHOLOGY
(Industrial/Organizational)

Approved by:

____
Advisor Date

Dedication

To the life changing teachers that helped me find Psychology and Graduate School:

Joseph Hibbeln, M.D.

Susan Ratwik, Ph.D.

Tim Sawyer, Ph.D.

Ralph Barnes, Ph.D.

ACKNOWLEDGMENTS

I would like to give my sincere thanks to the many people that supported me throughout my graduate career and during my dissertation research. First, to my family and friends, who have helped me in innumerable ways – e.g., my parents, John and Anne, for buying me the suits that I wore to my graduate school interview, and six years later, to my dissertation defense, and (along with my sister Danielle) for believing in me and supporting my aspirations all the while and in between; my wife, Michele for her daily patience, understanding, and sense of humor, and most importantly, for helping me with numerous iterations of R command; my colleagues at Wayne State who helped with the data collection for this study, and my colleagues at Denison Consulting for giving me the flexibility and support needed to finish strong.

Second, I'd like to thank the members of my dissertation committee for their time, critical insights, and mentorship throughout this project: Boris, Marcus, Larry, and Sub. I greatly appreciate the guidance I received from Larry James and James LeBreton when developing my thinking about implicit personality and when learning how to write conditional reasoning items. I'd also like to thank John Arnold for encouraging me to pursue this line of research, and Marcus Dickson for crucial guidance with the study's design and sample. Lastly, I'd like to thank my graduate advisor and the chair of my dissertation committee, Sebastiano (Sub) Fisicaro for supporting my ideas and helping me to learn how to think thoroughly.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Woody Allen is often quoted as having said that, "95% of life is *just showing up.*" One might go two steps further and suggest that showing up *regularly* and *on time* are basic prerequisites for the majority of life's commitments. Work is perhaps the most obvious application of this axiom, as attendance is a fundamental aspect of job performance in the vast majority of work settings. This is likely true to some degree even in workplaces that allow employees flexibility in the scheduling and location of work (Schmitt, Cortina, Ingerick, & Wiechmann, 2003). Behaviors that are contrary to "good attendance" are typically contained under the spectrum of counterproductive work behaviors, including lateness, absenteeism, and partial work sessions resulting from leaving early, taking excessive breaks, and so on (Spector, Fox, Penney, Bruursema, Goh, & Kessler, 2006). Of course, quitting one's job altogether represents the most severe form of deviation from good attendance behavior.

Collectively, the term *withdrawal* has been used to describe a range of physical, as well as psychological, forms of separation from work. However, the focus of the present study is on the former – behaviors that result in physical absence from one's job. This includes behaviors that can be enacted over multiple occasions, such as absenteeism and lateness, as well as more permanent separations from work, such as the decision to leave a job. Withdrawal behaviors are associated with substantial financial losses for organizations (Cascio, 1987; Mobley, 1982), lost productivity, and negative effects on employee morale and organizational culture (Mueller & Price, 1989; Staw, 1980). Thus, it is unsurprising that employee withdrawal is a serious concern to organizations and has been the focus of a great deal of research in the organizational sciences (Campion, 1991; Maertz & Campion, 2004; Muchinsky & Morrow, 1980; Steers & Mowday, 1980; Lee & Mitchell, 1994). The research literatures on turnover and absenteeism, in particular, are substantial, having enjoyed a period of continuous acceleration since the mid-1970s (Harrison & Martocchio, 1998; Rotundo & Spector, 2010).

A long list of potential strategies has been advanced for reducing the occurrence of employee withdrawal, including incentive programs (e.g., Shotzhauer & Rosse, 1985), realistic job previews (Morse & Popovich, 2009; Wanous, 1973), job enrichment, and transformational leadership, among others. An attractive possibility, which has nevertheless received comparatively less research attention to date, involves explicit consideration of predictors of withdrawal as part of the employee selection process. The guiding question posed by this line of research is the following: Is it possible to use psychological assessments to reliably select employees who are unlikely to have chronic attendance problems or quit a job altogether? Although research has identified several pre-hire variables that are associated with “withdrawal proneness” (Froggatt, 1970), such as personality traits from the Five Factor Model (Judge, Martocchio, & Thoreson, 1997; Taylor, 1968), selection-validation research is nonetheless limited, with available studies reporting modest predictive validity coefficients against withdrawal criteria (e.g., Barrick & Zimmerman, 2005; 2009). Moreover, a number of drawbacks associated with existing assessment techniques are apparent, such as the susceptibility of overt self-report measures to response distortion. Given these limitations, research on new constructs and testing formats has the potential to make an important contribution, particularly if new approaches can complement existing methods to improve prediction of withdrawal.

Overview

The research reported here involves the development and validation of a conditional reasoning test tailored specifically to predict employee withdrawal behaviors (i.e., the Condition Reasoning Test-Withdrawal or CRT-W). Over the past decade, Lawrence James and colleagues have pioneered the theoretical and technological development of conditional reasoning testing and constructed specific measures to assess dispositional achievement motivation (James, 1998) and dispositional aggression (Frost, Ko, & James, 2007; James, McIntyre, Glisson, Bowler, & Mitchell, 2004; James, McIntyre, Glisson, Green, Patton, LeBreton, 2005; James & LeBreton, 2011; LeBreton, Barksdale, & James, 2007). Although published research from this domain has focused solely on these two measures, the conditional reasoning system can be extended to other constructs and could be particularly useful for assessing negatively valenced characteristics such as dispositional withdrawal tendencies (James et al., 2004). This is primarily due to the indirect nature of conditional reasoning tests, which use inductive reasoning problems to infer the presence of implicit cognitive biases.

The overarching goals of the proposed research highlight unique theoretical and practical contributions. A general criticism of the dominant process models for withdrawal is that they are overly reliant on rational decision processes and fail to describe the spontaneous or automatic features of withdrawal (Lee & Mitchell, 1994). The present study borrows existing theory from several domains in order to specify novel hypotheses regarding the implicit cognitive processes that underlie withdrawal behaviors (i.e., withdrawal *justification mechanisms*). This represents an important first step toward building a dual process theory of withdrawal, which recognizes both controlled and automatic cognitive antecedents. In addition, by studying the implicit biases that contribute to dispositional withdrawal tendencies, this research addresses a critical gap in knowledge of the cognitive basis of personality (Greenwald & Banaji, 1995), and contributes to the marrying of implicit and explicit social cognitive features within the general personality framework (Bing, LeBreton, Davison, Migetz, & James, 2007).

The conditional reasoning measurement system provides an appropriate method for assessing implicit social cognitions and the posited role of such cognitions in withdrawal behaviors. Specifically, implicit biases related to marginalization of withdrawal, revocable conceptions of commitment, and predispositions to inequity, are proposed as novel justification mechanisms that support dispositional withdrawal tendencies. Thus, a major focus of the present investigation was the theoretical specification of these biases and their posited role in different types of withdrawal behaviors. The empirical component of this study comprises an initial effort toward developing and validating a conditional reasoning test to assess these justification mechanisms. Accordingly, a larger set of preliminary items were developed and administered than were ultimately retained. A “revised CRT-W scale was achieved by retaining those items that demonstrated favorable psychometric properties and contributed to the predictive validity of the scale.

Within an employee selection context, the rationale for developing a conditional reasoning test tailored specifically to employee withdrawal behaviors can be summarized in terms of three potential practical advantages, each of which is more fully elaborated in a later section of the proposal: (a) predictive validity for withdrawal criteria, (b) resistance to socially desirable responding, and (c) enhanced user/administrator acceptability in comparison to existing conditional reasoning tests (e.g., the conditional reasoning test of aggression). The present study was focused primarily on the issue of predictive validity. A secondary aim was to test the incremental validity of the CRT-W above existing personality constructs (e.g., conscientiousness and emotional stability) and biodata items (e.g., items assessing social embeddedness and past job changing behavior) that have been found to predict withdrawal criteria in previous research (e.g., Barrick & Zimmerman, 2005; 2009).

To examine the effects of the proposed justification mechanisms on withdrawal behaviors, a predictive validity study was conducted using a student sample in a university classroom setting. Following the administration of the CRT-W, student lateness, absenteeism, and permanent withdrawal from an introductory psychology course was tracked for one semester. Before detailing the specific hypotheses and method for this research, a review of relevant theory and research from the withdrawal domain is first presented. Given the expansive literature on employee withdrawal (Rotundo & Spector, 2010), the review is focused around the following key issues: (a) defining individual withdrawal behaviors conceptually and operationally, (b) highlighting the consequences of employee withdrawal, and by extension, the importance of studying withdrawal criteria from an employee selection perspective, (c) describing models of withdrawal and comparing and contrasting the general withdrawal construct perspective versus a multiple criteria perspective, and (d) summarizing research and theory on the antecedents of withdrawal with a particular focus on the cognitive processes underlying withdrawal.

Following these sections, implicit cognitions are introduced as the focus of the present study and as a relatively new direction for withdrawal research. Drawing from research and theory on social cognition, organizational commitment, and equity and fairness, three cognitive biases are proposed that may contribute to individual differences in withdrawal proneness, and therefore, prove to be useful as predictors in an employee-selection context. Upon building the necessary theoretical background, the CRT-W is introduced as a new scale for assessing implicit cognitions specific to behavioral withdrawal tendencies.

Review of Employee Withdrawal Literature

Specific withdrawal behaviors that result in an employee's physical separation from work are varied. However, most can be grouped into one of three categories or sub-types: lateness, absenteeism, and turnover. These subtypes provide a logical initial framework for categorizing withdrawal behaviors based on *time lost* and the *severity of consequences*. In terms of the typical amount of lost time due to single episodes of these behaviors, lateness subtracts minutes or hours from regular work attendance, absenteeism subtracts day(s), and turnover represents a permanent loss. A similar ordering generally applies to the severity of consequences of individual episodes of these behaviors, with single instances of lateness and absenteeism typically being less disruptive to work than the permanent departure of an employee (Sagie, Koslowsky, &

Hamburger, 2002). This is not to suggest, however, that day-to-day forms of withdrawal are not consequential in specific instances, as well as in the aggregate (Adler & Golan, 1981; Harrison & Martocchio, 1998).

Day-to-day Withdrawal: Lateness and Absenteeism

Lateness and absenteeism share several conceptual similarities and a few differences. Both behaviors result in temporary suspension of expected work duties and can be enacted by the same employee repeatedly over time (Harrison, 2002). Consistent with common usage, lateness (or tardiness) has been traditionally defined in withdrawal research as arrival to work after a scheduled start time (Blau, 1994). Subsequent definitions have expanded lateness to include leaving work early and taking excessive breaks. Absenteeism also represents a temporary separation from work, but usually the amount of time missed is an entire workday. Alternatively, some researchers have described lateness and absenteeism as being on a continuum, whereby lateness becomes a "partial absence" once some threshold of time lost is exceeded. For example, some have used the term *half-day absence* to describe lateness that exceeds four hours. Harrison (2002) suggested a threshold of two hours to demarcate lateness from partial absence, but noted that the issue of setting exact time limits is not a definitional priority and may be more meaningful in relation to specific organizational policies or work context.

One important difference is that, whereas lateness is a universally negative behavior, specific forms of absenteeism are acceptable under certain organizationally defined circumstances. Thus, the excused-unexcused dimension underlying absence-taking behavior is unique in comparison to other withdrawal behaviors. Due to the associated negative consequences, research has focused predominantly on unexcused absences or days missed that were not permissible according to organizational policy. Excluded from this class are approved personal days, sick days, and vacation days. Attempts to further partition the absence criterion has focused on the reasons or motivations underlying absence. In contrast to lateness, which is viewed as a volitional behavior that can be avoided by using proper time management (Blau, 1994), some absenteeism results from non-volitional sources (e.g., certified illness, family death; Sagie, Koslowky, & Hamburger, 2002). The volitional dimension of absence has been labeled *avoidable-unavoidable* (Nicholson, 1977), or alternatively, *voluntary-involuntary* to reflect whether the absence was under the reasonable control of the employee (e.g., resulting from choice behavior) or due to factors beyond the employee's control. Common examples cited as unavoidable (or involuntary) include problems with modes of transportation (e.g., a car accident), emergencies involving family members (e.g., a sick child), and certified sickness (Dalton & Mesch, 1991). Applying this distinction was intended to focus organizational researchers on avoidable or voluntary absences as the specific kind of absence behavior that is of greater theoretical interest. In contrast, unavoidable absences have little opportunity to be explained by psychological theory because they are seen as stemming mainly from random environmental factors (Nicholson, 1977).

Consistent with a management perspective, the avoidable-unavoidable dimension reflects a belief about the types of absence for which employees should be held responsible. This has been criticized by some as representing a blurry distinction, particularly in light of various causes that do not fit neatly into either category (e.g., some health-related concerns; Harrison, 1998). Furthermore, some research has underscored the subjective nature of the avoidable-unavoidable judgment. For example, a study by Payne and Nicholson (1987) found a fundamental attribution pattern, such that employees were more likely to attribute their absences to unavoidable causes, whereas supervisors attributed these same absences to avoidable causes. Finally, adding a volitional component into measures of absence has been criticized as problematic for studies investigating volitional constructs as antecedents of absenteeism (Harrison, 2002; Johns, 1998). Harrison (2002) explained that, “defining a behavior in terms of its cause(s) makes gathering data about those causes meaningless… The cause is already given” (p. 99). From this perspective, any additional variance accounted for as a result of studying volitional antecedents of voluntary (avoidable) absence can be viewed as tautological.

The avoidable-unavoidable dimension also carries through in Blau’s (1994) classic paper, which presented and tested a taxonomy of lateness behavior. However, in Blau’s taxonomy the category *unavoidable lateness*, as well as the categories *increasing chronic* and *stable periodic lateness*, refers to a particular pattern, duration, and frequency of observed lateness behavior rather than attributions about the reasons for particular instances of lateness. Blau’s tripartite model suggests that the relationships between lateness and other withdrawal behaviors (voluntary absenteeism and voluntary turnover) and the antecedents for lateness depend on the type of lateness exhibited.

Increasing chronic lateness is defined as a nonrandom pattern in which the frequency and duration of lateness increases steadily over time. This category is reflective of the progression of withdrawal model (Herzberg, Mausner, Peterson, & Capwell, 1957), which suggests that withdrawal is a temporal process ordered from less severe and infrequent acts (e.g., occasional tardiness) to more severe and increasingly frequent acts (e.g., routine absenteeism) and eventually culminating in turnover. Given this progression, lateness is expected to correlate positively with other forms of withdrawal including voluntary absenteeism and voluntary turnover. The proposed antecedents of increasing chronic lateness are disaffected job attitudes and low organizational commitment (Blau, 1994).

Stable periodic lateness also reflects a nonrandom pattern of lateness but without increasing frequency or duration. In other words, this category represents steady occasional lateness. Rather than being driven by deteriorating work attitudes, stable periodic lateness is thought to reflect a leisure-income trade-off or work-family conflict. A leisure-income trade-off suggests that employees choose occasional lateness or absenteeism because the disincentives for being late are acceptable in exchange for increased leisure time. Alternatively, working adults may accept a certain amount of lateness in order to accommodate the family's needs (e.g., getting children to school, attending a child's after school event, etc.). Blau (1994) suggested that taking these occasional liberties in punctuality should be unrelated to voluntary absenteeism and voluntary turnover.

As alluded to previously, unavoidable lateness is random in its pattern, duration, and frequency. It is caused by unforeseen events, such as traffic, weather, and illness. Consistent with the idea that unavoidable lateness is not due to deteriorating work attitudes or as part of a trade-off between work and life involvements, Blau (1994) posited no relationship between unavoidable lateness and other withdrawal behaviors. Although unavoidable lateness is precipitated by external causes, Blau noted that such instances can be curtailed by appropriate planning and time management and that by failing to do so, “an individual is more susceptible to any type of less controllable antecedent causing lateness” (p. 967).

Using recorded lateness behavior over an 18-month period in two samples, one with bank employees and one with hospital employees, Blau (1994) demonstrated empirical support for the predicted relationships using the proposed taxonomy. Following administration of a survey that contained the measures for work attitudes (i.e., job satisfaction, job involvement, and organizational commitment), leisure-income trade-off, work-family conflict, and scales to assess transportation, weather, and personal illness, lateness records were accrued by both organizations for an 18-month window. Based on this window, the lateness behavior was coded according to pattern, frequency, and time lost, from which individuals were placed into one of the three categories specified by the taxonomy. Results were largely supportive of the taxonomy and resulting predictions across both employee samples. Regarding antecedents, work attitudes were predictive of increasing chronic lateness, leisure-income trade-off and work-family conflict were predictive of steady periodic lateness, and transportation concerns and illness (but not weather) were predictive of unavoidable lateness. Moreover, antecedents were generally not predictive of other categories of lateness behavior (e.g., work attitudes did not predict stable periodic or unavoidable lateness). Partial support for the progression-of-withdrawal hypothesis was obtained in that increasing chronic lateness was associated with higher voluntary absenteeism and leaving work early (but not voluntary turnover).

The study by Blau (1994) has been cited and lauded frequently in the withdrawal literature (Harrison, 2002). The study demonstrates the complexity of lateness behavior and calls into question prior research that has operationalized lateness behavior using a single overall measure. Provided evidence of different patterns of relationships with antecedents as well as other withdrawal behaviors, it is likely that improved prediction of lateness behavior will result from attention to the pattern, frequency, and duration of such behavior. Despite the promise of this initial work, no follow-up studies have used Blau's taxonomy directly or attempted to replicate the original findings. Practical constraints (e.g., lack of highly detailed organizational records) and competing views about withdrawal behavior (e.g., a preference for aggregated rather than spliced measures of withdrawal) are likely reasons why subsequent applications have not been reported (Harrison, 2002).

Consequences of lateness and absenteeism. The outcomes most commonly associated with lateness are negative, whereas a slightly more nuanced perspective emerges for absenteeism. Although individual episodes of employee lateness may not be consequential in all circumstances, the potential for damaging outcomes is apparent. For the offender, being late can result in financial, disciplinary, and social sanctions at work, particularly when lateness becomes habitual. Lateness can also have a cascading effect, such as when arriving late for a meeting results in the loss of an important client, causes a disruption in work processes, produces a backlog of work that results in overtime expenditures, or creates added administrative pressures for calling in substitutes and spending time disciplining or counseling the offending employee (Adler & Golan, 1981; Cascio, 1987; Dalton & Todor, 1993). Furthermore, employee lateness can set a negative precedent for other employees and contribute to a negative organizational culture surrounding attendance issues (Nicholson & Johns, 1985).

In the aggregate, financial costs to the organization can be substantial. Sagie, Birati, and Tziner (2002) estimated the total financial loss attributable to employee withdrawal for a technology firm in Israel. The costs estimated for lateness included direct costs due to the lost time of the late employee, indirect costs associated with the resulting delays in other employees' productivity, and indirect costs associated with an increased propensity for absence (based on the idea that lateness may precipitate absence in a progression of withdrawal framework). The costs estimated for absenteeism included direct costs associated with recovering the lost productivity of the absent employee, indirect costs related with the potential for others' absenteeism rates to increase as a result, and indirect costs associated with an increased propensity for turnover (also based on the progression of withdrawal). Summed with the costs attributed to turnover and withholding effort (a form of psychological withdrawal from work), the total estimated cost of withdrawal was $2.8 million per year, which amounted to 16.5% of the company's before-tax income in 1997. Together, absenteeism (49%) and lateness (4%) accounted for 53% of the total estimated costs of withdrawal.

The financial and non-financial consequences of absenteeism were further described in a review by Harrison and Martocchio (1998). These authors described the short-, mid-, and long-term outcomes associated with absenteeism for the individual, the individual's immediate social environment, and the organization. Some of their preliminary conclusions were that absenteeism leads to more absenteeism for the individual (i.e., a snow-balling effect) and negative co-worker and supervisor responses (short-term), decreased job performance and higher likelihood of turnover for the individual (mid-term), reduced work-unit performance, although this might be moderated by the extent to which group members are interdependent on one another (mid-term), lower pay and fewer promotions for the individual (long-term), decreased customer service (long-term), and poor firm financial performance (long-term).

Interestingly, their analysis, based in part on earlier work by Goodman and Atkin (1984), also identified positive outcomes of absenteeism and in some cases, direct trade-offs where absenteeism is simultaneously beneficial to one party (e.g., the individual or the individual's family) and detrimental to another (e.g., the organization). Most notably, some researchers have described a compensatory process, whereby occasional absence is adaptive and actually prevents turnover that would occur in lieu of adequate time away from work (Dalton & Todor, 1993; Hill & Trist, 1955; Staw & Oldham, 1978). Two studies cited as tentative evidence found that absence led to short-term improvements in affect (Hackett & Bycio, 1996) and cognitive functioning upon returning to work (Totterdell, Spelten, Smith, Barton, & Folkard, 1995). In addition, a study by Staw and Oldham (1978) found that absence was positively related to job performance for employees who were incompatible with their jobs on the basis of mismatched job complexity and personal growth need strength and negatively related to job performance for compatible employees. These authors explained that, "absenteeism may serve a maintenance function for individuals who have difficulty coping with their work roles" (p. 542). From a similar perspective, employees might take occasional unexcused days away from work as a way to restore work-family balance or attend to other non-work commitments (Blau, 1994).

The compensatory model points out that, through stress reduction, absence may occasionally have positive effects for employees. Specific instances of lateness could serve a similar purpose (e.g., sleeping in could prevent an accident that might otherwise occur due to exhaustion; Hill & Trist, 1955). However, it is unlikely that such effects would be sustained over any length of time given the host of negative consequences already described. Harrison and Martocchio (1998) concluded that empirical evidence for a compensatory process was preliminary and that the weight of evidence points toward a progression of withdrawal model, in which withdrawal behaviors are positively intercorrelated and absence increases (not decreases) the probability of turnover (Koslowsky & Dishon-Berkovitz, 2001). Applying Blau's (1994) taxonomy, such a trajectory might only be anticipated when it stems from deteriorating job attitudes and the pattern of withdrawal behavior is increasing and chronic. Alternatively, stable periodic absence may reflect a compensatory function.

Measurement of lateness and absenteeism. Operationally, many researchers have described a strong preference for lateness and absenteeism measures based on organizational records rather than self- or peer-report (e.g., Johns, 1994a; 1994b). Organizational records are generally considered "cleaner" sources of attendance data given a number of sources of inaccuracy underlying subjective judgments about self and others (e.g., self-serving attributions; Harrison & Shaffer, 1994). However, Martocchio and Harrison (1993) noted that many organizational data systems are only partially automated, such that it should not be assumed that attendance data from organizational records are devoid of judgment, or in some cases, intentional distortion. Nonetheless, organizational records remain the most likely source for “objective” data on lateness and absenteeism.

The two most frequently used and researched indexes of absenteeism from organizational records are *frequency* and *time lost* (Harrison, 2002). Frequency is the number of absence spells within a time period, ignoring the length of absence within a spell. In other words, spells can be one or more days long, and only the number of spells – not their duration – is counted. Time lost refers to the total number of days missed within a time period (Chadwick-Jones, Brown, Nicholson, & Sheppard, 1971). Thus, the two indices are positively correlated, though not necessarily perfectly. Test-retest reliabilities for both indices are generally modest (Latham & Pursell, 1977), but appear higher for frequency measures (Chadwick-Jones et al., 1971; Hackett & Guion, 1985). In their meta-analysis of the job satisfaction-absenteeism relationship, Hackett and Guion (1985) reported a mean test-retest reliability of .55 for frequency and .40 for time lost. To be considered alongside these estimates is the possibility that fluctuating levels of absenteeism reflect rank order changes in individuals’ underlying withdrawal propensities over time, such that low test-retest reliabilities reflect both true and error variance components (Latham & Pursell, 1977). For example, using survival analysis methods, Harrison and Hulin (1989) demonstrated how unique temporal, situational, and historical factors influence individuals’ absence behavior over time in addition to a stable behavioral component. Other measurement issues include low base rates and unreliability in the organizational record keeping process (Harrison, 2002).

Availability and practical constraints have been the primary determinants of the measures used to assess lateness. In comparison to absenteeism, fewer examinations of the psychometric properties of various measures of lateness have been undertaken (Harrison, 2002). Harrison (2002) noted that the imperative for using organizational records is not as strong in the lateness literature because of reduced availability. He went on to recommend that, if available, either frequency or time lost measures may be appropriate, but that more informed decisions await future methodological research. The previously described work by Blau (1994) suggests that the antecedents of lateness and the interrelationships between lateness and other physical withdrawal behaviors depend on the pattern, frequency, and duration of lateness behavior demonstrated by an individual over time. Accordingly, optimal objective (or self-report) measures would allow one to classify lateness behavior according to Blau's typology (Blau, 2002).

Blau (2002) noted that, in lieu of detailed organizational records or insufficient sample sizes, self-report measures could be necessary. A study by Koslowsky and Dishon-Berkovitz (2001) examined the comparability of self-report measures of lateness and objective personnel files in a sample of white-collar employees in Israel. Consistent with the organization's policy, lateness was defined as arriving at work at least 1-minute after the scheduled start time. Objective data in personnel files were gathered automatically by a time-punch clock over a 10-month period prior to collection of the self-report lateness measure, as well as during a 3-month follow-up period afterwards. To minimize demands on employees' memory, the self-report measure was worded to ask about lateness behavior over the preceding 2-week period.

Consistent with prior research on the comparability of self-report and objective absenteeism measures (e.g., Johns, 1994b), results indicated that employees systematically under-reported the average number of times per week that they were late. A moderate positive correlation was noted between the self-report measure of lateness and the pre (r = .31) and post objective measures (r = .32). Together, these results suggest that self-report and objective indices of lateness are moderately intercorrelated and should not be viewed as interchangeable. However, the different time periods used for self-report and objective measures prevents stronger conclusions. Based on these findings, Koslowsky and Dishon-Berkovitz (2001) and others (e.g., Blau, 2002), have suggested methods for reducing socially desirable responding and improving the accuracy of self-report lateness measures (e.g., employ a counting rather than estimation procedure). Finally, it is interesting to note that the objective lateness measure showed strong correspondence over time, as indicated by test-retest reliability across the two measurement periods equal to .82. This provides preliminary evidence for the superior test-retest reliability of lateness in comparison to absenteeism behavior, and is generally consistent with the view that lateness is reflective of personal characteristics of employees (i.e., a volitional perspective) more so than situational determinants.

Permanent Withdrawal: Turnover

In contrast to absenteeism and lateness, turnover is a single behavioral act that results in “the termination of formal relations between an employee and an organization” (p. 54, Krausz, 2002). Thus, turnover is a permanent form of withdrawal. Turnover is a behavioral construct involving the act of termination and needs to be distinguished from cognitive constructs such as withdrawal cognitions (i.e., thinking about quitting) and turnover intentions (i.e., formulating a plan for quitting) (Harrison, 2002). Like absenteeism, turnover is a diverse behavior warranting certain qualifications.

The most basic distinction to be made, with direct implications for the antecedents of turnover, is the difference between voluntary and involuntary types of turnover. Voluntary turnover is defined as a choice behavior carried out by the employee, whereas involuntary turnover is a choice behavior carried out by the organization (or an authority figure within the organization). Although both result in termination, the behavioral construct is different for the act of quitting (i.e., voluntary) versus the act of being dismissed, laid off, or forced into retirement (i.e., involuntary). Possible reasons for voluntary turnover include disaffection with a current work role or organization, desire to change careers, changing family or personal circumstances (e.g., a spouse takes a job in a different location), a more attractive job opportunity elsewhere, and so on. Possible reasons for involuntary turnover include poor job performance, violations to company policy, and financial necessity for the firm (e.g., downsizing or permanent layoffs).

Other reasons for employee turnover can be difficult to classify as purely voluntary or involuntary (Campion, 1991; Krausz, 2002). Krausz (2002) describes an ambiguous situation in which an employee is dismissed due to not accepting a job transfer. Early retirement offers and buyouts are similarly difficult to disentangle because they involve a negotiated departure rather than a definitive rejection by the employee or organization. On the other hand, routine job transfers are generally not considered turnover because the employee remains with the organization, albeit at a different geographical location, and retirement is usually considered a separate special type of turnover (Harrison, 2002). Other examples for which there could be disagreement include turnover due to pregnancy and health-related concerns, neither of which necessarily connotes an explicit choice (Campion, 1991). Due to the murkiness of a simple dichotomy, Campion (1991) suggested a continuum to allow a moderate classification for instances of mutual agreement between organizations and employees.

Although somewhat crude, the function served by the traditional voluntary-involuntary distinction is to identify or reduce sources of heterogeneity when studying employee withdrawal behaviors. Consistent with the focus of withdrawal research, the vast majority of studies have been interested in voluntary turnover as a motivated choice behavior. In other words, the development and testing of theories of motivated choice behavior conceptualize the appropriate outcome variable as voluntary and not involuntary turnover, and indeed, empirical findings indicate that the two are differentially predicted by antecedents. An additional reason to focus on voluntary rather than involuntary turnover has to do with the differential consequences of turnover for organizations, and by extension, the types of turnover that organizations are most interested in preventing. However, in order to fully elaborate these consequences, categorization along two additional dimensions is needed: *avoidable-unavoidable* and *functional-dysfunctional* (Campion, 1991). These dimensions are operationally valuable and have measurement implications in certain research contexts, such as when organizational consequences are the focus of investigation (Harrison, 2002).

Attributions about avoidable-unavoidable employee turnover are somewhat different than was previously described in reference to absenteeism. Here, the attribution is focused on what, if anything, the *organization* (not the employee) could have done to prevent the turnover (Harrison, 2002). Avoidable voluntary turnover suggests that the organization could have prevented the loss, such as by increasing pay, offering promotion, improving management-employee relationships, or initiating flexible work policies. Unavoidable voluntary turnover suggests that the reasons for quitting could not have been prevented by the organization, such as when a geographical move is necessary for the employee's family or the employee decides to change careers (Campion, 1991). The implication is that knowledge about avoidable voluntary turnover and underlying reasons carries unique change implications for organizations.

The consequences of avoidable voluntary turnover likely depend on organizations' case-by-case evaluation of the employees who leave voluntarily. Regardless of whether an organization *can* prevent a loss (i.e., avoidable-unavoidable), only those losses that the organization would *want* to prevent are seen as problematic. Along these lines, Dalton and colleagues (Dalton, Krackhardt, & Porter, 1981; Dalton & Todor, 1979; Dalton & Todor, 1982; Dalton, Todor, & Krackhardt, 1982) differentiated *functional turnover* from the traditional *dysfunctional* view, to underscore the possibility that turnover can actually benefit organizations. Functional turnover comprises the loss of employees for whom the organization holds a negative evaluation. Campion (1991) assessed functionality of turnover by asking supervisors about the quitting employee's job performance (e.g., Would you re-hire the employee who left?) and the difficulty of replacement (e.g., In general, how easy would it be to find someone who would do as good a job as the employee who left?). Borrowing from utility models of turnover (e.g., Cascio, 1987), functionality is highest when poor employees leave and are replaced by good employees in a cost effective manner, and functionality is lowest when good employees leave and are difficult or even impossible to replace.

Consequences of turnover. Turnover is financially costly to organizations when it results in the loss of good employees who then need to be replaced (Mobley, 1982). As a consequence, the organization incurs the costs associated with recruiting, selecting, socializing, and training a new employee to fill the previous employee's position, as well as any lost productivity in the interim. The total estimated cost of replacing a lost employee can vary widely depending on a range of factors, including the industry and job type, the labor market, whether the position can be filled by someone already in the organization, and so on (Staw, 1980). As one example, the cost of replacing a single registered nurse has been estimated at between $10,000 and $60,000 depending on the nurse specialty (Hayes, O'Brien-Pallas, Duffield, Shamian, Buchan, Hughes, et al. 2006). Organization-level analyses demonstrate the seriousness of this problem when the cost of individual turnover episodes is aggregated. Sagie et al. (2002) estimated that turnover costs accounted for nearly one-third of financial losses due to employee withdrawal in a single organization.

Other consequences can be more difficult to estimate, such as opportunity losses resulting from the unrealized productivity of lost employees and the productivity of replacement employees in comparison to lost employees (Boudreau & Berger, 1985). Taking all factors into consideration, it should be apparent that not all instances of turnover result in a net loss for the organization, such as when the productivity gained by the new employee offsets the costs associated with loss and replacement (i.e., functional turnover).

From a more macro perspective, the attraction-selection-attrition (ASA) model (Schneider, 1987; Schneider, Goldstein, & Smith, 1995) describes *some* regularly occurring turnover as part of a natural attrition process in which poor-fitting employees (i.e., low person-organization or *P-O* fit) exit the organization and leave behind a gradually more cohesive organizational culture over time. On the other hand, as the rate of attrition increases beyond “healthy” levels, the positive effects of turnover are likely counterbalanced by financial losses and accelerated organizational homogeneity, a potentially maladaptive consequence of ASA processes. Thus, although turnover should not be regarded as universally negative, positive implications are likely evident only when specific instances represent functional turnover and when overall base rates are relatively low.

In terms of non-financial outcomes, turnover has been described as having mainly negative effects at the individual, work-unit, and organization level, although the potential for moderators has been recognized (Staw, 1980; Hausknecht, Trevor, & Howard, 2009). For example, turnover may have negative consequences for the morale of “stayers” or those who remain with the organization, such as reduced job satisfaction and organizational commitment (Staw, 1980). Due to the disruption of work-units, turnover has been linked to decreased quality of group communication, reduced commitment (Mueller & Price, 1989), and poorer customer service (Hausknecht et al., 2009; Kacmar, Andrews, Van Rooy, Steilberg, & Cerrone, 2006). Consistent with these findings, Krackhardt and Porter (1986) found that turnover was increased within the work groups of recently turned fast-food employees, and in particular, those whose social networks were most similar to that of the former employee. These results were described as evidence for a “snowballing” effect of turnover within work groups. As with absenteeism, the negative effects of turnover on work-units likely depend on the degree to which individuals are interdependent and the centrality of the lost individual within this interdependence.

Finally, at the organization level, some authors have described the implications of a *turnover culture*, which is characterized as the “set of shared understandings about the legitimacy of leaving an organization” (p. 378, Deery & Shaw, 1997). In essence, organizations with negative turnover cultures reinforce and perpetuate turnover as an expected employee behavior. Accordingly, turnover culture has been studied primarily as an antecedent of turnover intentions and behavior (e.g., Iverson & Deery, 1997). However, it is likely that such cultures emerge in concert with, or as a consequence of, historically high turnover rates in the organization. For example, Iverson and Deery (1997) posited that turnover cultures result from strong industry norms that lead to frequent job changing. For this reason, it seems plausible that negative turnover cultures are both an antecedent and consequence of high turnover environments.

Measurement of turnover. With a few exceptions, previous studies have extracted individual-level turnover data from organizational records and archives (Harrison, 2002). Campion (1991) described a variety of operational difficulties with such measures and ultimately advocated a combination of archival data and follow-up questionnaires with supervisors and employees. Often organizational records capture information about the reason(s) employees leave in addition to the occurrence and timing of turnover events. Harrison (2002) suggests that organizations are motivated to maintain accurate turnover data due to the high financial costs associated with handling the turnover process (e.g., administration of severance pay). However, such information can be inaccurate if employees respond to exit interviews dishonestly, if data systems force administrators to enter a single reason for turnover when multiple reasons exist, or if the categories applied are too broad to be very useful (e.g., “personal reasons”).

Campion (1991) compared organizational records with employee and supervisor provided reasons for turnover and found that data sources yielded corresponding information approximately 70% of the time. This indicates that overreliance on any single source, such as organizational records, carries the potential for some inaccuracy. Furthermore, researchers’ attempts to apply appropriate qualifiers to individual turnover data points (e.g., voluntary-involuntary) can be inaccurate when based on distorted or erroneous organizational records, or such efforts can be impossible when organizational records do not provide sufficient detail. As an alternative, Campion developed continuous measures to assess the voluntariness, avoidability, functionality, and utility of turnover behavior based on employee and supervisor self-report. Such measures could be used instead of, or in addition to, organizational archives. A major disadvantage for using these, as well as other survey based approaches to assessing turnover, is that they require follow-up contact with employees who have left an organization. In addition, such instruments are reliant on individuals’ accurately remembering the circumstances surrounding turnover and assume that such events are not subject to different interpretations over time.

From a psychometric standpoint, analyzing turnover data can be problematic when the proportion of stayers greatly exceeds leavers within a given study window. A low base rate (e.g., very few leavers) can severely attenuate the ability of a dichotomous turnover variable to correlate with other variables (Harrison, 2002). More generally, this raises an issue of research design and the need to study turnover as a longitudinal process rather than as an either-or event. Traditional research methodologies and analysis strategies produce an arbitrary turnover rate that is a function of the length of the study window (i.e., the follow-up period after initial measures were collected). Logistic regression using a dichotomous turnover criterion does not differentiate between employees who leave immediately and those who leave just before the study is discontinued. In this way, traditional methods are not equipped to differentiate between early- and late-leavers, despite that organizations likely view these cases differently and theoretical withdrawal models may offer different understandings. In response to these shortcomings, several authors have called for use of survival analysis analytic strategies for modeling turnover as a longitudinal phenomenon (Morita, Lee, & Mowday, 1989; Singer & Willett, 1991; Somers & Birnbaum, 1999). Application of survival analysis requires more detailed turnover data, including information about event timing (e.g., tenure) in addition to event occurrence.

Summary. The literature on absenteeism and turnover is voluminous and continuing to grow (Rotundo & Spector, 2010). Lateness has been the focus of comparatively less systematic investigation, but is important nonetheless, and particularly so in light of a progression-of-withdrawal model (Blau, 2002). At present, each behavior and its consequences are understood in terms of various taxonomies and “qualifiers,” contributing to an overall nuanced picture of withdrawal. Although such nuance was intended mainly to improve the predictive capacity of various withdrawal models (e.g., models of volitional choice), there are lingering questions as to whether partitioning the withdrawal criterion space has improved or detracted from the utility of existing theories (Harrison, 2002). Having explored much of this diversity in the preceding section, what follows is a description of withdrawal models and competing perspectives on a unitary versus partitioned withdrawal construct.

Models of Withdrawal

There are differing perspectives in the literature as to whether lateness, absenteeism, and turnover represent a single underlying withdrawal construct versus distinct behavioral dimensions with unique antecedents and consequences (Porter & Steers, 1973). As noted by Harrison (2002), specific withdrawal behaviors have long been measured separately (e.g., within organizations’ record keeping) and accepted as distinct behavioral indices or variables. However, some commonality is implied simply by use of the term *withdrawal* as a label for these behaviors as a collective (Clegg, 1983; Mobley, 1982). Operationally, the appropriateness of a single-construct versus a multiple-criteria conceptualization likely depends on the level of specificity called for in a particular research context (Harrison, 2002).

Taking a macro orientation (i.e., low specificity), withdrawal behaviors can be considered a special class of counterproductive work behaviors (CWBs), and more broadly still, CWBs can be viewed as a special class of job performance behaviors (Rotundo & Spector, 2010). Recent taxonomies of CWBs group withdrawal behaviors together as a single dimension (Sackett & Devore, 2001; Spector et al., 2006). Spector et al. demonstrated that the pattern of relationships with a host of antecedent constructs (e.g., job satisfaction, justice perceptions) was different for withdrawal than for other types of CWB. These authors described withdrawal as, “contrasting with these other forms of behavior [other CWBs] because it is an attempt to avoid or escape a situation” (p. 450). However, turnover was not included as a type of withdrawal behavior in this study because the focus was on CWBs that could be enacted on repeated occasions (e.g., daily). Nonetheless, a CWB perspective highlights the escape-avoidance feature that is common to withdrawal behaviors and which distinguishes withdrawal from other damaging employee behaviors.

Moving to a higher level of specificity (i.e., the withdrawal research domain), some researchers, most notably Hanisch, Hulin, and colleagues (Hanisch, 1995; Hanisch & Hulin, 1990; Hanisch & Hulin, 1991; Hanisch, Hulin, & Roznowski, 1998), have argued for a multiple-behavior construct. These authors have defined a general withdrawal construct as:

> composed of a variety of acts, or surrogate intentions, that reflect both the negativity of the precipitating job attitudes and the target of these negative job attitudes (p. 111, Hanisch & Hulin, 1991).

From this perspective, individual withdrawal behaviors are indicators of a single underlying propensity for withdrawal. This propensity comprises negative attitudinal (e.g., job dissatisfaction), cognitive (e.g., turnover intentions) and behavioral features, and as such, is broader than the strict behavioral focus of the multiple criteria perspective (Harrison, 2002). Even within this general construct framework, a distinction is apparent regarding the specific manifestations of withdrawal. Employees can seek to minimize time spent on task performance without affecting their overall employment status (i.e., *work withdrawal*) or seek to leave the organization (i.e., *job withdrawal*). The target of an individual's dissatisfaction – the nature of the work or the organization – is related to whether a work or job withdrawal response is more likely. Note that this distinction is not altogether different from a taxonomic multiple-criteria perspective, in which lateness and absenteeism are viewed as interrelated (i.e., day-to-day forms of withdrawal) and somewhat dissimilar from turnover, which reflects a permanent separation (Harrison, 2002).

The main arguments for a general withdrawal construct are improved predictability by general work attitudes and improved generalizibility (Harrison, 2002). Aggregating behavioral episodes of withdrawal should yield a composite criterion with improved predictability, particularly for those antecedents that are equally broadly defined, such as general work attitudes (Hanisch et al., 1998). This is consistent with both theoretical and statistical rationales for forming composite criteria (Schmitt et al., 2003). From a theoretical standpoint, predictors and criteria should be matched in scope (Fisher, 1980), such that general work attitudes are expected to be more predictive of a general withdrawal construct than individual withdrawal behaviors. Hanisch et al. (1998) review several studies comparing aggregated and individual withdrawal criteria, and which demonstrate support for this basic pattern. From a statistical standpoint, individual withdrawal behaviors (e.g., turnover) can suffer from low stability, skewed distributions, and low base rates (Johns, 1998). Aggregating across behavioral indices has the potential to ameliorate some of these disadvantages and yield a more predictable criterion (Hanisch et al., 1998). Finally, the promise of a more abstract, and therefore, portable theory of withdrawal is attractive from a generalizability standpoint (Harrison, 2002; Schmitt et al., 2003).

Arguments against the general withdrawal construct focus on the complexity of individual withdrawal behaviors and heterogeneity that results from their combination (Blau, 1998; Johns, 1998). Based on the taxonomic diversity of individual withdrawal behaviors and the differential prediction of antecedents of these behaviors, it is conceptually unclear what a fully aggregated construct represents and questionable as to why this should result in improved predictability (Blau, 1998). Explanations focus on improved reliability and the appropriateness of widening the scope of criteria to reflect broadly defined predictors. Along these lines, Johns (1998) suggests that the reasons offered by Hanisch and colleagues for an aggregated construct are decidedly psychometric. In contrast, less theoretical justification is offered, including a clearly specified nomological network of proposed antecedents and consequences.

An additional concern is the conflating of the withdrawal criterion with proposed antecedents of withdrawal, including job attitudes and withdrawal intentions (Martocchio & Harrison, 1993). John's (1998) review indicated that, in addition to behaviors, previous measures captured information about employees' intentions, expectations, desires, and feelings. As a consequence, the current specification of the construct appears tautological or contaminated with variance from the predictor space, which might account for the increased predictability. In balance, a possible counter-argument is that a disaggregated approach has achieved a similar purpose (e.g., better predictability by volitional constructs; Campion, 1991) by splintering individual withdrawal behaviors along several dimensions. Offering a similar point, Harrison

(2002) explains that, as a result of contamination, individual behaviors are only considered indicators of withdrawal if dissatisfaction is inferred as the cause. In addition to the tautological consequences, it is unclear how less intentional forms of withdrawal would fit into this framework. In summary of the opposition voiced by Johns, Blau, and others (e.g., Martocchio & Harrison, 1993; Harrison, 2002), as currently constructed, replacement of the taxonomic view of withdrawal with a general withdrawal construct seems preliminary. Nevertheless, the difficulties associated with measuring individual behaviors points to the need for further development of construct-oriented perspectives of withdrawal (Schmitt et al., 2003).

Importantly, the potential for a general withdrawal construct depends on the existence of positive covariance among individual withdrawal behaviors (Harrison, 2002). That is, statistical aggregation presupposes that individual components are positively correlated. Somewhat controversially (e.g., Blau, 1998; Johns, 1998), Hanisch et al. (1998) cite high internal consistency among self-report measures of job and work withdrawal as evidence of strong positive covariance. Recall that these authors' general withdrawal construct subsumes behavioral and attitudinal features of withdrawal. For this reason, organizational records of lateness, absenteeism, and turnover alone are deficient indices of withdrawal because each can be enacted without the implied job dissatisfaction. Accordingly, Hanisch and colleagues have developed self-report measures of job and work withdrawal to incorporate attitudinal as well as behavioral indicators, although both are subjective perceptions of withdrawal (Harrison, 2002). Thus, common method bias due to sampling the perceptions of single individuals at single time points is a problematic alternative explanation for the high internal consistency reliabilities reported, or at least contributes to an upward bias for these estimates (Blau, 1998; Harrison, 2002; Johns, 1998).

Construct and measurement issues temporarily aside, the degree of positive covariance among withdrawal *behaviors* remains a critical issue for the viability of a general withdrawal construct. Overall, empirical research supports a moderate positive correlation among different physical forms of separation from work (Clegg, 1983; Koslowsky et al., 2001). Mitra, Jenkins, and Gupta (1992) conducted a meta-analysis of 17 studies (33 correlations) of the relationship between absenteeism and turnover with samples comprising employee from a variety of manufacturing and nonmanufacturing settings. They reported an uncorrected mean correlation of .23 and a corrected mean correlation of .33 (adjusted for unreliability in absenteeism and unequal sample sizes) between indices of absenteeism and turnover. The finding that employees who are absent more frequently are also more likely to turnover is consistent with several prior and subsequent reviews (e.g., Hanisch, 2002; Muchinsky, 1977). At the same time, sufficient non-artifactual effect heterogeneity was observed, suggesting the possibility of moderators. Moderation tests indicated that the type of absenteeism measure (i.e., frequency versus time lost) did not moderate the strength of the absenteeism-turnover relationship. Unfortunately, a similar analysis based on the type of turnover measure (e.g., voluntary versus involuntary) was not possible. An additional moderation test indicated that positive covariation among absenteeism and turnover was somewhat stronger for studies with durations of less than 12-months (mean corrected r = .37) in comparison to studies longer than 12-months (mean corrected r = .29). Supplemental analyses suggested that job market conditions (e.g., unemployment rates) might mitigate the strength of the absenteeism-turnover relationship over more extended time periods. This is generally consistent with Harrison and Martocchio's (1998) time-based view of withdrawal, which suggests that the antecedents of withdrawal depend on the time window of interest.

A meta-analysis by Koslowsky, Sagie, Krausz, and Singer (1997) provided mixed support for the positive covariance among withdrawal behaviors. In total, these authors meta-analyzed 118 correlations drawn from 30 studies examining the correlates of employee lateness. Correlates included other withdrawal behaviors and intentions (i.e., absenteeism, turnover, and turnover intentions), as well as work attitudes (e.g., job satisfaction and organizational commitment), demographics, and job performance. Overall, absenteeism and turnover showed stronger relationships with employee lateness than the other correlates investigated, with overall mean corrected correlations (adjusted for sampling error and unreliability) of .40 for the lateness-absenteeism relationship and .27 for the lateness-turnover relationship. However, the magnitude of intercorrelations differed by the type of absenteeism and turnover measure used. Specifically, correlations with lateness were higher for voluntary absenteeism than for involuntary (corrected mean *r*s of .41 versus .16, respectively), as well as for turnover intent (e.g., based on a turnover intention questionnaire) in comparison to actual turnover behavior (corrected mean *r*s of .46 versus .07, respectively). Furthermore, it is not specified how frequently absenteeism or lateness measures were based on self-report versus organizational records (i.e., objective data); however, Harrison (2002) suggested that approximately 20% of the studies included in the meta-analysis by Koslwosky et al. used self-report, such that estimates might be somewhat inflated due to a common method bias. Nonetheless, the overall pattern of results suggests a moderate positive relationship between lateness and absenteeism and a somewhat weaker positive relationship between lateness and turnover. Koslowsky et al. (1997) interpreted this pattern of relationships as tentative support for the progression-of-withdrawal model.

Four models have been advanced to describe different possible relationships between lateness, absenteeism, and turnover (Rosse, 1988). An independence model suggests that individual withdrawal behaviors are uncorrelated; the occurrence of one (e.g., lateness) has no effect on the probability that another will occur (e.g., absence). A compensatory model suggests that individual behaviors are negatively correlated, such that the occurrence of one (e.g., absence) depresses the likelihood of another (e.g., turnover). Neither model has fared well in light of the consummate finding that withdrawal behaviors tend to be positively (albeit modestly) intercorrelated. As previously described, a compensatory process may be an accurate description of withdrawal as a coping mechanism under a constrained set of circumstances (e.g., in short-term and in high-stress situations). However, the bulk of evidence is inconsistent with the compensatory view, and indicates that absenteeism, lateness, and turnover tend to operate in the same direction (Koslowsky et al., 1997; Mitra et al., 1992).

Two alternative models, the spillover model and the progression-of-withdrawal model, are consistent with positive interrelationships among withdrawal behaviors. The spillover model accounts for positive interrelationships among specific withdrawal behaviors by suggesting that the stable individual differences (i.e., “withdrawal proneness”) that contribute to one behavior are likely to contribute to the others. That is, individuals who are prone to being late are also prone to being absent, and so on. Moreover, to the extent that withdrawal behaviors are in response to poor work attitudes, it is likely that such attitudes will manifest in several different forms of withdrawal (Beehr & Gupta, 1978). Alternatively, a progression of withdrawal model employs a similar rationale but also describes a hierarchical and temporal ordering of withdrawal behaviors, such that lateness leads to absenteeism and absenteeism leads to turnover (Rosse, 1988). Evidence bearing on the dispositional component underlying withdrawal is reviewed briefly below.

Withdrawal proneness. Seeking to understand the personal factors that underlie withdrawal is by no means a new objective for researchers (Steers & Rhodes, 1978). Early studies by Taylor (1968) and Froggatt (1970) demonstrated stable individual patterns among industrial workers' use of sick days. Other studies have found that past withdrawal behaviors were a strong predictor of future withdrawal behaviors (e.g., Barrick & Zimmerman, 2005, 2009; Judge et al., 1997), including when temporal stability is examined across situations (Harrison & Price, 2003). Finally, different types of withdrawal behaviors are correlated within individuals, such that employees who are absent more often also have a higher probability of quitting (Griffeth et al., 2000). Together, these findings suggest that withdrawal tendencies differ across people and may be relatively stable within persons across employment contexts.

Additional evidence for this perspective is gained by studies that link stable individual characteristics (e.g., personality traits) to withdrawal behaviors, suggesting a dispositional source of withdrawal. Three traits from the Five Factor Model (FFM) in particular have received support. Studies by Barrick and Zimmerman (2005, 2009) demonstrated that conscientiousness and emotional stability measured pre-hire were positively correlated with turnover up to two years post hire among employees in finance. Judge et al. (1997) reported a positive relationship between extraversion and absence and a negative relationship between conscientiousness and absence among university employees. Taylor (1968) studied medical absences among refinery workers and found that neuroticism was positively related to length of absence, and extraversion was positively related to frequency of absence. Other stable individual constructs that have been linked to withdrawal behaviors include positive and negative affectivity (Iverson & Deery, 2001), confidence (Barrick & Zimmerman, 2005), impulsivity (Porter & Steers, 1973), self-monitoring (Harrison & Price, 2003), and anxiety (Bernardin, 1977).

Cognitive antecedents. The dominant cognitive process models of withdrawal stem from March and Simon's (1958) early conceptual model, which emphasized employees' evaluation of the job (most commonly operationalized as job satisfaction) and the perceived availability of alternative jobs. A subsequent influential model by Mobley (1977) built on the March and Simon model by specifying a temporal process of withdrawal that involves several stages. The Mobley model proposed that job dissatisfaction leads to withdrawal cognitions, which in turn leads to an evaluation of the expected utility of quitting. If the expected utility is high (i.e., high anticipated benefits and low anticipated costs associated with quitting), a search is begun, and job alternatives are compared with the current job along several dimensions. If the alternative jobs compare favorably, an intention to quit is formed, which can ultimately lead to the turnover event. In their review, Lee and Mitchell (1994) concluded that research evidence provides mixed support for the sequential cognitive process laid out in the Mobley model. However, a consummate finding from the research stimulated by this model is that turnover intentions tend to be a stronger predictor of actual withdrawal behaviors than job attitudes, although effect sizes are generally modest for both predictive relationships (Lee & Mitchell, 1994).

With some exceptions, subsequent models have mainly added mediators or moderators to the basic sequential process described by Mobley (1977). For example, Steers and Mowday (1980) incorporated met expectations as an antecedent of job attitudes and broadened job attitudes to include organizational commitment and job investment in addition to job dissatisfaction. Hulin, Roznowski, and Hachiya (1985) incorporated economic factors (e.g., unemployment rates) and the career orientation of employees, as key determinants of the relative importance of job attitudes versus perceived alternatives in shaping withdrawal intentions.

Process models to date have described withdrawal as preceded by a series of explicit cognitive evaluations (e.g., of the job, alternatives, and more complex comparative analyses). Lee and Mitchell (1994) presented an alternative perspective based on the idea that the judgment and decision-making processes used by people rarely reflect a systematic, extensive, or rational approach. Following from work on image theory, these authors proposed several decision pathways that precede withdrawal and reflect varying degrees of automaticity versus mental deliberation. Beach (1990; as cited by Lee and Mitchell) described decisions as rarely involving extensive searches and consideration of alternatives and as motivated by maximizing utility in only a subset of cases. More commonly, choices are somewhat rare and behavior is generally pre-programmed. Consistent with this perspective and in contrast to previous withdrawal models, Lee and Mitchell's model describes some turnover decisions as occurring relatively automatically, such as when a "shock" occurs that connotes value incongruence between the person and the organization and results in the initiation of a script-driven decision to leave (i.e., if X were to occur, leaving would be the appropriate response).

Research evidence bearing on the decision pathways specified by Lee and Mitchell (1994) is still preliminary. However, their "unfolding" model of turnover makes an important contribution to the withdrawal literature by describing how automatic and unconscious cognitions can affect the withdrawal process. These authors pointed out a major limitation of prior models, which have supplied an overly rational and deliberate account of withdrawal, particularly with respect to absenteeism and tardiness, which often have a spontaneous or impulsive component (Porter & Steers, 1973). As Lee and Mitchell described, it is still the case that relatively little attention has been paid to the automatic cognitive processes that guide withdrawal decisions and behaviors. The present study proposes that implicit cognitive biases provide an explanation of how scripted behaviors are enacted automatically in response to organizational events. Westen (1998) and others (e.g., Greenwald & Banaji, 1995) have advanced strong arguments for the importance of implicit cognitions by reviewing research that indicates that, whereas conscious processes (e.g., conscious attitudes and motivations) guide the limited subset of behaviors that are consciously chosen, unconscious processes guide a much wider range of behaviors that are not consciously chosen and which occur "over the long run" (p. 338).

Implicit social cognitions. Indirect evidence for this proposition can be drawn from parallel research on the role of implicit cognition in counterproductive work behaviors (including withdrawal) and behavioral aggression. Research on conditional reasoning testing provides the measurement framework for assessing specific implicit biases in the context of pre-employment selection testing. Next, a brief introduction to the concept of conditional reasoning and the measurement system used by conditional reasoning tests is provided. Thereafter, extant research is reviewed demonstrating the conditional reasoning test of aggression (CRT-A) as a useful parallel for conceptualizing the implicit biases that may shape withdrawal behaviors and as evidence of the validity of the conditional reasoning approach for predicting negative workplace behaviors.

Conditional Reasoning

Conditional reasoning asserts a social cognitive explanation for personality that is different from the traditional trait perspective. Whereas trait perspectives are focused mainly at the level of behavior, social cognition provides one perspective for understanding the motives behind behaviors. Specifically, social cognition describes the patterns of thinking individuals use to interpret their social environments and which serve as a precursor to stable patterns of behavior in response to those environments. From this perspective, behavior is a combined result of expectations, perceptions, attributional sense making, cognitive re-interpretation of events (e.g., retrospective understanding), and so on. In total, these processes might be considered features of an individual's *reasoning* about his or her social environment. Central to the concept of conditional reasoning is the idea that people are neither unbiased nor passive observers of their social environments. Instead, reasoning is shaped by a host of factors, which are intrinsic and potentially idiosyncratic to the individual. In other words, the individual uses a personalized basis for judging, understanding, and responding to his or her surroundings. This is the essence of the label *conditional reasoning*, which describes the process by which reasoning is contingent on these factors (James & Mazerolle, 2002).

At various times, reasoning can be affected by idiosyncratic social cognition above or below conscious awareness. The label *implicit* is used to refer to those sources of biased reasoning that are thought to generally operate below an individual's conscious awareness. Unabated, social responding may be shaped somewhat automatically (or heuristically) by certain implicit biases. As a result, the individual may not readily understand or acknowledge the contribution that such factors play in their judgment and behavior. Note that, in this context, the term "bias" does not necessarily connote "bad" or negatively valenced motives. The bias acts to shape perception and reasoning in a manner that is consistent with implicit motives whether those motives are prosocial (or adaptive) or anti-social (or maladaptive).

Implicit cognitive biases serve both retrospective and prospective functions within a person. Retrospectively, biased cognition can serve an ego-protective function by offering a re-interpretation of events to block negative self-information. For example, self-serving attributions focus ones attention on the role of external factors when accounting for a personal failure (Johns, 1994) and cognitive dissonance describes a process of reconciling seemingly inconsistent information about self-concept and behavior (Festinger & Carlsmith, 1959). Prospectively, implicit cognitive biases act as information filters by shaping future expectations, focusing selective attention, framing perception of social information, and priming the likelihood of certain responses. As an example, individuals who use aggressive schemata to understand social interactions might aggress in anticipation of hostility from others or may seek information that justifies an aggressive response. To the person with an aggressive motive, relationships can appear as power struggles, challenges from others can be seen as personal attacks, and so on (James et al., 2004; 2005). Whether serving retrospective or prospective purposes, implicit biases have the sum effect of “enhancing the rational appeal of motive-based or dispositional behavior” (James & Mazerolle, 2002, p. 92). In other words, people generally want to believe that their actions and choices are reasonable and even logical, as opposed to irrational (James, 1998).

James et al. (2004) referred to the cognitive biases that are held to rationalize specific behavioral tendencies as justification mechanisms (JMs). JMs are conceptually linked to personality because they contribute to stable behavioral tendencies and are held to varying degrees across individuals. James (1998) described JMs as developing in response to specific dispositional tendencies such that the presence of various JMs in one’s social reasoning is expected as a consequence of different personalities. More recent work has emphasized that JMs might contribute to implicit aspects of personality that are thought to be distinct from the explicit features assessed by overt self-report measures (Frost, Ko, & James, 2007). JMs are conceptualized as implicit biases because individuals are generally unaware of their effect on reasoning and behavior. For example, a person who holds a hostile attribution bias – an aggressive JM – might respond aversively to negative feedback from a manager, having viewed the feedback as the manager’s attempt to demean him and establish superiority. Furthermore, the same person might be surprised to learn that others would perceive the same feedback as merely constructive criticism (James & LeBreton, 2011). Although it might be possible to bring JMs into an individual's conscious awareness, more commonly they are believed to operate at the level of automatic social cognitions of which the individual is generally unaware (James & Mazerolle, 2002).

To date, published research on conditional reasoning has focused on the implicit biases underlying three constructs: aggression, achievement motivation, and fear of failure. Drawing primarily from social cognition research and theory, James (1998) introduced six specific JMs that contribute to (or are a consequence of) dispositional aggression, six JMs for dispositional achievement motivation, and eight JMs for fear of failure. Whereas the JMs for achievement motivation and fear of failure were specified in direct opposition to one another (i.e., as a set of oppositely valenced JMs), JMs for aggression were specified without an oppositely valenced set of prosocial JMs. The JMs for these constructs illustrate the earlier point that implicit biases can be either positively or negatively valenced. As an example, one JM for achievement motivation was labeled the *personal responsibility inclination* and defined by James (1998) as the:

> tendency to favor personal factors such as initiative, intensity, and persistence as the most important causes of performance on demanding tasks (p. 134).

This JM proposes that individuals with a strong motive for achievement make attributions that support their intense drive to succeed and their belief that personal success is dictated by controllable factors. In contrast, individuals with a strong fear of failure prefer external explanations for personal success (and failure). The implicit belief that personal success is largely uncontrollable supports these individuals' avoidance orientation. Following from these definitions, it would be anticipated that reasoning about causal attributions for success is conditional on the presence of an achievement versus fear of failure JM and that the presence of these JMs stems from an implicit motive to achieve versus avoid failure (James, 1998).

An example JM for aggression is the *derogation of target bias* or tendency to attribute deservingness or blame to the victims of aggression. A popular culture illustration involves the women who “deserved” to get sexually abused for having dressed risqué on a night out with friends. In other words, the derogation of target bias might conclude that she “got what she wanted,” or misled the perpetrator in some way. On the other hand, a non-aggressive interpretation might assert that women should be allowed to dress as they please without expecting a physical confrontation from males and that the perpetrator of the crime is unequivocally at fault. In this case, reasoning about the intentions and circumstances of the actors in the described scenario is conditional on the presence of a *derogation of target bias*, which should be held to a greater degree by individuals with a strong motive to aggress. As with the previous constructs, this is because individuals with particular dispositional tendencies (i.e., the motive to aggress) prefer to view their behavior as sensible and justified (James, 1998).

Specification of JMs serves as the basis for the operational measures of conditional reasoning and the resulting inferences about implicit dispositional tendencies. The following section introduces the conditional reasoning measurement system in greater detail and uses the CRT-A as the primary illustration of this method.

The Conditional Reasoning Measurement System

The measurement system developed by James et al. (1998) represents a psychometrically rigorous advancement within the longstanding tradition of using indirect or projective testing to assess personality, attitudes, and stereotypes (see Campbell, 1950; Greenwald & Banaji, 1995). Unlike many indirect assessments which require open-ended response formats and semi-structured or un-structured scoring protocols, conditional reasoning tests use “traditional” item formats with a series of multiple-choice items that are presented and scored in a standardized format according the theoretical development of JMs. Inductive reasoning problems of this type are the vehicle for inferring the presence or absence of specific JMs (James & Mazerolle, 2002).

Inductive reasoning involves a logical and probabilistic extrapolation from available information in the problem. Probabilistic conclusions are judgments about what is most likely to be true following from several premises or pieces of information. At the same time, other solutions could be possible provided additional information. This is unlike deductive reasoning, in which the conclusion must be true following from true premises (Moore, 1998). LeBreton et al. (2007) described the difference between inductive and deductive reasoning as an issue of *generality*:

> with deductive reasoning the inferred answer is of lesser (or equal) generality than the premises, but with inductive reasoning the inferred answer is of greater generality than the premises (p. 3).

A classic example taken from Moore (1998, p. 5-6) serves to further illustrate the difference between inductive and deductive reasoning. The examples shown below are written in argument form in which the premises are listed in a series of statements appearing above the horizontal line, and the conclusion that follows is shown below the line.

Deductive Example:

All men are mortal.

Socrates is a man.

Therefore, Socrates is mortal.

Inductive Example:

Socrates was mortal.

Sappho was mortal.

Cleopatra was mortal.

Therefore, all people are mortal.

In both examples, the conclusions offered follow from the assumed truth of the premises. In the deductive example the conclusion that Socrates is mortal must logically follow if it is the case that all men are mortal and that Socrates is a man. In contrast, the opposite conclusion (i.e., that Socrates is *not* mortal) is logically invalid. The deductive solution does not require a generalization beyond the information contained in the premises. In the case of the inductive argument, the conclusion that all people are mortal is a generalization from three specific examples of mortal people (i.e., Socrates, Sappho, and Cleopatra) to all people. The probability that the inductive solution is correct is enhanced by multiple instances of mortal people, but cannot be stated conclusively. Until the premises comprise all possible people, the inductive solution offered remains probabilistic.

These examples underscore several unique features of the inductive reasoning situation. First, the absolute truth of an inductive conclusion cannot be known definitively. Although some conclusions may connote strong probabilistic statements, such that alternative conclusions carry very low probabilities of being true, all that is necessary for disconfirmation is a single contradictory case (e.g., a single immortal person). Second, consistent with the increased generality of inductive reasoning, information beyond that contained in the premises can be leveraged to determine the likelihood of particular solutions. These might be considered un-stated premises. For example, a scientific understanding of the limitations of the human body might enhance one's confidence in the conclusion that all people are mortal, or alternatively, a belief in science fiction might decrease such confidence. Accordingly, there is a degree of uncertainty inherent to the inductive reasoning situation that is not true of deductive situations. Despite a certain degree of ambiguity underlying inductive reasoning, the *Principle of Induction*, which holds that past events provide a useful basis for judging future events, ensures that induction is not a frivolous activity. Instead, given accurate information about past events (e.g., the sun has risen every day so far), it is possible to make probabilistic inferences with reasonably high confidence (i.e., the sun will rise tomorrow) (Moore, 1998).

With the goal of couching conditional reasoning items within an inductive reasoning paradigm, item stems for conditional reasoning problems consist of a short paragraph containing a series of statements or premises. Item content is written to be provocative or increase the opportunity for individuals to display JMs that are consistent with their implicit dispositional tendencies. Four response options are provided, each of which presents a possible, although not necessarily plausible, conclusion. The participant's task is to solve the inductive reasoning problem by choosing the most logical or valid inference following from the stated premises. Among the four response options, two reflect clearly illogical conclusions and are included only to maintain the face validity of the inductive task. In other words, they represent inductively incorrect solutions. The other two options reflect logical conclusions that are contingent on the influence of a JM. That is, one of the options is designed to appear logical to someone whose reasoning is influenced by a JM (e.g., a hostile attribution bias), and the other is designed to appear logical to someone whose reasoning is not influenced by the JM or is influenced by an oppositely valenced bias (e.g., a socially adaptive JM) (James, 1998).

Because responding to any particular conditional reasoning item could reflect the influence of the JM of interest or some extraneous factor (e.g., prior knowledge or experience regarding item content), inferences about the individual are based on the pattern of responding observed across several items, each of which are designed to tap into one or more JMs (James & Mazerolle, 2002). A stronger implicit motive is inferred when an individual consistently favors solutions that reflect the underlying JMs, whereas a weaker (or lack of) implicit motive is inferred when an individual infrequently chooses these solutions or solely uses solutions that have the opposite valence of JMs. Accordingly, scale scores are derived by summing across item responses. It is also important to note that, although conditional reasoning tests use item formats that are consistent with inductive reasoning problems, scores on the CRT-A are uncorrelated with cognitive ability (James et al., 2005).

Faking. As with previous indirect measures, the major advantage attributed to conditional reasoning tests is resistance to faking. LeBreton et al. (2007) tested the fakeability of the CRT-A under different instructions and across samples with ostensibly different levels of motives for performing well on the test. First, these authors investigated the necessity of maintaining the indirect nature of the assessment by not disclosing its actual purpose in the instructions. Under the usual (i.e., indirect measurement) instruction set, test-takers are told that they will be completing a test to assess their inductive reasoning capacity and that they should choose the most logical solution to each problem. Importantly, they are not informed that the test is actually assessing aspects of their personality, or more specifically, their latent aggressive tendencies. Therefore, it is assumed that participants are unaware of the test's actual purpose under indirect testing conditions. In comparison to the usual instructions, LeBreton et al. (2007) hypothesized that disclosing the actual purpose of the assessment would influence test-taker responding. Two alternative instruction sets were tested. The disclose-fake instructions disclosed the purpose of the assessment (i.e., as an indirect assessment of aggression designed to look like an inductive test) and instructed participants to endorse the aggressive response option. The disclose-logic instructions disclosed the purpose of the assessment but did not instruct participants to fake bad or good, but rather to choose the most logical conclusion, as in the usual instruction set. It was anticipated that, in comparison to the typical testing conditions, disclose-fake instructions would yield higher mean aggression scores (i.e., indicating higher aggression) and that disclose-logic instructions would yield lower mean aggression scores.

Results provided mixed support for the anticipated effects of disclose-fake and disclose-logic testing conditions. Consistent with expectations, participants in the disclose-fake condition were able to inflate their aggression scores. Moreover, the magnitude of the effect was large, as indicated by a mean in the disclose-fake condition of 17.82 (*SD* = 3.83) versus a mean of 3.63 (*SD* = 2.02) under typical instructions, with possible scores ranging from 0 to 22. This suggests that test-takers are able to identify the aggressive response options when made aware of the test's actual purpose and instructed to "fake bad." On average, test-takers correctly identified the aggressive option on nearly 18 out of 22 items. A significant difference was also observed between the disclose-logic (*M* = 4.49, *SD* = 2.51) and typical instruction conditions, but in the opposite direction of the authors' hypothesis. Ultimately, it was unclear why test-takers in the disclose-logic condition responded more aggressively. LeBreton et al. (2007) speculated that participants might have been confused as to whether the logical or aggressive option was appropriate. However, the mean in this condition was comparable to the mean obtained under typical testing conditions in these authors' second study (see below), such that the difference observed is probably not meaningful. Nonetheless, these results clearly suggest that it is essential to maintain the indirect nature of conditional reasoning tests by using typical administrative procedures, which does not involve disclosing the true purpose of assessment.

A second study by LeBreton et al. (2007) tested the effect of fake good instructions while maintaining typical indirect testing conditions. This involved a within-subjects design in which undergraduate psychology majors completed the CRT-A twice, separated by a 1-week delay. For the first administration, test-takers were given the typical instructions and asked to select the most logical solution to each problem. For the second administration, test-takers were asked to complete the test as though they were applying for a customer service job that they really wanted. This was designed to mimic the motivated testing conditions typical in a job application scenario. A second condition with opposite ordering of the two instruction sets was used to achieve counterbalancing. In addition to the CRT-A, at both time points participants completed self-report measures assessing conscientiousness, agreeableness, emotional stability, extraversion, achievement motivation, and aggression. It was hypothesized that mean scores across the traditional self-report measures but not the CRT-A would be inflated under the job applicant testing condition.

Consistent with expectations, scores on the CRT-A were not systematically higher when participants were instructed to complete the measure as though they were applying for a desirable job (M = 4.32, SD = 2.40) versus under the usual instruction set (M = 4.63, SD = 2.49). In contrast, scores for all of the overt measures other than extraversion were significantly shifted in a socially desirable direction in the job applicant condition. LeBreton et al. (2007) interpreted this pattern of results as support for the idea that, unlike traditional self-report measures, the CRT-A does not appear fakeable when job applicant testing conditions are simulated with college students.

A third study, attempted to replicate and extend these findings by comparing CRT-A scores across groups that theoretically differ in their motivation to fake. This study compared mean test scores for incumbents of call centers, job applicants for package handling positions and temporary staffing agency positions, and college students majoring in business. The CRT-A was administered under usual testing conditions in all three groups (i.e., maintaining indirect measurement). Consistent with the findings from Study 2, it was hypothesized that these groups would not differ in their mean CRT-A scores, despite that the groups ostensibly differ in their self-presentation motives.

Results supported the authors' prediction, such that scores did not differ across the combined applicant sample (*M* = 3.32, *SD* = 2.15), incumbent sample (*M* = 3.30, *SD* = 2.13), and student sample (*M* = 3.55, *SD* = 2.22). A one-way ANOVA using a between-subjects group factor, as well as pair wise comparisons, were non-significant despite a large sample size of 966. These findings indicate that job applicants and incumbents do not systematically inflate their scores on the CRT-A in comparison to student samples. Taken together with the results of Studies 1 and 2, LeBreton et al. (2007) concluded that, instructions supporting the indirect nature of the assessment are necessary, and as long as indirect testing conditions are supported, the CRT-A does not appear susceptible to faking.

In addition to not disclosing the test's purpose, several actual inductive problems are interspersed among the conditional reasoning items. For example, the 22-item CRT-A includes three actual inductive items, with the idea that this will reinforce test-takers' belief that the test is assessing reasoning ability. Similarly, in order to prevent test-takers from over thinking or attempting to dissect items and to guarantee that test-takers do not enlist the help of others, it is recommended that conditional reasoning tests be administered under the same testing conditions as tests of cognitive ability, including time constraints and proctoring.

Predictive validity of the CRT-A. Conditional reasoning tests of aggression and achievement motivation were developed to be predictive of workplace relevant criteria. The focus of discussion here is on the predictive validity of the CRT-A specifically, as this measure has been the dominant focus of published research on conditional reasoning. The results of a number of validation studies were summarized in reviews by James et al. (2004, 2005), and more recently, meta-analyzed by Berry, Sackett, and Tobares (2010). These authors' draw divergent conclusions about the predictive validity of conditional reasoning tests of aggression for aggressive and counterproductive workplace behaviors.

James et al. (2005) summarized the results of 11 validation studies yielding 14 correlations between scores on a conditional reasoning test and a criterion involving behavioral manifestations of aggression. Five studies used the CRT-A (i.e., the final 22-item version) and six studies used earlier test versions, including a developmental version of the CRT-A that included trial items that were later discarded and the developmental VCRT, an alternative test format that uses visual presentation of test content and sixth grade (or lower) language requirements for written content. The total sample size was 1,538 and included undergraduate college students, patrol officers, nuclear facility operators, restaurant employees, package handlers, and employees in a variety of temporary jobs. Study designs were predictive (6), concurrent (1), postdictive (2), and experimental (2).

The criteria studied varied widely. Included were indices of aggressive behavior (e.g., the number of "hard" fouls committed by college students during intramural basketball games), supervisor ratings of overall job performance (among police officers), absenteeism, employee turnover, student conduct violations (e.g., lying and plagiarism), and theft. At first glance, several of these appear only weakly related to aggression, and indeed Berry et al. (2010) have challenged their appropriateness as “aggressive” criteria. James et al. (2005) have argued that, although illustrations of workplace aggression tend to focus on extreme or violent interpersonal acts (assault, vandalism, theft, etc.), aggression is more frequently manifest in subtle or passive actions, such as taking unauthorized absences or quitting a job, rudeness behaviors, and withholding of information or effort. A burgeoning literature on incivility supports this rationale and demonstrates the seriousness of these subtle forms of workplace deviance (see Caza & Cortina, 2007; Cortina, Magley, Williams, & Langhout, 2001). Thus, indirect forms of aggression are clearly important behaviors to be studied and can be included under the rubric of counterproductive workplace behaviors. However, the degree to which certain counterproductive behaviors reflect manifestations of aggression versus a different source(s) remains an open empirical question. This seems particularly true in the case of withdrawal behaviors, which have been linked to a range of non-aggressive antecedents (Spector et al., 2006).

Uncorrected validity coefficients ranged from .32 to .64 with an overall mean predictive validity of .44. These results demonstrate that scores on conditional reasoning tests of aggression are valid predictors of the criteria investigated, with observed effects representative of statistically significant and practically meaningful relationships. Following a similar argument as the one suggested above, James et al. (2005) divided the sample based on criteria that appeared directly or indirectly related to aggression. As tentative support for the scale’s construct validity (due to small number of studies in the two resulting samples), the mean uncorrected validity was higher among studies that focused on more obvious indices of aggression ($r = .50$ versus $r = .40$). Included in the sample demonstrating somewhat lower validity (i.e., the indirect aggression criteria) were studies of absence, turnover, performance, and unreliability behavior. James et al. concluded that the CRT-A is a valid instrument for predicting a range of counterproductive workplace behaviors and has predictive relationships commensurate with other established instruments, including integrity tests, overt personality tests, and tests of cognitive ability.

Berry et al. (2010) conducted a meta-analysis in follow-up to the review by James et al. (2005). These authors were interested in three potential moderators in addition to conducting their meta-analysis with a larger sample of existing validation studies. The moderators tested included: (a) dichotomous criteria versus continuous criteria, (b) undergraduate versus employee samples, and (c) specific test version, including the published version of the CRT-A, developmental versions of the CRT-A, and the visually based VCRT. The meta-analytic sample was nearly double the size of the James et al. review, including 12 studies, yielding 21 correlations of conditional reasoning test scores with criteria. Studies were grouped according to two categories of the criteria examined: performance criteria (e.g., supervisor ratings of performance) and CWB criteria (e.g., indices of aggression, theft, lying, and withdrawal). Due to small number of studies examining performance criteria, moderator analyses were conducted only within the CWB criteria category.

Berry et al. (2010) reported overall mean uncorrected validity coefficients of .16 for CWB criteria and .14 for performance criteria. These results differ markedly from the average validity of .44 reported by James et al. (2005). As possible sources of the observed divergence, Berry et al. describes several decision rules (e.g., coding and extraction of effects from primary studies) in their meta-analysis that diverged from those used in the prior review. However, the handling of individual studies was ultimately ruled out because Berry et al. demonstrated that re-estimating validity coefficients using the coding rules applied by James et al. had no appreciable effect on estimates. Therefore, the divergence of overall estimates was attributed to differences in samples, with Berry et al. including several additional studies that reported smaller validity coefficients than those included by James et al. (2005).

Berry et al.'s (2010) overall validity estimate for CWBs should be qualified by the results of moderation tests. First, the exclusion of studies with dichotomous criteria and low base rates (e.g., less than 10% committing the aggressive behavior) resulted in a higher estimate (r = .26) that is roughly equivalent to the predictive validity of integrity tests for CWB criteria. Second, the specific test version used served as a moderator, with the validity estimate for the CRT-A (r = .11) significantly lower than for developmental versions of the CRT-A (r = .35) and the VCRT (r = .24). Finally, studies' use of undergraduate versus employee samples did not systematically affect the predictive validity of conditional reasoning tests. Overall, Berry et al. called for additional research evaluating the predictive validity of conditional reasoning tests for aggression with special attention to the specification of criteria, but noted that the initial validity of .44 reported by James et al. (2005) "appears to have been "overly optimistic" (p. 379). The need for additional research is further underscored by Berry et al.'s finding that the final scale version yielded the weakest evidence for validity. However, the reader is reminded that the results demonstrated by Berry et al., as with those of the earlier review by James et al., should be viewed as preliminary evidence given the relatively small number of studies that were available.

Additional research is ultimately needed to determine the predictive validity of the CRT-A for counterproductive and aggressive workplace behaviors. Perhaps more importantly, additional research is needed to apply the conditional reasoning approach to the specification and measurement of additional constructs pointed at different behavioral criterion. The potential of this approach is demonstrated in the research on the specific measures for achievement motivation and aggression and echoed in repeated calls for the development of personality-based measures that do not rely on self-report (e.g., Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). Conditional reasoning tests may be particularly well suited for assessing negatively valenced constructs in testing contexts that are likely to engender strong self-presentational motives, such as in pre-employment screening.

A Conditional Reasoning Test of Withdrawal

Previous findings that demonstrate a predictive relationship between conditional reasoning tests of aggression and absence and turnover provide preliminary evidence that implicit cognitive biases may shape the enactment of withdrawal behaviors. In the work of James and colleagues (2004, 2005), the implicit biases assessed and linked to withdrawal behaviors were conceptually related to dispositional tendencies to aggress. The validity coefficients reported by James et al. for studies using absence criteria (*r*s = .34, .37, and .42) and turnover criteria (*r* = .32) reinforce these authors' point that withdrawal can reflect hostile motives in certain circumstances, such as when an employee intentionally misses work as a form of retribution. However, full consideration of the voluminous literature on employee withdrawal suggests that hostility might be an important antecedent for a narrow range of possible circumstances surrounding a withdrawal decision. Highlighted in the earlier literature review is a long list of non-aggressive motives, cognitive processes, situational factors, and personality constructs that have been shown to precede withdrawal.

Following Spector's work that distinguishes withdrawal from other forms of counterproductive work behavior (Spector et al., 2006), it is likely that a dispositional tendency to withdraw (previously described as "withdrawal proneness") is accompanied by specific *JMs for withdrawal* that develop within the individual in a parallel process as JMs for aggression, which James (1998) described as arising from a motive to aggress. A model of the rationalization process for aggression presented by James et al. (2005) provides a useful framework for describing a unique rationalization process associated with withdrawal. These authors' original figure is adapted (see Figure 1) to depict a parallel rationalization process for withdrawal.

Central to the model is the idea that conflict arises within individuals from simultaneous tendencies to withdraw and to maintain a favorable view of self. From an ego-protective vantage, people generally want (or need) to view themselves as good, moral, ethical, motivated by benevolent intentions, and under control (Bersoff, 1999). Self-esteem theories suggest that the motive to maintain a positive self-image and sense of self-worth is among the strongest and most basic drivers of behavior and reasoning (Greenwald, Bellezza, & Banaji, 1988). This motive has been referred to as "self-enhancement" and is evidenced by a variety of judgment biases (e.g., the better-than-average effect), each of which have the sum effect of enhancing information that is consistent with a positive self-image (Kunda, 1990; Silvera & Seger, 2004). At the same time, challenges to a positive self-image are deflected by a host of ego-protective defense mechanisms (Bersoff, 1999).

Stemming from the finding that past withdrawal behaviors tend to be the best available predictor of future withdrawal behaviors (e.g., Harrison & Hulin, 1989), research has focused on characterizing the “withdrawal prone” individual. Froggatt (1970) was among the first to study the dispositional sources of absence behavior. In describing the chronically absent employee, Froggatt suggested that withdrawal was, “the overt expression of a desire to work discontinuously” (p. 310) which might be only partially mitigated by factors external to the withdrawal prone worker. In other words, withdrawal stems from an avoidance tendency that is inherent to the person and which can be expected to manifest fairly consistently across situations. Subsequent research has found some support for a stable individual difference explanation, demonstrating modest relationships between personality traits and withdrawal behavior (e.g., Barrick & Zimmerman, 2005, 2009; Iverson & Deery, 2001; Judge et al., 1997; Ones et al., 2003; Salgado, 2002; Taylor, 1968). However, explicit traits may capture only part of the stable intrinsic component of withdrawal proneness and may be only indirectly related to an implicit withdrawal tendency. In reference to dispositional aggression, James et al. (2005) suggested that individuals’ motive to aggress resides at a largely implicit (or un-recognized) level because explicit recognition would conflict with a positive self-image. Reflecting an implicit dispositional perspective, I define the dispositional tendency to withdraw as *the propensity to withdraw chronically and in response to specific environmental events*.

This definition reflects the idea that the dispositional tendency to withdraw is continuously present and further stimulated by events in the individual's external environment (e.g., work situations). In this way, it is proposed that the dispositional tendency creates a "readiness" that can lead to automated or scripted withdrawal over time or in response to the environment. Without any necessary environmental stimulus, a withdrawal prone individual may become increasingly anxious or dissatisfied if he or she remains in the same job or role for an extended period. In other words, individuals may differ in their conceptions of time and commitment as well as the permanence of decisions, each of which is likely to guide job-changing and permanent withdrawal behaviors. It is anticipated that the dispositional tendency to withdraw is accompanied by implicit beliefs that support short-term and malleable conceptions of time, commitment, and decisions. At the same time that an implicit tendency to withdraw is expected to predispose the individual to shorter-term commitments in general, it is also likely to enhance the probability that behavioral withdrawal will be chosen as a "coping" response for the demands and conditions of continuous employment or continuous involvement in an ongoing commitment (e.g., school). In response to specific events, a dispositional tendency to withdraw is expected to shape perception and reasoning so that absence, tardiness, and quitting are viewed as justifiable actions. It is further likely that a personal history of withdrawal from prior commitments increases the availability of withdrawal within the behavioral repertoire, making future withdrawal more likely and more easily accessible via automatic or scripted routes (Lee & Mitchell, 1994).

Because withdrawal behaviors are commonly viewed as deviant by others and may be attributed to a range of personality “defects,” (e.g., laziness, aloofness, non-committal, etc.) a dispositional withdrawal tendency and the motive for positive self-image are at odds within the person (Johns, 1994). An intraindividual conflict arises due to the difficulty of reconciling a behavioral disposition that is outwardly negative with a desire to believe that one’s behavior is rational and good. JMs for withdrawal function as a mechanism for allowing the expression of withdrawal without sacrificing positive self-regard. JMs focus accounts of withdrawal on specific features of the social environment while deflecting personal responsibility. JMs provide a behavioral script for enacting withdrawal as a reasonable response to injustices or as part of a “natural” progression within a formal relationship (e.g., a script for changing jobs within a pre-specified time frame). JMs reinforce a short-term and malleable view of concepts surrounding personal obligation and commitment, thereby softening the perceived damage caused by contractual breaches. Figure 1 demonstrates how implicit JMs shape (or distort) the manner in which withdrawal behaviors are expressed and rationalized explicitly.

Three proposed implicit biases associated with behavioral withdrawal tendencies are summarized in Table 1 and described in greater detail in the sections below. Collectively, these JMs for withdrawal are anticipated to serve the retrospective and prospective functions of framing withdrawal as a socially acceptable or normative behavior and as a reasonable or even logical response to organizational events. From an ego-protective standpoint, individuals should be motivated to frame instances of withdrawal in a manner that reduces the common deviance account for such behaviors (e.g., absenteeism stems from laziness or malingering). For example, Johns (1994) suggested that, in an effort to avoid negative self-knowledge and co-worker disapproval:

> employees will be motivated to disassociate themselves from exhibiting too much absence and will attribute elevated levels of this negatively valued behavior to others (p. 229).

Withdrawal JMs are expected to operate within an individual through similar biased social cognitive routes as the JMs that support aggression. For example, withdraw JMs may operate below conscious awareness vis-à-vis biased attributional processes, selective attention and information filtering, differential framing, and enactment of schemata that support withdrawal. The proposed JMs described below reflect features of each of these information-processing routes. *Marginalization* is the self-serving attribution pattern specific to the rationalization of withdrawal behaviors. *Revocable commitment* is a motive to enact short-term and high malleability schemata in reference to commitment and related concepts (e.g., contracts, obligation, and reciprocity). *Social Injustice* is an overreliance on referent cognitions and a tendency to filter information using fairness themes, which subsequently predispose an individual to feel inequity. As with aggressive JMs, the sum anticipated effect of withdrawal JMs is to enhance the appeal of withdrawal as a behavioral response to organizational events and to do so without necessary conscious awareness.

Marginalization of withdrawal. As already stated, marginalization of withdrawal is a self-serving attribution pattern specific to the rationalization of withdrawal behaviors. The conceptualization of this JM arises from a fundamental tenet across social, cognitive, and clinical domains of psychology – that information processing is typically biased to minimize the impact of negative self-information. This premise is central to a range of well-described cognitive phenomena, such as cognitive dissonance, defense mechanisms, ego-protective biases, rationalization, self-serving attributions, the fundamental attribution error, self-presentation biases, and so on. Marginalization of withdrawal can be thought of as a specific application of these more general processing biases. Because withdrawal behaviors are widely regarded as a negative behavior in a variety of domains, it follows that people will be generally motivated to rationalize, conceal, downplay, and legitimize specific instances of withdrawal. The pressure to legitimize withdrawal likely arises from forces that are internal (e.g., guilt) and external (e.g., co-worker perceptions) to the individual (Hammer, Landau, & Stern, 1981). Following from the idea that legitimizing withdrawal serves an ego-protective or self-presentational purpose within an individual, it stands to reason that such processes will be most evident among individuals that withdraw most frequently (i.e., the withdrawal prone). In other words, the JM might be “built up” as a consequence of, and in order to support, frequent withdrawal behavior. Thus, although most individuals might be expected to engage in biased processing in order to justify or rationalize a specific instance of withdrawal, withdrawal prone individuals in particular may be hyper-prepared to do so.

Evidence of marginalization of withdrawal can take on several related forms. The first is a self-serving bias, in which the frequency of others’ withdrawal behavior is overestimated relative to one’s own. This pattern has been well documented in self-report assessments of absence and lateness. In accordance with a marginalization JM, perceiving high absenteeism among others in one’s social environment (whether via selective attention or retrospective biases) provides a basis for viewing one’s one absenteeism as normative rather than deviant. Studies by Harrison and Shaffer (1994) and Johns (1994) provide important illustrations of a possible marginalization of withdrawal process among withdrawal prone individuals.

In their aptly named “Lake Wobegon” study, Harrison and Shaffer (1994) investigated self-reports of absenteeism relative to reports of peers’ absenteeism and objective absenteeism data. Across seven field studies using diverse student and employee samples, their results provided compelling evidence that individuals generally underestimate the frequency of their own absenteeism while overestimating the frequency of absenteeism of an average peer or co-worker. On average, between 85 and 90% of participants indicated that they were absent less often than an average peer or co-worker, with the estimated frequency of absence for oneself approximately half that of the estimated frequency of absence for the counterpart. By comparing self-reported frequency of absence and estimated frequency of peer absence against corresponding objective absenteeism data (i.e., for the participant and the participant’s peer group), a large upward bias was demonstrated for ratings of others’ absenteeism whereas a moderate downward bias was demonstrated for self-ratings of absenteeism. In other words, the large discrepancy between estimates of self- and other-absenteeism is due to a simultaneous overestimation of others’ absence and underestimation of one’s own absence, although the overestimation of others is more robust. Furthermore, ratings of peers’ absenteeism were positively correlated with participant’s absenteeism over a follow-up period (contributed an additional 10%), above and beyond the participant’s previous absenteeism record (r = .78). Finally, the downward bias for self-rated absenteeism was somewhat reduced when ratings were solicited during a casual conversation among peers as opposed to from a mailed survey or face-to-face interview with a researcher.

A study by Johns (1994) observed a similar self-serving attribution pattern in employees (utility workers and teachers) estimates of their own absence behavior relative to company records of their absence over a matched time period and relative to the average employee in their organization. On average, self-reported absence was underestimated in comparison to matched objective data by roughly two-fold. As with the Harrison and Shaffer (1994) study, self-reported absence was also significantly less than estimates of an average co-worker's absence. An important finding from this study was that the degree of under-reporting in the self-report measure was positively associated with the number of actual absences. In other words, the most frequently absent employees were also the most egregious in terms of under-reporting their own absence relative to the average other. Johns explained that, “individuals with elevated absence records would be motivated to engage in more extreme underreporting” (p. 233). This finding provides tentative evidence of an underlying marginalization process. Furthermore, qualitative accounts by these employees indicated mainly external attributions for own-absenteeism behavior (i.e., a fundamental attribution pattern).

In addition to shaping strong self-serving attribution patterns that support withdrawal, the marginalization of withdrawal JM is expected to frame the circumstances surrounding withdrawal using “soft” language, which supports specific beliefs about the seriousness (lack thereof) of absenteeism and lateness. For example, a marginalization JM would support externally expressed beliefs that manager reactions to withdrawal should be lenient, that withdrawal (e.g., turnover) has many positive consequences for organizations and individuals, or that some level of withdrawal is simply a “fact of life.” The JM might also lead to selective attention toward factors that are thought to compensate for withdrawal. For example, one might implicitly believe that starting work an hour late is permissible as long as one works twice as hard after arriving. Similarly, the absence taking process may be framed as a coping behavior that prevents more severe negative consequences including turnover, burnout, and work-family conflict. Indeed, similar thinking has guided the development of compensatory models of withdrawal (Dalton & Todor, 1993).

Revocable commitment. Revocable commitment stems from turnover models that emphasize commitment as a core, mediating construct in the process of withdrawal (Steers & Mowday, 1980). Although commitment can be viewed as unfolding over time from one perspective, from another vantage people differ in how they think about various commitment-related concepts, including decisions, obligation, and reciprocity. Research on normative and continuance facets of commitment, both of which imply cognitive rather than affective evaluations, provide an illustration of this idea. For example, research has begun to investigate how individuals differ in their attention toward organizational norms regarding commitment and their tendency to continue as a function of prior investments (e.g., Harrison & Price, 2003). A revocable commitment bias follows from this individual difference perspective, by describing some individuals’ proclivity to focus on the short-term and evolving nature of commitment concepts. For example, this bias is anticipated to frame job-related commitments as malleable and continuously evolving, and obligations as loose and conferring low demands for satisfying reciprocity. Such beliefs would support fluid job changing and withdrawal behavior.

A study by O'Reilly and Caldwell (1981) demonstrates how perceptions of the revocability of job decisions may influence subsequent withdrawal behaviors through job attitudes and behavioral commitment. These authors found increased job attitudes and reduced turnover among individuals who perceived their initial decision to accept a job as volitional (i.e., free from external constraints) and irrevocable (i.e., very difficult to change jobs). The explanation offered for these findings centered on the idea that job attitudes and behavioral commitment may be formed via retrospective rationalization or justification processes that bring beliefs into alignment with the circumstances surrounding previous behavior. O'Reilly and Caldwell likened this process to cognitive dissonance. Subsequent research has supported the importance of perceptions of job-decision revocability on the development of organizational commitment. Meyer, Bobocel, and Allen (1991) concluded that perceived revocability of the job acceptance decision was the strongest pre-hire correlate of later continuance commitment.

Although the irrevocability of a job decision has traditionally been viewed as resulting from applicant perceptions of the situation surrounding job acceptance (e.g., the number of alternative job options and, by extension, the ease of changing one's decision), it is also likely that dispositional differences influence the perceived irrevocability of a job acceptance decision. From this perspective, some individuals might generally view their decisions as strong forms of commitment that may be difficult to undo. Based on O'Reilly and Caldwell's (1981) study, this might serve as a predisposition to rationalization processes that result in positive job attitudes and strong behavioral commitment. Conversely, individuals who generally view their decisions as revocable would not engender the positive consequences of dissonance and would be expected to remain relatively un-attached to the organizations in which they work. Consequently, weak perceptions of obligation and continuance commitment would be expected to create little or no barrier to job mobility and liberal absence taking.

Social injustice bias. Individual difference concepts have evolved over time within fairness and equity theories. Equity theory (Adams, 1963) suggests that individuals judge the suitability of work outcomes (e.g., status, pay, and so on) in relation to the outcomes experienced by referent others. These cognitive evaluations are referred as referent cognitions. A state of inequity is thought to occur when one perceives that they are either overcompensated or undercompensated relative to others. A shortcoming of initial applications of equity theory was the universal implications for employees. That is, individual differences in fairness perceptions were largely un-recognized.

In response to this shortcoming, Huseman and colleagues (Huseman, Hatfield, & Miles, 1985, 1987; Miles, Hatfield, & Huseman, 1989, 1994) posited that individuals vary in the degree to which undercompensation and overcompensation are personally distressing. Their construct – *equity sensitivity* – reflects an individual's preference for different input to output ratios, where inputs refer mainly to expended work effort and outputs refer mainly to the tangible rewards that are received in exchange. Individuals can be classified along a continuum as Benevolents, Equity Sensitives, or Entitleds. In the middle of the continuum, Equity Sensitives prefer a balanced ratio of inputs to outputs. For these individuals, both undercompensation and overcompensation are thought to represent states of inequity, which in turn, promotes negative cognitive and affective outcomes (e.g., job dissatisfaction). Alternatively, Benevolents are most satisfied when their input exceeds their output, and Entitleds are most satisfied when the reverse is true. For these individuals, only certain types of inequity are thought to result in distress. For example, Benevolents are comfortable or tolerate working harder than is justified by their level of compensation (Miles et al., 1994), whereas Entitleds are skewed toward expecting greater reward for less effort. Finally, referent others' input to output ratios serve as the context for judging whether one's own ratio indicates equality, undercompensation, or overcompensation (Adams, 1963; 1965).

The social injustice JM proposed here reflects individuals' willingness to invoke perceptions of inequity to rationalize withdrawal. The basic propositions underlying this JM are that individuals will differ in their reliance on referent cognitions and that individuals who more frequently engage referent cognitions to understand ongoing information from their social environment will be generally predisposed to perceived inequity. For these individuals, all stages of reasoning are heavily shaded by fairness and justice concepts, which are continually "just below the surface," so to speak. Therefore, the social injustice bias represents a cognitive preparedness to engage fairness concepts in a self-serving manner. For example, selective attention and strategic choice of a referent other provide ready mechanisms for rationalizing withdrawal behaviors as an appropriate response to inequity. As an illustration, a teacher who engages a social injustice JM may implicitly justify taking occasional unexcused absences as being commensurate with a $30,000 (low) annual salary. Similarly, increased use of referent cognitions coupled with a fairness interpretation might imply that one's own withdrawal behavior is justified by management's failure to discipline others for being late or missing work.

Among the JMs proposed here, a social injustice bias has the most obvious conceptual overlap with JMs for aggression. For example, James et al. (2005) described the retribution bias as tendency to attribute importance to revenge and righting perceived wrongs. Clearly, a perceived inequity could engender withdrawal as an act of retribution against the organization, and indeed, this parallels closely with the explanation offered by James et al. for studying the predictive validity of the CRT-A using withdrawal criteria. Therefore, some overlap is anticipated between the social injustice JM for withdrawal and aggressive JMs that promote retribution. However, a distinction is also maintained because a social injustice explanation for withdrawal need not stem from dispositional aggression. That is, an employee with low dispositional aggression could justify withdrawal as fair (or socially just) without any underlying hostility. More broadly, social injustice simply provides a mechanism for conceiving of withdrawal as a fair and justified response.

The Present Study

The overarching goal of the present research is to develop and validate the CRT-W as a new measure that assesses dispositional withdrawal tendencies. Potential applications for the CRT-W include a range of contexts in which predicting withdrawal behaviors could be useful, the most obvious of which is employment testing. For example, it is anticipated that pre-hire applicant scores on the CRT-W would demonstrate predictive relationships with subsequent employee withdrawal behaviors, including turnover, absenteeism, and tardiness. If this pattern holds true, organizations could use the CRT-W to screen out job candidates who would be likely to demonstrate withdrawal behaviors as employees. Thus, the CRT-W is developed primarily with employment testing applications in mind.

Other potential applications include predicting attendance and successful program/course completion in training and educational contexts. It is the latter setting – withdrawal behaviors within a large introductory psychology course – that provides the basis for the present investigation. A few comments are warranted regarding use of a student sample and the likelihood that results would generalize to an employee sample. First, several parallels have been drawn between the nature of student and employee roles (e.g., see Munson & Rubenstein, 1992). For example, both students and employees work independently and in groups to complete performance tasks within a specified time frame under the supervision of an instructor. Most importantly for the current study, a parallel can be drawn between the importance of attendance behavior and the underlying incentive systems that influence attendance in academic and work settings. Like employees, students are expected to arrive on time and remain in class until sessions are finished. Although one might argue that the disincentives for not attending are less severe in a classroom (e.g., resulting in the loss of participation credit, missing points on a quiz, or being chastised by an instructor), good attendance behavior is clearly an important aspect of academic performance for the vast majority of students. Likewise, permanent withdrawal from a course has the potential to carry negative consequences for the student, such as when financial loss or loss of credit results or when a permanent record of the drop appears on the student’s transcript.

Therefore, withdrawal behaviors have a negative connotation in both settings, and students, like employees, should be motivated to exhibit good attendance.

Along the same lines, it is expected that the constructs assessed in the present study are applicable to withdrawal behavior across a variety of domains including work and school. Accordingly, items for the CRT-W reflect dispositional tendencies assessed broadly (i.e., not specific to work contexts), even though item content is rooted in specific withdrawal-relevant scenarios, some of which focus on employment contexts. Finally, as with the approach taken to developing and validating the CRT-A, studying withdrawal across a range of academic and non-academic contexts ultimately enhances the generalizibility of validation evidence for the CRT-W (James et al., 2005).

Four main issues provide the rationale for developing a conditional reasoning test tailored specifically to withdrawal proneness: (a) predictive validity, (b) incremental validity, (c) resistance to faking, and (d) broader appeal and user acceptance than the CRT-A. The present study was designed to test the first two issues, whereas the latter two reflect anticipated advantages of a conditional reasoning approach to assessing dispositional withdrawal tendencies in employment testing contexts. These potential advantages are described in more detail below, followed by sections that present specific study hypotheses in reference to points 1 and 2 from above.

The central reason for advancing a conditional reasoning measurement approach is the possibility of reduced fakeability. Research evidence confirms that faking is reduced or eliminated in comparison to traditional self-report measures as long as indirect testing conditions (e.g., inductive instructions) are maintained (LeBreton et al., 2007). Faking is expected when test-takers are highly motivated to present themselves favorably, such as would be expected in “high stakes” testing scenarios. This may be further exacerbated in response to items that assess negatively valenced traits such as aggression or withdrawal. Despite the stance of some industrial psychologists that faking is a “red herring” issue for personality testing because criterion-related validity coefficients are relatively unaffected (e.g., Ones, Viswesvaran, & Reiss, 1996), the past decade has seen continued research emphasis on developing new methods and test formats that combat participants’ ability to inflate test scores (Berry, Sackett, & Wiemann, 2007). Beyond the issue of criterion-related validity, faking raises concerns associated with fairness and may detract from the perceived legitimacy of personality testing in employment settings. For example, organizations may be skeptical about using tests that appear easily fakeable, and consultants may be reluctant to recommend use of personality tests that appear un-sophisticated.

Another reason for developing a conditional reasoning test that focuses specifically on withdrawal tendencies has to do with the practical issue of user acceptability. Some organizations or practicing psychologists may be reluctant to use or recommend use of the CRT-A for applicant screening when the primary criteria of interest are withdrawal behaviors. Despite the fact that the CRT-A has been statistically validated against withdrawal criteria in four studies to date (James et al., 2005), it may be difficult to make an intuitive case to managers regarding the connection between dispositional aggression and attendance behaviors, particularly if other measures under consideration have a more straightforward connection to the criterion of interest. For these reasons, the CRT-A may be difficult to implement in situations where organizations are specifically focused on preventing employee withdrawal behaviors (e.g., turnover) and have less immediate interest in preventing aggressive CWBs. Alternatively, the CRT-W may offer no practical advantage over the CRT-A if overt measures of equivalent or superior validity are available. In short, additional research is needed on the user acceptability of conditional reasoning testing in employment contexts in general, as well as the acceptability of specific tests given the organization's emphasis on predicting specific criteria. However, it is anticipated that, because the CRT-W is conceptually aligned with withdrawal criteria, which are broadly emphasized across a wide range of job and organizational contexts, the CRT-W would exhibit favorable acceptance by users.

Predictive Validity of the CRT-W

The present study is designed as test of the predictive validity of the CRT-W for employee turnover, absenteeism, and lateness. Figure 2 shows a conceptual model linking withdrawal JMs to specific withdrawal behaviors and illustrates the expected relationships that constitute *Hypotheses 1* through *8*. The study's hypothesized relationships reflect the conceptual development of the JMs and the types of withdrawal behaviors that are most directly relevant for each. The specification of the marginalization of withdrawal bias as broadly implicated in several forms of temporary withdrawal and permanent withdrawal leads to the following hypotheses:

> *Hypothesis 1*: Individuals with a stronger marginalization of withdrawal bias exhibit a higher frequency of lateness behaviors.

> *Hypothesis 2*: Individuals with a stronger marginalization of withdrawal bias exhibit a higher frequency of absenteeism behaviors.

> *Hypothesis 3*: Individuals with a stronger marginalization of withdrawal bias have a higher likelihood of permanent withdraw.

Unlike the marginalization of withdrawal JM, a revocable commitment bias is proposed to influence how an individual thinks about and behaves in relation to commitments that are typically viewed as binding, long-term, and contractual. For this reason, revocable commitment is most directly relevant for permanent withdrawal behaviors and may be less directly relevant for temporary withdrawal behaviors (i.e., lateness and absenteeism). Therefore, *Hypothesis 4* specifies a positive relationship with permanent withdrawal, and no hypotheses are offered with respect to lateness and absenteeism (although exploratory analyses will investigate these pathways):

> *Hypothesis 4*: Individuals with a stronger revocable commitment bias have a higher likelihood of permanent withdraw.

The specification of the social injustice bias as broadly implicated in several forms of temporary withdrawal and permanent withdrawal leads to the following hypotheses:

> *Hypothesis 5*: Individuals with a stronger social injustice bias exhibit a higher frequency of lateness behaviors.

> *Hypothesis 6*: Individuals with a stronger social injustice bias exhibit a higher frequency of absenteeism behaviors.

> *Hypothesis 7*: Individuals with a stronger social injustice bias have a higher likelihood of permanent withdraw.

With marginalization, revocable commitment, and social injustice JMs underlying the implicit dispositional tendency to withdraw, it is expected that individuals with a stronger overall tendency to withdraw will display a higher frequency of all forms of withdrawal. In this way, the following hypotheses are in reference to participants' overall scores on the CRT-W (rather than their scores on specific JM sub-scales, as specified previously):

> *Hypothesis 8*: Individuals with a stronger dispositional tendency to withdraw (as evidenced by stronger overall levels of marginalization, revocable commitment, and social injustice biases) exhibit a higher frequency of lateness behaviors.

> *Hypothesis 9*: Individuals with a stronger dispositional tendency to withdraw exhibit a higher frequency of absenteeism behaviors.

> *Hypothesis 10*: Individuals with a stronger dispositional tendency to withdraw have a higher likelihood of permanent withdraw.

Incremental Validity of the CRT-W

A related goal involves testing for the incremental validity of the CRT-W (at the total scale level) beyond existing measures that have been found to reliably predict withdrawal behavior. For this purpose, the present research will also investigate the predictive validity of personality traits from the FFM model (conscientiousness and emotional stability) and a withdrawal-tailored biodata measure. Examining the incremental validity beyond these measures will help to determine the usefulness of combining the CRT-W with personality and biodata predictors in a selection battery.

Obviously, the ability of the CRT-W to outperform or add incrementally to other measures is partially dependent on the predictive validity of the other measures, as well as the intercorrelations among the measures. Personality traits from the FFM have been related to several forms of withdrawal behavior in previous research. Conscientiousness and emotional stability in particular have been supported. For example, a meta-analysis by Salgado (2002) reported mean corrected (uncorrected) validity coefficients for turnover criteria of .35 (.25) and .31 (.23) for emotional stability and conscientiousness, respectively. A meta-analysis of 28 studies using personality-based integrity tests to predict absenteeism reported a mean corrected validity of .33 with absenteeism criteria (Ones, Viswesvaran, & Schmidt, 2003). Personality-based integrity tests assess a confluence of traits associated with conscientiousness and honesty (Berry et al. 2007). More recently, studies by Barrick and Zimmerman (2005, 2009) obtained validity coefficients in the low .20s for personality traits predicting turnover (e.g., r = -.21 for conscientiousness, r = -.22 for emotional stability). Therefore, the following hypotheses were advanced:

> *Hypothesis 11*: Individuals with higher levels of conscientiousness exhibit a lower frequency of lateness behaviors.
>
> *Hypothesis 12*: Individuals with higher levels of conscientiousness exhibit a lower frequency of absenteeism behaviors.
>
> *Hypothesis 13*: Individuals with higher levels of conscientiousness have a lower likelihood of permanent withdrawal.
>
> *Hypothesis 14*: Individuals with higher levels of emotional stability exhibit a lower frequency of lateness behaviors.
>
> *Hypothesis 15*: Individuals with higher levels of emotional stability exhibit a lower frequency of absenteeism behaviors.
>
> *Hypothesis 16*: Individuals with higher levels of emotional stability have a lower likelihood of permanent withdrawal.

Previous research indicates that a strong predictor of future withdrawal behavior is prior withdrawal behavior (Harrison, 2002). Therefore, biodata measures have been constructed to assess the frequency with which individuals have demonstrated job-changing as well as temporary withdrawal behaviors. For example, Barrick and Zimmerman (2005, 2009) found that two specific facets of a biodata inventory administered pre-hire – *prior job changing behavior* and *social embeddedness* – predicted voluntary turnover up to two years post-hire in separate predictive validation studies using organizational samples (*r*s in the .20s). Social embeddedness referred to the number of social contacts (i.e., friends and relatives) an individual had within the organization prior to joining. Therefore, it is hypothesized that an adapted biodata measure assessing previous withdrawal behaviors in academic contexts and social embeddedness in the course (i.e., number of friends and/or relatives taking the course at the same time) will predict withdrawal criteria:

> *Hypothesis 17*: Individuals who have higher levels of prior withdrawal exhibit a higher frequency of lateness behaviors.
>
> *Hypothesis 18*: Individuals who have higher levels of prior withdrawal exhibit a lower frequency of absenteeism behaviors.
>
> *Hypothesis 19*: Individuals who have higher levels of prior withdrawal have a higher likelihood of permanent withdrawal.
>
> *Hypothesis 20*: Individuals who are more socially embedded exhibit a lower frequency of lateness behaviors.
>
> *Hypothesis 21*: Individuals who are more socially embedded exhibit a lower frequency of absenteeism behaviors.
>
> *Hypothesis 22*: Individuals who are more socially embedded have a lower likelihood of permanent withdrawal.

Conscientiousness, emotional stability, and tailored biodata measures have demonstrated predictive validity with behavioral withdrawal criteria; however the observed relationships (i.e., *r*s ranging from the low .20s to the low .30s) are not so strong as to preclude the incremental validity of other measures. This possibility is further underscored by prior research that has demonstrated low intercorrelations among measures of implicit and explicit personality (e.g., Frost et al., 2007). Therefore, it is hypothesized that the CRT-W will contribute incrementally to the prediction of withdrawal behaviors above and beyond measures of conscientiousness, emotional stability, and previous withdrawal behaviors and social embeddedness as assessed by a biodata inventory:

> *Hypothesis 23*: A dispositional tendency to withdraw (as assessed by the CRT-W) explains variance in behavioral withdrawal criteria above and beyond explicit measures of personality and biodata measures of prior withdrawal behaviors and social embeddedness.

In addition to testing the study's hypothesized relationships, several additional analyses were undertaken on an exploratory basis. These included the following: (a) exploratory tests of the predictive relationships between the remaining three traits within the Five Factor Model of personality (i.e., extroversion, openness to experience, and agreeableness), (b) exploratory tests of the predictive validity of additional facets of an adapted biodata inventory (commute method and difficulty, withdrawal intentions), and (c) exploratory analyses testing course grades as a predictor of withdrawal behaviors and potential covariate for analyses examining the predictive validity of the CRT-W.

Chapter 2

Method

This study was designed for the dual purpose of developing a new conditional reasoning test of withdrawal and providing an initial examination of the test's predictive validity. This study was conducted with a sample of college students over the course of a 16-week academic semester. The CRT-W and all other predictor measures were administered during testing sessions held throughout the semester. Students' withdrawal behaviors in an introductory psychology course were gathered over the same time and served as the criteria for testing hypotheses.

Sample

Participants included undergraduate students enrolled in an introductory psychology course at a large university in the Midwestern United States. Among the 374 students enrolled in the course, a total of 253 completed the predictor measures for the study, representing a 67.6% participation rate. A total of 213 of these participants also consented, on a separate occasion for an ostensibly separate study, to participate in the criteria collection portion of the present study. Thus, matched predictor and criterion data were available for 213 participants. The larger sample of 253 participants provided the baseline sample for scale development analyses, whereas the smaller sample provided the baseline sample for hypothesis testing. Prior to analysis, samples were further winnowed on the basis of data screening (see Results for additional details). Table 2 displays the demographic characteristics of the baseline samples for scale development analyses and hypothesis testing. As shown in the table, the samples have highly similar demographic compositions.

The baseline sample for hypothesis testing (n = 213) is consistent with the enrollment target based on power analysis. Using the formulas provided by Cohen, Cohen, West, and Aiken (2003), it was estimated that between 190 and 200 participants would be needed to achieve adequate statistical power to detect a bivariate correlation of .17. In reference to a predictive validity, this corresponds to 3% of the variance in withdrawal behaviors, which was deemed the smallest predictive effect connoting potential practical significance. Based on one-tailed significance tests and alpha set to .05, the power to detect this effect with the present sample size was .81.

Study Design

The present study followed a modified predictive validation design in which the timing of predictor collection generally preceded the accumulation of criterion data. However, the exact timing of predictor collection varied across participants. The introductory psychology course spanned a total of 16 weeks beginning with the first day of lecture (i.e., week 1) and concluding with the final exam (i.e., week 16). Predictor measures, including the CRT-W, were collected during testing sessions that were conducted in weeks 3 through 15. The number of participants completing testing sessions each week was equal to the following: week 3 = 55, week 4 = 40, week 5 = 13, week 6 = 24, week 7 = 12, week 8 = 24, week 9 = 25, week 10 = 27, week 11 = 8, week 12 = 12, week 13 = 0, week 14 = 5, and week 15 = 8. Behavioral withdrawal criteria accrued over the 16 week semester based on students' lecture and lab attendance.

Measures

Conditional reasoning test of withdrawal. The test booklet and 30 items comprising the full CRT-W that was administered in the present study are shown in

Appendix A. The test included 25 conditional reasoning items and five actual inductive items. The ratio of inductive items to conditional reasoning items (i.e., 5:25) was based roughly on the ratio used in the CRT-A (i.e., 3:22). Among the conditional reasoning items, 9 items assessed the *marginalization of withdrawal* JM, 10 items assessed the *revocable commitment* JM, and 6 items assessed the *social injustice* JM. To facilitate the "realness" of the testing experience, the instructions contained with the test booklet displayed an actual inductive item with the correct response shown for illustrative purposes. The first two items on the test were actual inductive items, and the other three inductive items were positioned roughly equidistant throughout the remainder of the test. The ordering of the conditional reasoning items was determined randomly, and adjustments were made to ensure that items from the same JM did not appear consecutively. Table 4 shows the final ordering of items within the test.

Item generation. The item generation process followed the recommendations and examples provided by James and colleagues (e.g., James & Mazzerolle, 2002; James & LeBreton, 2011). Items were drafted by the study author and reviewed by two subject matter experts with extensive experience writing and implementing conditional reasoning tests. Over several iterations, the subject matter experts reviewed the items for face and content validity based on the operational definitions of the JMs that were provided (see Table 1). The test was also piloted with a small number of undergraduate students to verify that the instructions and items were clear and to ensure that the time required for each item did not greatly exceed 1 minute.

Item examples. In this section, an example item from each JM is presented to illustrate the design of CRT-W items and describe in greater detail the proposed rationale behind each JM's effect on the inductive reasoning process. The first example item is from the marginalization JM:

> Many universities are now using a delayed schedule in which classes begin 10 minutes after the hour rather than on the hour. For example, 8am classes start at 8:10am, 9am classes start at 9:10am, and so on. According to several universities, this has led to a reduction in attendance problems, and students report liking this schedule more than the traditional one. Which of the following is the most appropriate advice for a business considering a delayed schedule?

Response options were: (a) Be careful not to compromise important aspects of a productive workplace, (b) Provide reasons for going green at work, (c) Use delayed schedules because they are more closely aligned with people's natural tendencies, and (d) Consider how time zones may affect travel.

This item presents an inductive generalization problem. Participants are told about a policy that delays the start time of college courses by 10 minutes and are asked to consider several assertions regarding use of a similar policy in a business context. Thus, the key issue whether the positive effects of a delayed attendance policy can be expected to generalize from one context to another. Options *b* and *d* have no logical connection to the premises in the item stem. Options *a* and *c,* on the other hand, provide competing solutions and are designed to attract participants with different justification processes.

The scenario described in the item stem is designed to evoke a positive response from individuals who hold a marginalization of withdrawal bias. Underlying the scenario are two interrelated ideas both of which undermine the importance of punctuality. The first is that punctuality is imposed and runs counter to individuals' natural inclination, which is to be late. The second is that institutional policies on attendance could be improved if they were more accommodating of how people naturally behave. For the individual who holds a marginalization bias, both themes are logically attractive and provide a strong basis for accepting the inductive generalization from academic to business contexts. Thus option *c* is the favored solution.

On the other hand, individuals who do not hold a marginalization bias should be motivated to find logical fallacies in the item's premises and underlying assumptions, and consequently favor a solution that does not permit the inductive generalization to business contexts. For example, one might challenge the idea that rules should be accommodating by bringing to mind a number of examples in which rules force people to act unnaturally (or against preference) to their benefit (e.g., being forced to eat broccoli as a child). Option *a* represents a logical extension of this thought process by drawing a connection between a more disciplined view of punctuality and increased productivity in the workplace. As a result, the academic context comes to be viewed as representing a "special case" rather than having generalized applicability, and response option *a* is the favored solution.

The second example item is from the revocable commitment JM:

> The old saying that, "there are a lot of fish in the sea …," suggests that one should consider many possibilities before making a choice, and that the process of exploring alternatives leads to better final decisions. For example, one should try many different jobs before deciding on a career and come up with a number of ideas before deciding what to write about for a senior thesis. Which of the following is the biggest problem with the fish-in-the-sea saying?

Response options were: (a) Other cultures have different sayings, (b) It still implies that one eventually has to make a "final" decision, (c) It overlooks the possibility that the best option will be discovered early on, and (d) It is rare that online dating leads to a meaningful relationship.

The item stem explains the meaning of the “fish-in-the-sea” saying as the importance of trying many alternatives before committing to one and asks the participant to point out the logical fallacy with this advice. Options *a* and *d* are illogical distractors with no logical connection to the inductive premises of the item. Options *b* and *c,* on the other hand, provide competing solutions and are designed to attract participants with different justification processes.

The message contained in the item stem – that one should not overcommit – was written to appeal to the individual that holds a revocable commitment bias. This individual is likely to reason affirmatively that it is logically advantageous to remain noncommittal. For example, by staying open to all possible courses of action, one avoids potential missteps and can change courses freely. Consequently, this individual has difficulty finding a logical fallacy with the saying and rather agrees with its logical tenability. However, option *b* provides this individual with an opportunity to discredit the premise by agreeing with it to an extreme level. In other words, although the item stem promotes the idea of eschewing over commitment, it still suggests that a “final” decision or commitment is inevitable and that “playing the field” is a temporary advantage. Hence, option *b* is attractive because it allows the individual to discredit the idea that a final decision or commitment is necessary.

Alternatively, option *c* is attractive to the individual with the opposing tendency to reason in favor of strong and binding commitments. This individual would likely agree that the act of commitment focuses one’s efforts and increases behavioral intensity toward goal attainment. This individual would also be inclined to point out the many problems associated with “failure to commit” and “indecision.” For the present item, option *c* is logically attractive because the individual may see lost opportunity as an important risk associated with “playing the field” for too long. Instead, staying open to all possible courses of action is only advantageous until an acceptable course of action is secured – keeping one’s options open is a means to a better end, rather than an end in and of itself.

The third example item is from the social injustice JM:

> An old story is said to involve a man and the man’s wife who is dying and desperately needs medication. However, the medication is expensive, and the man has no way of paying for it or taking a loan. Therefore, the story presents the question of whether or not the man should steal the medication in order to save his wife. Which of the following is the most important question for determining whether stealing is okay in this example?

Response options were: (a) Has the man been rejected by society? (b) Is the man attractive? (c) What is the woman’s favorite food? and (d) Can the man make amends (or make up for) for his crime?

The item stem describes a classic morality conundrum, in which a poor man is confronted with a difficult choice to steal or risk the possibility of losing his sick wife. The participant is not asked to provide a solution to the conundrum, but rather to determine the most important question to ask in order to determine whether stealing is moral or immoral in this situation. Options *b* and *c* are illogical distractors with no logical connection to the inductive premises of the item. Options *a* and *d,* on the other hand, provide competing solutions and are designed to attract participants with different justification processes.

The message contained in the item stem suggests that what might be considered immoral in one situation could be considered moral in another. While most people will agree that stealing is wrong in general, the morally ambiguous scenario described is designed to tip the balance in favor of viewing stealing as morally defensible in certain rare and extreme instances. The task for participants is to find the "nail in the coffin" that proves this case as exceptional and justifies the man's choice to steal. Individuals who hold a social injustice bias are likely to frame this situation, as most other situations, in terms of equity: Do the parties involved receive equitable treatment and outcomes? By framing situations in terms of equity, inequity becomes the routine justification for behavior that is on the surface immoral. Therefore, to justify this *poor* man's immoral act of stealing, the individual with a social injustice bias wants most to point to an injustice that has been committed (e.g., by society), thus favoring option *a* as the logical solution.

Option *d*, on the other hand, is designed to appeal to a different thought process. For some individuals, being poor may have little connection with a social injustice, particularly if an internal attribution is made (e.g., the man and his wife are lazy). Others may see the connection to potential inequity but display resistance to using inequity as an excuse for "bad behavior," instead preferring a different sort of rationalization. One such rationalization, reflected by option *d*, involves the man repaying his debts to society in the future. In this case, rather than seeking a retrospective excuse (i.e., stealing is justified by past inequity), a prospective excuse is preferred (i.e., stealing is justified by later repayment). From a slightly different perspective, option *d* is logically appealing for those who rely less on referent thinking – a central feature of a social injustice bias – when making sense of others' behavior (or their own) because it focuses on the man as the source of the problem/solution rather than others surrounding him (e.g., society).

Item scoring. Two methods have been reported for scoring conditional reasoning items in past research (James & LeBreton, 2011). For the achievement motivation measure, items were scored using a 3-point scale, with 1 assigned to JM responses, 0 assigned to illogical distractor responses, and -1 assigned to non-JM responses. Accordingly, scale scores for the achievement motivation measure were derived by summing all the JM (+1) and non-JM (-1) responses across items. In this way, JM and non-JM responses had a cancelling effect within the computation of scale scores. The second method was implemented for the aggression measure and used a 2-point scale, in which illogical distractor responses and non-JM (i.e., non-aggressive) responses were assigned a 0 and JM responses were assigned a 1. Following from this method, scale scores for the aggression measure were derived by summing all the JM (+1) responses. Based on a preliminary comparison of the two methods, James and LeBreton noted that the choice of scoring procedure had little effect on the psychometric characteristics of the aggression scale, thus the simpler 2-point format was adopted.

A modified scoring procedure was adopted for the present study. JM responses were scored 1, non-JM responses were scored 0, and illogical distractor responses were treated as missing item-level data (see Table 4 for the item scoring key). Scale scores were then derived by taking the mean, rather than the sum, across items. In this way, a stronger dispositional tendency for withdrawal is inferred from a higher proportion of items on which the JM response is endorsed.

The main reason for preferring this approach is the type of inference underlying the illogical distractor option. Scoring this option a 0 within a 3-point format (i.e., +1, 0, -1) assigns the illogical distractor response as the mid-point of an ordinal variable, for which a +1 indicates greater reliance on the JM and a -1 indicates opposition to the JM. The implication is that a 0 should connote a moderate or neutral level in reference to the JM. However, given that the distractor response option is illogical other inferences are equally plausible, including poor reasoning on the item, lack of attention, misinterpretation of the item stem or response options, or a random guess. Alternatively, scoring both the non-JM and distractor options as 0 on an ordinal scale implies that both responses carry the same inference – that is, a lesser reliance on the JM. However, equating these two responses is also problematic in light of the multiple interpretations note above for the illogical distractor response.

For these reasons, the present scoring procedure was adopted as a way to differentiate the illogical distractor response option from both the JM and non-JM response, while avoiding the inferential problem associated with positioning the illogical response as the mid-point of an ordinal scale. Although this method has the disadvantage of creating missingness at the item level, two key factors minimize the impact of this decision. First, the percentage of respondents endorsing the illogical distractor option was relatively low across items (see Results), such that the amount of missing item-level data as a result of this decision was also low. Second, hypothesis testing and associated inferences about participants' level of dispositional withdrawal tendencies are based on scale scores which were derived as the mean across multiple items. As a result, a missing item-level response does not impact the direct comparability of scale scores across participants. Finally, item and scale characteristics were examined based on all three of the scoring procedures described above. The modified scoring procedure adopted here produced similar conclusions overall;

however, for a small subset of items, the item-criterion correlations were somewhat higher using the modified procedure. In general, this reinforces James and LeBreton's (2011) conclusion that different scoring procedures have a minimal impact on substantive findings for conditional reasoning items. At the same time, the new procedure is better aligned with the inference underlying an illogical distractor response, and may confer a small benefit in the measure's predictive validity.

IPIP personality scales. Personality traits from the Five Factor Model (FFM) were assessed using the shortened version of the International Personality Item Pool (Goldberg, Johnson, Eber, et al., 2006). The scales include 10 items per trait, each of which uses a 7-point response format ranging from 1-*very inaccurate* to 7-*very accurate* (see Appendix B). Goldberg et al. reported good convergent validity between Costa and McCrae's (1992) NEO-PI-R and corresponding traits from the IPIP (mean r = .73 across traits). Consistent with previous research (see http://ipip.ori.org/), internal consistency reliability was adequate in the present study, with coefficient alphas of .77 for conscientiousness, .85 for emotional stability, .86 for extraversion, .79 for openness to experience, and .72 for agreeableness.

Separate scale scores were derived for each trait as the mean across corresponding items, with higher scores indicating higher levels of each trait. Scale score means were 4.87 (SD = 0.89) for conscientiousness, 4.03 (SD = 1.12) for emotional stability, 4.45 (SD = 1.13) for extraversion, 5.20 (SD = 0.78) for openness to experience, and 5.55 (SD = 0.73) for agreeableness. The distribution of scores on each trait was fairly normal, with the exception that agreeableness exhibited a moderate negative skew (skewness = -0.52, standard error = 0.17).

Biodata. The biodata measure that was used included adapted items from the instrument developed by Barrick and Zimmerman (2005, 2009) in employee selection contexts. As previously described, Barrick and Zimmerman's instrument assessed prior job changing behavior and social embeddedness in the organization (i.e., the number of pre-existing friends and relatives in the organization at the time of hire). Likewise, the biodata items in the present study were organized around the following main conceptual themes: *prior job changing behavior*, *prior university and course changing behavior*, *prior attendance behaviors at work and school*, and *social embeddedness* in the introductory psychology course. In addition, the following themes were examined on an exploratory basis using biodata items: *method/difficulty of one's daily commute* to the college campus and *intentions to withdraw*. A detailed mapping of items to each of the six biodata themes is shown in Table 5, as well as corresponding details regarding item scaling. All items are shown in appendices.

Prior job changing behavior. Prior job changing behavior was assessed by the following three items (labeled items 1, 2, and 3 in Appendix C): For how long have you worked in your current job? For how long did you work in your most recent job prior to the one you have now? How many jobs have you held in the past five years? For the first two items, participants were asked to indicate the number of years and months. Responses to these two items were summed to indicate the number of years and months spent in participants' current job and most recently held former job (*job tenure*). The mean for job tenure was 2.28 years (*SD* = 2.59) and ranged from 25 students with no job tenure to one student with 15 years of job tenure. Job tenure exhibited a significant positive skew (skewness = 2.05, standard error = 0.17). [2]

For the third item relating to prior job changing behavior (*number of jobs held*), participants were asked to report the number of jobs they had held within the last five years. The mean number of jobs held was 2.10 (*SD* = 1.51) and ranged from 25 students with no jobs held to one student with 10 jobs held. Number of jobs held also exhibited a significant positive skew (skewness = 1.11, standard error = .17).

The correlation between the number of jobs held and students' job tenure was .09 (p = .22). Due to the low correlation observed and the different response scales used (i.e., time versus frequency), *job tenure* and *number of jobs held* were maintained as separate variables for subsequent analyses.

Prior university and course changing behavior. Given the present study's context, a set of items were adapted to focus specifically on university and course changing behaviors. University changing behaviors were assessed by two items (labeled items 5 and 6 in Appendix C) focusing on the number of full-time universities and part-time universities attended since graduating high school. For both items, participants were asked to indicate the number of universities attended. Responses to these items were summed to indicate participants' total number of full- and part-time universities attended since beginning college (*universities attended*). The mean number of universities attended was 1.62 (*SD* = 0.95) and ranged from 125 students who attended only one university to one student who had attended a total of seven universities. Universities attended exhibited a significant positive skew (skewness = 2.18, standard error = 0.17).

Prior course changing behavior was assessed by two items (labeled items 7 and 8 in Appendix C) focusing on the number of high school courses and college courses dropped previously. For both items, participants were asked to indicate the number of courses they had dropped previously. Responses to these two items were summed to indicate the total number of high school and college courses dropped (*courses dropped*). The mean number of courses dropped was 1.32 (SD = 1.78) and ranged from 87 students with no dropped courses to one student with 13 dropped courses. *Courses dropped* exhibited a significant positive skew (skewness = 2.68, standard error = 0.17).

Finally, plans changing behavior was assessed by two additional items (labeled items 4 and 11 in appendix C) focusing on the number of times a participant changed their college major and their future career plans. For both items, participants were asked to indicate the number of changes. Responses to these items were summed to indicate the total number of major and career changes (*plans changed*). The mean number of plans changed was 3.12 (SD = 2.40) and ranged from 25 students who had never changed their major or career plans to two students who changed plans a total of 11 times. *Plans changed* exhibited a significant positive skew (skewness = 0.93, standard error = 0.17).

The number of courses dropped was significantly correlated with the number of universities attended (r = .31, p < .001) and number of times plans were changed (r = .32, p < .001). However, the number of universities attended was not significantly correlated with the number of times plans were changed (r = .09, p = .19). In light of the rather modest interrelationships among these factors, separate variables were maintained for *universities attended*, *courses dropped*, and *plans changed*.

Prior attendance behaviors at work and school. The following two items assessed prior attendance behaviors at work and school: About how many times have you been scolded or disciplined in previous jobs for showing up late or missing a scheduled work shift? Since the time that you started high school, about how many times have you been scolded or disciplined for showing up late or missing a class? For both items (labeled items 12 and 13 in Appendix C), participants responded by indicating the number of times. Responses were summed for these two items to indicate the total number of times students have been disciplined for attendance problems across work and school contexts (*attendance problems*). The mean for attendance problems was 3.22 incidents (*SD* = 4.12) and ranged from 57 students with no incidents reported to one individual with 25 incidents. *Attendance problems* exhibited a significant positive skew (skewness = 2.65, standard error = 0.17).

Social embeddedness. The biodata items assessing social embeddedness were modified to focus on students' social relationships within the introductory psychology course rather than in an organization, as in the two studies by Barrick and Mount (2005; 2009). Participants' social embeddedness was assessed by two separate items (labeled items 9 and 10 in Appendix C) focusing on the number of friends and familial relatives students had in the present course. For both items, participants were asked to indicate the number. Responses to these two items were summed to reflect the total number of friends and relatives each student had in the introductory psychology course (*friends and relatives*). The mean number of friends and relatives was 2.39 (*SD* = 2.65) and ranged from 56 students with no friends or relatives in the course to two students with a total of 15 friends and relatives in the course. *Friends and relatives* exhibited a significant positive skew (skewness = 2.00, standard error = 0.17).

Method and difficulty of commute. Two items assessed the difficulty of participants' daily commute to the college campus where introductory psychology was held (items 14 and 16 in appendix C): Approximately how far from WSU's campus do you live currently? How would you describe your commute to a friend who was considering living where you currently live? For the first item (commute *distance*), participants were asked to indicate the number of miles. The mean for *distance* in miles was equal to 12.19 (SD = 12.73) and ranged from 52 participants who lived on campus (and therefore indicated 0 miles) to one participant who commuted 80 miles. *Commute distance* exhibited a significant positive skew (skewness = 1.59, standard error = 0.17). For the second item (*commute difficulty*), participants were asked to rate the difficulty of their commute on a 4-point scale ranging from 1=*very easy* to 4=*very difficult*. The mean for *difficulty* was 1.74 (SD = 0.74), approximating a moderately easy commute on average, and ranged from 88 participants with a very easy commute to five participants with a very difficult commute. *Commute difficulty* exhibited a significant positive skew (skewness = 0.84, standard error = 0.17).

The correlation between ratings of commute difficulty and commute distance was positive and statistically significant (r = .39, p < .001). However, because the degree of interrelationship was moderate and because these two items had different scale properties, separate variables based on *commute difficulty* and *commute distance* were maintained for subsequent analyses.

One additional item assessed the method that participants' used to commute to campus (*commute method*): How do you usually get to campus? The multiple-choice response options for this item were: driving oneself, riding with a friend/family, walking, biking, or using public transportation (see item 15 in Appendix C). In total, 115 students drove themselves to campus, 51 walked, 34 got a ride with a friend/family, 10 used public transit, and 2 biked.

For subsequent analyses this item was coded in three ways, reflecting three somewhat different perspectives about how commuting method might affect attendance outcomes. One possibility is that attendance is interrupted by unreliable commute methods. In terms of the different modes of transportation that were examined, walking would seem the most reliable because it connotes proximity to the college campus and also because it is the least susceptible to mechanical failures and traffic incidents. As indirect support for this explanation, a one-way ANOVA was conducted with *commute method* as the factor and *difficulty* of commute (as described above) as the dependent variable. Results indicated that method of commute was significantly related to perceived difficulty, $F(4,207) = 5.90$, $p < .05$. Post-hoc comparisons based on Tukey's HSD test revealed that walking was perceived as significantly less difficult than either driving or riding with others. No other significant pairwise differences were obtained. Following these results and the rationale that walking is perhaps the most reliable method of getting to class on time, commute method was first coded dichotomously so that walking was assigned a 2 and all other methods a 1 (*commute method-reliable*).

From a slightly different perspective, methods also differ by how much independence they afford the commuter, both in planning when to leave, the route to take, and how to handle any situations that come up in route to campus. Along these lines, driving, walking, and biking give the individual commuter the greatest independence. In contrast, commuters are dependent on others' decision making and schedule when riding in someone else's car or taking public transportation. Following this rationale, commute method was also coded dichotomously so that walking, biking, and driving were assigned a 2 and all other methods a 1 (*commute method-independent*).

A third perspective that was examined involved the possibility that some methods of commuting might dissuade student attendance more readily under moderately inclement weather conditions such as rain. In particular, commuters that walk, bike, or use public transportation (which often involves some walking) may be less inclined to attend class when weather conditions are moderately poor. In contrast, rain or other mild inclement conditions would be unlikely to dissuade commuters that spend the majority of their commute sitting in a car. Although the opposite pattern might be anticipated as inclement weather becomes increasingly extreme (i.e., commuters that drive in a car would be less likely to attend in extreme weather), this is somewhat less plausible in the present study because data collection took place in the fall term (which is typically characterized by mild weather) and because the university is closed to all commuters in the event of extremely poor weather conditions. Thus, the third coding scheme that was explored assigned a 2 to commuters that either drive themselves or ride in a car with others versus a 1 to commuters that walk, bike, or use public transit (*commute method-weather*).

Examining the interrelationships among dichotomously coded commuting method variables revealed significant correlations for *commute method-reliable* and *commute method-independence* (r = .29, p < .001) and for *commute method-reliable* and *commute method-weather* (r = -.86, p < .001) versus a non-significant correlation for *commute method-independence* and *commute method-weather* (r = -.06, p = .36). To further consider potential redundancy, the interrelationships among commuting method variables, *commute distance*, and *commute difficulty* were also examined. The strongest correlations were observed for *commute method-weather* and *commute distance* (r = .56, p < .001) and for *commute method-reliable* and *commute distance* (r = -.54, p < .001). Based on this pattern of interrelationships, *commute method-weather* and *commute method-independence* were ultimately retained for further exploratory analyses, whereas *commute method-reliable* was dropped due to being highly correlated with *commute method-weather* and having the strongest overlap with *commute difficulty* (r = -.29, p < .001).

Withdrawal intentions. Intentions to withdraw from the course were assessed with three items (items 7, 8, and 9 in Appendix D): How frequently do you expect to miss the introductory psychology class this semester? How frequently do you expect to be late to the introductory psychology class this semester? How likely are you to drop this introductory psychology course this semester? For the first two items, participants were asked to indicate their expected frequency using a 5-point scale ranging from 1=*never* to 5=*very frequently*. The correlation between intention to miss class and intention to be late for class was positive and statistically significant (r = .45, p < .001). The mean of responses to these two items was computed to reflect the judged frequency of day-to-day withdrawal behaviors (*intentions-daily withdrawal*). The mean for *intentions-daily withdrawal* was equal to 1.65 (*SD* = 0.71) indicating that, on average, students expected to be late or absent between *never* and *rarely*. Responses ranged from 70 students who expected to never be late or absent to one student who expected to be late and absent very frequently. *Intentions-daily withdrawal* exhibited a significant positive skew (skewness = 1.70, standard error = 0.17).

For the third item (*intention-drop*), participants were asked to indicate their expected likelihood using a 5-point scale ranging from 1=*extremely unlikely* to 5=*very likely*. The mean for *intention-drop* was 1.36 (*SD* = 0.71) and ranged from 157 students who indicated that it was extremely unlikely they would drop the course to one student who indicated it was very likely she would drop the course. *Intention-drop* exhibited a significant positive skew (skewness = 1.52, standard error = 0.17).

Background questionnaire. An additional questionnaire was included to assess participants' age, sex, and ethnicity for descriptive purposes (see Appendix D). In addition to the demographic items, several exploratory items were included to assess current status at the university (e.g., year in school and full-time versus part-time status) and expectations regarding withdrawal behaviors during the semester (e.g., expected frequency of absence from class and likelihood of dropping the class).

Behavioral withdrawal criteria. Criteria were developed to assess several types of student withdrawal behaviors during the introductory psychology course. Specifically, the following types of withdrawal behaviors were examined: (a) permanent withdrawal from the course (i.e., a "drop"), (b) lecture absenteeism, (c) lab absenteeism, (d) lateness to lectures (frequency and time lost indices), (e) lateness to labs, and (f) early departures from lectures. Permanent withdrawals were single event occurrences in the present context – that is, students could not permanently withdraw and resume participation in the course at a later time. Episodes of lateness, absenteeism, and early departure were aggregated within-persons and across the days of the semester. Table 3 shows the lecture and lab schedules and indicates the specific days that were aggregated to form the absenteeism, lateness, and early departure criteria.

Separate criteria were maintained for absenteeism and lateness behaviors in the lecture versus the lab because there were several important differences between the two domains. For example, whereas no official attendance policy was used for the lecture section, students were penalized for arriving late or failing to attend lab meetings (e.g., reductions to homework points). The meeting times and locations also differed, with lectures generally meeting twice weekly for all 16 weeks of the semester and labs generally meeting once weekly for 12 weeks of the semester. The settings and potential social pressures for attendance also differed. Lectures were held in a large auditorium with an approximate capacity of 500, whereas labs were held in classrooms with small groups of students generally ranging from 15 to 30. Finally, the modes of data capture were different in the two domains (clickers in the lecture versus lab instructors in the labs), and are therefore subject to different sources of error (e.g., clicker malfunction versus instructor leniency).

Consistent with previous studies (e.g., Foust, Elicker, & Levy, 2006), withdrawal criteria were specified as a proportion based on the total number of withdrawal episodes divided by the total number of opportunities for withdrawal. For example, lateness to lectures was computed for each student as the total number of lateness episodes divided by the total number of lectures attended. Deriving a proportion in this manner accounts for the nesting or non-independence of different types of withdrawal behaviors. In other words, this adjusts for the fact that it is not possible for a student to be late (or depart early) on a day when the student was absent and helps to differentiate lateness and absenteeism as separate types of behaviors. Similarly, absenteeism was computed as a proportion of the total lectures/labs that could be attended taking into consideration students' withdrawal dates. In this way, each withdrawal behavior represented a frequency estimate – that is, how regularly a withdrawal behavior was exhibited by the student – with the exception of lateness to lectures, for which separate frequency and time lost estimates could be examined.

Permanent withdrawal. Permanent withdrawal was coded dichotomously to indicate whether students completed or withdrew from the introductory psychology course (course completion = 0, course withdrawal = 1). Specifically, permanent withdrawal was defined as a student terminating all aspects of her participation in a course before the course's completion. Permanent withdrawal was inferred when a student's pattern indicated early termination or when a formal withdrawal request appeared in the course grade book. In total, 30 students withdrew from the course, including 18 who completed a formal course withdrawal request and 12 who terminated their participation without completing a formal withdrawal request. In contrast, 183 students completed the course.

A specific permanent withdrawal date was established as the student's last act of participation in the course (e.g., the last assignment/exam completed or the last lecture/lab attended). This was determined by triangulating across several data sources, including students' attendance records for the lectures and labs, students' use of clickers in the lecture, and grade book records indicating completion of course assignments and exams. The course ran for a total of 106 days. For students who permanently withdrew from the course, the number of days in the course prior to withdrawal ranged from 26 for one student to 98 for four students, with a mean of 76.13 days. The median date of withdrawal was November 18th or approximately one month before the semester's conclusion. For survival analysis purposes, a timing variable was derived as the number of days from the first day of class to the withdrawal event.

Lecture absenteeism. An absence was defined as a whole lecture session missed. Absenteeism was computed as a frequency index based on the proportion of lectures that each student missed during the semester. Specifically, for students that permanently withdrew from the course, absenteeism was derived as the number of sessions missed divided by the total number of lectures that were held prior to the student's permanent withdrawal date. For students that completed the course, absenteeism was derived as the number of sessions missed divided by 26.

Lecture absenteeism was coded from students' use of clickers in the classroom during lectures (see http://www.iclicker.com/dnn/). Clicker questions were posted by the lecturer within the context of his PowerPoint presentations to the class. After each question was posted, students were given between 30 seconds and one minute to respond by using their clicker to select one of several response options. Individual-level response data were captured and stored using the iclicker software and database tools. The response data collected included the exact time at which a student's response was registered, as well as the specific response option selected for each question. An absence was inferred from a lack of responding to all clicker questions posted within a lecture session.

As shown in Table 3, clickers were used during 26 of the 29 lectures held during the semester. The number, type, and timing of clicker questions varied across lectures. On average, 5.35 clicker questions were posted per lecture session; the most for any single session was 10 questions, and the least was one question. Because it could be suggested that days in which only a single clicker question were used provide a relatively less reliable indication of student absenteeism (i.e., because missing a single question connotes whole-session absence rather than a pattern of missingness across several questions throughout the session), it is important to describe the context for these particular sessions. Three sessions were exam days, in which students were required to use the clickers to provide their acknowledgment of the academic integrity policy for exams, for which students were able to register a response for the full session (i.e., with no time limit). One additional day involved a single peer instruction question (see below), which was posted approximately halfway through the lecture session (i.e., 52 minutes after the session started). Because of the timing and context for these specific instances, the four sessions in question were retained for the aggregated index of lecture absenteeism.

Clicker questions were generally dispersed throughout the duration of the lecture sessions, with the earliest question posted 2 minutes and 23 seconds after the start of lecture and the latest question posted 91 minutes and 3 seconds after the start of lecture. Twelve sessions included a 5-item quiz shortly after the start of lectures (e.g., 2 – 5 minutes). Quiz scores were based on the number of correct responses. Other types of questions included peer instruction and general participation. Peer instruction questions were dispersed throughout the lectures and were generally administered on an alternating basis with the quizzes (i.e., on non-quiz days). Peer instruction questions provided students with the opportunity to communicate with other students about their responses prior to registering a final decision. General participation questions included opinion polls and opportunities for students to provide feedback on various topics. Peer instruction and general participation questions provided an opportunity for participation points but were not scored in terms of correctness.

Use of the clicker data to infer student attendance behaviors rests on two noteworthy assumptions: (a) a failure to respond indicates physical absence, and (b) responses were only made by students using their own pre-registered clicker and identification number. There are several reasons to speculate that these assumptions were reasonably met in the present context. First, a clicker response could only be registered from the classroom. In other words, “remote” responding was not possible. Second, acquiring and registering a clicker was a course requirement, as specified in the syllabus and reinforced by the instructor. Moreover, 200 points in the course involved in-class use of the clickers, including successfully registering the clicker at the start of the semester and using the clicker for in-class quizzes and peer instruction activities thereafter. These 200 points accounted for just under 20% of students’ grade for the course. Third, students were encouraged to respond to clicker questions even if their response was a “best guess,” as there was no penalty associated with incorrect guesses (e.g., on quizzes). Therefore, it was unlikely that students present in the classroom would choose not to respond. And fourth, the instructor brought several extra clickers to each session in the event that a student forgot or lost his or her clicker. On those occasions, the student was assigned a temporary identification number that could be linked back to their permanent identification number at a later time.

A related concern involves the possibility of students committing “clicker fraud” such as responding on behalf of an absent friend. The instructor took several precautions to minimize this possibility, including clearly prohibiting and detailing serious negative consequences for students caught using others’ clickers and by having teaching assistants count the number of students physically present at each lecture meeting. On no occasion was the number of clicker responses greater than the number of students physically present at the lecture.

Lab absenteeism. Lab absenteeism was based on students’ attendance records from the course grade book. There were 12 lab meetings during the semester, with attendance constituting 30% of students’ grade in the lab and roughly 8% of students’ grade in the course. Lab instructors used sign-in sheets to keep track of student attendance on a weekly basis and periodically updated the course grade book throughout the semester. Students were marked absent if they missed an entire lab session. Using the same procedures as described for the lecture, lab absenteeism was computed as a frequency index based on the proportion of labs that each student missed during the semester.

Lateness to lecture. Lateness was defined as arrival after the start of the lecture. Episodes of lateness were coded from students’ use of the clickers, as indicated by the pattern of responding (and non-responding) across chronological questions within a lecture session. Lateness is inferred by a pattern in which one or more non-responses occur at the beginning of the sequence. For example, considering five chronological clicker questions, each of the following example response patterns would be coded as episodes of lateness, with *O*s indicating non-responses and *X*s indicating responses:

Student 1: O, X, X, X, X

Student 2: O, O, X, X, X

Student 3: O, O, O, X, X

Student 4: O, O, O, O, X

Frequency and time lost indices were computed, with frequency equal to the number of lateness episodes divided by the total number of lecture sessions attended, and time lost equal to the average duration of lateness episodes. It was possible to estimate time lost due to lateness based on the time associated with a students' first response. For illustration, the same example response patterns are repeated below but with the time of response also shown in terms of minutes and seconds after the start of lecture (i.e., in *mm:ss* format):

Student 1: O, X (02:33), X (3:56), X (5:01), X (6:54)

Student 2: O, O, X (3:56), X (5:01), X (6:54)

Student 3: O, O, O, X (5:01), X (6:54)

Student 4: O, O, O, O, X (6:54)

Using the method specified above, time lost would be equal to 2 minutes and 33 seconds for student 1, 3 minutes and 56 seconds for student 2, and so on. It should be noted that in an absolute sense this results in a potential overestimate of the actual time that each student was late because the time of first response provides an upper bound estimate of when the student arrived to lecture. For example, although it can be inferred that student 3 was present in lecture at 5:01 and not at 3:56, his actual arrival time is assumed to occur between the two times. Nevertheless, the first evidence of the student's presence in the lecture is indicated by his response at 5:01. At the same time, this method is not sensitive to lateness occurring prior to the first clicker question. That is, students who arrived after the lecture's start but before the first clicker question were not coded as late.

Only a subset of the 26 days on which the clickers were used provided information that was useful for coding student lateness behavior. In particular, 16 days in which the clickers were used within the first 15 minutes of lecture were identified as relevant for coding student lateness behavior (see Table 3). Alternatively, days in which the first clicker question was posted after the first 15 minutes of lecture were excluded. Furthermore, only the clicker data from the first 15 minutes of these 16 lectures was considered when forming the aggregate indices of lateness. In other words, these indices (frequency and time lost) capture the lateness behaviors that occurred within the first 15-minutes of lecture. This was done for three reasons. First, definitions of lateness have focused on relatively short-term tardiness at the beginning of work shifts and have suggested that longer-term forms of lateness (e.g., partial session absences) represent a different type of behavior with different potential causes. In this way, limiting the days included to those with clicker questions in the first 15 minutes ensured that the aggregated frequency measure was based on similar observations of lateness behavior at the episode level. Second, this ensured that the time lost measure was a meaningful index of average lateness behavior for each student and not simply a function of one or two extreme lateness episodes. And third, 15 minutes was a meaningful threshold for lateness behavior in the present context. Missing the first 15 minutes of lecture meant that a student typically missed all of the points for a quiz. Similarly, missing the first 15 minutes in the lab meant that a student lost attendance points for the day.

Lateness to lab. Lateness data was available for 9 of 15 lab sections. In each of these labs, instructors marked students late if they arrived more than 15 minutes after the start of lab. The lab instructors for the remaining sections either changed their policy mid-semester or did not track and penalize students' lateness. As a result, the frequency of lab lateness could be derived for 134 of 213 study participants in the baseline hypothesis testing sample. Frequency of lateness to lab for each student was computed as the number of lateness episodes divided by the total number of labs attended.

Early departures. Early departure from lecture – that is, leaving the classroom before the lecture had ended – was examined on exploratory basis as a final form of withdrawal behavior. Based on anecdotal evidence from the instructor, this form of withdrawal behavior was particularly evident on quiz days when a subset of students would leave shortly after completing the quiz. These occurrences of early departure represent a potentially detrimental form of withdrawal behavior since leaving just after completion of the quiz meant that students missed a substantial amount of new material each week, as well as several opportunities for participation points.

Early departure was derived as the proportion of lectures attended in which the student left before the end of the lecture. Because early departure was captured from students' use of the clickers, it was necessary to operationalize the "end of lecture" as the last clicker question posted for each session. On average, the last clicker question on quiz days was posted at 67 minutes in comparison to the total length of lecture, which was 85 minutes. A pattern of clicker responding in which there was one or more non-responses at the end of a session was coded as an early departure. Accordingly, the example response patterns shown below would be coded as episodes of early departure, with *O*s indicating non-responses and *X*s indicating responses. Note that student 4's pattern is indicative of both a lateness and early departure episode within the same lecture session.

Student 1: X, X, X, X, X, X, O

Student 2: X, X, X, X, X, O, O

Student 3: X, X, X, X, O, O, O

Student 4: O, O, X, X, X, O, O

Procedure

Separate recruitment and study procedures were maintained for the predictor and criterion collection portions of the study. Recruiting for the predictor portion of the study was conducted by the principal investigator. The principal investigator visited four introductory psychology lectures during the first two weeks of the semester and visited every lab section at least once during subsequent weeks. During the lecture and lab visits, a recruiting script was read aloud. Students were provided with generic information about the study and their role as a participant. Students were invited to participate in exchange for 1 hour of research credit for the course. Students were not informed that their participation in the study would be linked to other research studies including the separate criterion collection portion of the present study.

The recruiting message and consent process for the criterion collection portion of the study was handled by a different researcher, who also attended lectures during the first two weeks of the semester but on two separate occasions than the principal investigator. This researcher informed students about a different research study that was investigating students' use of i-clickers in the classroom and how this relates to performance in the course (e.g., student grades). Participation did not require anything from students beyond simply allowing the researchers to access their clicker data and the course grade book at the conclusion of the semester. After informing students about the purpose of the research study, the researcher posted a clicker question to the classroom that asked students whether they wished to participate in the study. Students provided consent using their clicker to indicate yes or no. No research credit was provided in exchange for participating in this portion of the study.

All predictor measures, including the CRT-W, were administered during group testing sessions that included between 1 and 34 participants. Participants signed up for a testing session using the online research system or in person during one of the principal investigator's visits to lectures and labs. Testing sessions were held in small classrooms located near the building where lectures were held. All aspects of the study were administered by the principal investigator, with one or two research assistants acting as proctors during administration of the CRT-W.

All sessions were administered according to the protocol shown in Appendix E, which included written verbal instructions that were read aloud by the principal investigator. Upon arrival to the testing session, participants were first asked to complete the informed consent document. Materials were administered using paper-and-pencil questionnaires contained in two packets. The first packet included the CRT-W. Prior to distributing the CRT-W, participants were encouraged to do their best and were instructed that the students achieving the highest 50 scores on the test would receive $20.00 gift certificates. Gift certificates were distributed on the basis of students' scores on the five real inductive items embedded in the measure. Because 174 participants (68.8%) correctly answered all five items, gift certificates were distributed to 50 of these individuals using random selection.

Next, participants were informed that they would have 30 minutes to complete the CRT-W. However, in actuality, participants were given as much time as needed to complete the measure. In each session, the researcher announced when students had 15 minutes, 10 minutes, and finally 5 minutes remaining, but did not announce when time was up, instead providing students with additional time if necessary to complete all items on the measure. Moreover, the instructions encouraged participants to work through the items in a timely fashion and that they would not be penalized for incorrect guesses. The announced time limit of 30 minutes was chosen based on initial pilot testing with the CRT-W and prior research with the CRT-A which has reported that test takers complete approximately one conditional reasoning item per minute (James & LeBreton, 2011). The mean time required to complete the CRT-W was 25 minutes, and only 14 participants (3.6%) exceeded 30 minutes.

Upon completing the CRT-W, participants were handed the second packet of materials, which contained the personality measure, biodata items, and a questionnaire containing items regarding participants' demographics and educational background. Instructions for all aspects of these measures were self-contained in the packet.

Participants were simply told to read the instructions and complete the items at their own pace. Participants typically completed the second packet of materials in approximately 15-20 minutes. Upon turning in the second packet, participants were thanked for their participation and informed that additional details about the study would be communicated to them via an email from the researcher at the completion of the study (i.e., after the semester's conclusion). Debriefing participants in this manner was deemed necessary in order to prevent students from revealing the underlying purpose of the CRT-W to other prospective participants and to prevent students from modifying their classroom attendance behaviors during the remainder of the semester. In total, all aspects of the testing sessions were completed within 60 minutes.

Data on students' withdrawal behaviors accrued continuously throughout the semester. As described in detail previously, withdrawal criteria were derived from the course grade book and students' use of clickers during lectures. At the conclusion of the semester, the principal investigator extracted withdrawal data for those students who consented to the criterion portion of the study. Predictor and criterion data were subsequently matched using student identification numbers.

CHAPTER 3

RESULTS

Results are presented in two major sections. The first section reports the results of scale development and refinement analyses. Examinations of item and scale reliability and validity provided the basis for eliminating the subset of items with the poorest psychometric properties and lowest predictive validities. The second section reports the results of hypothesis testing. Using the revised CRT-W, hypothesized relationships between test scores and behavioral withdrawal criteria are examined, as well as the incremental validity of test scores above and beyond previously validated predictors of withdrawal.

Scale Development and Refinement Analyses

Prior to scale development analyses, data screening procedures were conducted following the recommendations of James and colleagues (e.g., James & Mazerolle, 2002; James & McIntyre, 2000). The first step involved identifying and removing participants with incomplete or inappropriate response patterns on the conditional reasoning items. In addition to illogical distractor response options, missing item-level data were apparent for two participants, each of whom omitted a response to a single conditional reasoning item. Because these participants only missed a single item from the test, they were not eliminated from subsequent analyses. For item analysis, all missing item-level data were handled within the scope of pairwise deletion – that is, removing the participant only for the affected item rather than from all subsequent analysis.

Next, per James and McIntyre's (2000) recommendation, participants who endorse an unusually high number of illogical responses to the conditional reasoning items were removed. For the 22-item CRT-A, these authors suggested a cut-off of 5 or more illogical response options, which is indicative of a careless response pattern. In three studies described by LeBreton et al. (2007), applying this cut-off has resulted in dropping 7 participants from a total sample of 558 (1.3%), 9 from a total sample of 109 (8.3%), and 15 from a total sample of 966 (1.6%). In the present study, the cut-off was set at 7 or more illogical response options to reflect the length difference of the CRT-W in comparison to the CRT-A (i.e., 30 items versus 22). Specifically, this resulted in dropping 9 participants from the baseline scale development sample of 253 (3.6%). Among the participants dropped, the mean number of illogical responses endorsed out of 25 conditional reasoning items was 10.56 (*SD* = 5.22) in comparison to a mean of 1.14 (*SD* = 1.42) for the remainder of the sample.

The length of time participants took to complete the CRT-W was also considered as a potential screening mechanism. This resulted in the removal of one additional participant who completed the CRT-W in 12 minutes and also endorsed 6 illogical responses (i.e., one less than the cut-off of 7). For comparison, the median time required to complete the CRT-W for all participants was 25 minutes (*M* = 24.63, *SD* = 4.29). After dropping the 10 participants identified to this point, the total remaining sample included 243 participants. This sample served as the basis for examining item response characteristics, item-total correlations, and inter-item correlations (see Reliability and Validity Analyses below).

Data screening of behavioral withdrawal variables. Because a substantial portion of the scale development analysis involved estimating item-criterion correlations, it was also necessary to examine the behavioral withdrawal variables for potential univariate outliers and normality. This was done using the baseline sample for hypothesis testing (N = 213), comprising participants that consented to both the predictor and criterion collection portions of the study. Unlike the behavioral withdrawal criteria, it was not possible to examine outliers and consider potential transformations for the conditional reasoning items which were scored dichotomously (0-1). It was however, possible to examine item response characteristics more generally within the scope of scale refinement. Three main factors were considered when screening the behavioral withdrawal criterion variables: the presence of univariate outliers, skewness, and interrelationships among the variables.

Outlier analysis. First, outlier screening was conducted by standardizing the behavioral withdrawal variables, excluding permanent withdrawal which was a dichotomous variable. In total, 10 participants with standardized values exceeding ±3.29 on one or more withdrawal variables were flagged for closer examination. Tabachnick and Fidell (2007) recommended ±3.29 as a conservative basis for identifying values as having a very low probability of deriving from the population of interest (p < .001). Although each displayed relatively high levels of withdrawal, only three were ultimately removed as outliers using pairwise deletion.

For lecture absenteeism, the overall mean was .18 (SD = .16) or just over 4 of 26 lectures missed. One participant had a z-score of 3.41 corresponding to a proportion of .73 lectures missed. Specifically, this participant missed 8 of 11 lectures prior to a permanent withdrawal from the course. Although this participant was flagged as a potential outlier, closer examination of the frequency distribution revealed that the participant's attendance was not unusual or clearly separated from the remainder of the distribution. For example, the next highest proportion of lectures missed was .69. Therefore, this participant was not considered an outlier.

For lab absenteeism, the overall mean was .08 (*SD* = .13) or just under 1 of 12 lab sessions missed. Three participants had z-scores greater than 3.29, including one who missed 83% of labs (z = 5.62), one who missed 67% of labs (z = 4.41), and one who missed 58% of labs (z = 3.76). The next highest proportion of labs missed was .50. Examination of the frequency distribution suggested that the participants who missed 67% and 58% of labs were not clearly separated from the distribution, whereas the participant who missed 83% of labs did appear visually as a potential outlier. This participant missed 10 of 12 labs and submitted a formal withdrawal from the course during the last week of the semester. However, there was no indication that the participant's data was suspect or that the attendance record was spurious. Therefore, it was concluded that the participant indeed had very poor attendance to lab but should not be considered an outlier.

For lecture lateness frequency, the overall mean was .05 (*SD* = 0.10) indicating that, on average, students were late to 5% of lectures attended. Four participants had z-scores greater than 3.29, including one who was late to 67% of lectures attended (z = 6.33), one who was late to 57% of lectures attended (z = 5.35), one who was late to 45% of lectures attended (z = 4.15), and one who was late to 44% of lectures attended (z = 4.04). The next highest proportion for lecture lateness was .33. Further scrutiny indicated that the most extreme participant score (*z* = 6.33) represented an outlier and was eliminated from all subsequent analyses involving lecture lateness frequency. Several factors led to this judgment. First, whereas the other three participants flagged all had greater than 4 episodes of lateness behavior and a consistent pattern of higher than average withdrawal across the other behavior types assessed, this participant had only two episodes of lateness (out of three lecture sessions attended) and had good attendance records on the other behaviors assessed. Most problematic was the limited time window available to observe this participant's lateness behavior, having attended only three lecture sessions prior to formally withdrawing from the course in the third week of the semester. In contrast, the remaining three participants remained in the course for at least 75 days, allowing for a more reliable assessment of their lateness behavior. Consequently, this participant was removed using pairwise deletion from further analyses involving lecture lateness frequency but was included in analysis of other withdrawal behaviors including permanent withdrawal from the course.

For lab lateness, the overall mean was .05 (*SD* = 0.09), indicating that on average students were late to 5% of the labs they attended. Two participants had *z*-scores greater than 3.29, including one who was late to 56% of labs attended (*z* = 5.50) and one who was late to 50% of lectures attended (*z* = 4.90). The next highest proportion for lab lateness was .33. Further examination of the data for these participants did not suggest a spurious pattern; both displayed higher than average frequencies for the other withdrawal behaviors assessed and completed the course providing the full semester as a basis for observing their lateness behavior. Therefore, it was concluded that neither participant should not be considered an outlier despite very high frequencies of lateness behavior.

For early departures from the lecture, the overall mean was .11 (*SD* = 0.15) indicating that, on average, students left approximately 11% of attended lectures before the final clicker question had been posted. Three participants were flagged as potential outliers, two who left early on 67% of attended lectures (z = 3.84) and one who left early on 63% of attended lectures (z = 3.55). The next highest proportion for early departures was .55. Further examination revealed that the two most extreme participants left early on two of three of the lectures when it was possible to code for early departures. In contrast, the third participant left early on five of eight of the lectures when it was possible to code for early departures. The reduced number of opportunities for early departures was a result of two factors. First, both participants had higher than average frequencies of lecture absenteeism (.61 and .69). Second, one of the participants permanently withdrew from the course at 68 days (compared to the 106 days total) further reducing his opportunity for early departures. As a result, neither participant's high frequency is based on a comparatively robust sample of early departure episodes and opportunities. Consequently, both participants were removed using pairwise deletion from further analyses involving early departures but were included in analysis of other withdrawal behaviors.

Skewness. Each of the behavioral withdrawal variables exhibited a strong positive skew: permanent withdrawal (skewness = 2.08, standard error = 0.17), lecture absenteeism (skewness = 1.25, standard error = 0.17), lab absenteeism (skewness = 2.58, standard error = 0.21), lecture lateness frequency (skewness = 2.71, standard error = 0.17), lecture lateness time lost (skewness = 1.34, standard error = 0.17), lab lateness (skewness = 2.88, standard error = 0.21), early departures (skewness = 1.41, standard error = 0.17). Tabachnick and Fidell (2007) recommended ±3.29 as a conservative basis for identifying deviations from normality with very low probabilities of occurring simply due to chance ($p < .001$). Examination of the histograms for these variables pointed to generally low base rates for the withdrawal behaviors as the general reason for their positive skewness. That is, a large proportion of the students either did not exhibit the withdrawal behavior, or did so at a very low frequency, and the remaining students formed the tail of the positively skewed distribution. Counting those participants who exhibited one or more episodes of temporary withdrawal, the base rates observed were 31.3% for lab lateness, 36.8% for lecture lateness, 38% for lab absenteeism, 46.9% for early departures, and 82.2% for lecture absenteeism. The base rate for permanent withdrawal from the course was 14.1%.

With the exception of permanent withdrawal, transformations were attempted to determine if the normality of these variables could be restored for analytic purposes. Consistent with Tabachnick and Fidell's (2007) recommendations, square root, logarithmic, and then inverse transformation was attempted, with the logarithmic transformation $log(x+1)$ consistently resulting in the greatest reduction in positive skewness. However, due to the high number of ties at 0 for each variable, strong positive skewness remained after transformation; skewness was 1.01 for lecture absenteeism (versus 1.25), 2.66 for lab absenteeism (versus 2.58), 2.64 for lecture lateness frequency (2.71), 0.75 for lecture lateness time lost (versus 1.34), 2.47 for lab lateness (versus 2.88), and 1.32 for early departures (versus 1.41). Corresponding histograms for the original and transformed variables are shown in Figures 3, 4, 5, 6, 7, and 8. Because transformations had little impact on restoring normality to these variables' distributions and also because transformation had little effect on predictive relationships (see below), the un-transformed variables were ultimately retained for analytic purposes.

Interrelationships. The interrelationships among behavioral withdrawal criterion variables were examined as a final check prior to scale analyses. This was done for two main reasons, first to verify that behavioral withdrawal criteria were positively intercorrelated across participants, and second, as a check for potential redundancy among the variables. Intercorrelations of the behavioral withdrawal variables are presented in Table 6. Consistent with expectations, the average intercorrelation among withdrawal behaviors was positive and small-to-moderate in magnitude (mean r = .24). Two other noteworthy patterns provided further insights into the data's quality, supporting the progression of withdrawal pattern that has been observed in prior studies (Herzberg et al., 1957) and underscoring the consistency of withdrawal behaviors across the contexts studied in the present study. Specifically, permanent withdrawal was more highly correlated with absenteeism than with lateness behavior (mean r = .40 versus .16), and second, same-type behaviors (e.g., absenteeism-absenteeism) were more highly correlated across the lecture and lab contexts than different-type behaviors (e.g., absenteeism-lateness) within lab or within lecture (mean r = .45 versus .27). These patterns generally support the plausibility of the withdrawal data collected and provide one indication of the comparability of withdrawal behaviors in the academic context with studies from non-academic contexts.

Beyond these general patterns, the correlation between lecture lateness frequency and lecture lateness time lost (r = .61; *rho* = .94) was observed to be moderately to highly correlated. Consequently, the decision was made to exclude the lecture lateness time lost variable from further analyses. Although this variable carried potentially unique information in comparison to the remaining *frequency* measures of withdrawal, it was the author's judgment that this variable was somewhat less trustworthy as an index of behavior due to: (a) the somewhat higher number of judgment calls leading up to its computation, and (b) the discontinuous nature of the measurement of time using unequally spaced clicker responses.

CRT-W reliability and validity analyses. Consistent with the recommendations of James and LeBreton (2011), item and scale reliability and validity were evaluated using the following considerations: (a) item response characteristics, (b) item-total correlations, (c) item intercorrelations, (d) item-criterion correlations, and (e) internal consistency reliability estimates. Following from the first four considerations, several items from the initial scales were dropped to arrive at a revised CRT-W. Finally, internal reliability estimates for the initial and revised scales were compared.

Item response characteristics. Item response characteristics were examined within the full participant sample (N = 243) and are summarized in Table 7. Item p-values indicate the base rate of participants endorsing the JM response option. Based on the work of James and colleagues (2004; 2005) with the conditional reasoning test of aggression and also because the base rates of actual withdrawal behaviors was low in the present study, it was generally expected that items would exhibit low base rates of JM endorsement (i.e., less than .50). However, because no prior studies have examined these particular items or JMs, it was not possible to evaluate observed p-values against a firm pre-specified criterion. Instead, p-values were examined on an exploratory basis, and items with extreme base rates or p-values that differed substantially from the other items were flagged for further consideration.

The average p-value across all items was equal to .46. This is a higher average base rate than was reported for the CRT-A, which had an average p-value across 22 items equal to .18 (James & LeBreton, 2011). P-values also varied widely between items, ranging from item 4 from the social injustice JM (p-value = .08) to item 2 from the social injustice JM (p-value = .99) and item 10 from the revocable commitment JM (p-value = .98). In total, 16 of 25 items had base rates at or below .50. The lowest base rates (less than .25) were observed for items 5 and 8 from the marginalization JM, items 1, 2, 5, and 7 from the revocable commitment JM, and items 4 and 6 from the social injustice JM. The items that differed the most from these (e.g., p-values greater than .70) were items 1, 3, and 9 from the marginalization JM, items 8 and 10 from the revocable commitment JM, and item 2 from the social injustice JM.

Two items that clearly stood out with extreme p-values were item 10 from the revocable commitment JM and item 2 from the social injustice JM. Both items were heavily slanted toward the JM response, with only 2% and 1% of respondents endorsing the non-JM response, respectively. A re-examination of the items revealed different possible explanations for high JM endorsement. Item 10 from the revocable commitment JM asked respondents to identify the most reasonable explanation for a difference between parenting styles of the past and present generation. The competing solutions presented offered different negative evaluations of past generation parenting (i.e., for undervaluing the importance of being well-rounded) and current generation parenting (i.e., for not reinforcing "stick-to-it-iveness"). In this case, participants' overwhelming preference to evaluate past generation parenting negatively (and in doing so, endorse the JM response) may be due to a generation bias inherent to the present sample of college students. Alternatively, the use of the colloquialism "stick-to-it-iveness" may have been misunderstood or made this option a less appealing alternative.

Item 2 from the social injustice JM offers competing solutions to an incentive system that de-motivates more senior professors by giving larger salaries to less experienced, newly hired professors. The JM response option proposes the use of a salary cap for newly hired professors in contrast to the non-JM response option which proposes disciplining senior professors in order to restore their motivation. Although either intervention has a logical basis, the JM response might be described as a more direct solution to the problem as stated. In other words, the most logical solution to a problem with salaries is to focus on the salaries themselves rather than new modes of discipline. Alternatively, participants may have viewed the disciplinary solution as unfair because it proposed further penalizing the senior professors even though they are already the "victim" of a salary disadvantage.

These observations suggest ways that the items might be improved for future studies. Specifically, it might be possible to decrease the p-value of these items by eliminating a potential generational confound in item 10 from the revocable commitment JM and by writing a more logical and less punitive non-JM solution for item 2 from the social injustice JM. Nevertheless, for the purposes of the present investigation, these items have near zero variance greatly inhibiting their ability to correlate with an external criterion or with the other items comprising the CRT-W. Thus, the decision was made to eliminate these items from all subsequent analyses.

The percentage of participants endorsing a distractor response option was generally low across items (e.g., less than 5%) and ranged from 0% for item 6 from the revocable commitment JM to 14.8% for item 4 from the marginalization JM. The high percentage of participants endorsing a distractor option for the latter item appears to be somewhat problematic given that the second highest distractor endorsement percentage was roughly half of this (i.e., 7.8% for item 5 from the marginalization JM). Further examination revealed that responses were evenly split between the two distractor options, suggesting that in general, this item could be improved by writing distractor options that are more clearly distinguishable from the content of the item's premise. Alternatively, it is possible that the JM and non-JM response options for this item had lower inductive validity than the other items in the test. This item was flagged for further consideration.

Item-total correlations. Consistent with James and LeBreton's (2011) recommendations, biserial item-total correlations were examined. A biserial correlation adjusts for an artificial dichotomy in one of the two variables correlated, essentially estimating what the correlation would be had the measure reflected the variable's underlying continuous nature. In other words, estimation of the biserial correlation assumes the dichotomized variable is continuous and normally distributed. As discussed by James and LeBreton, biserial correlations, in comparison to point-biserial, provide more stable estimates and are less sensitive to attenuation when the dichotomized variable has an extreme response distribution (i.e., p-values that depart widely from .50). In contrast, the point-biserial correlation systematically underestimates the magnitude of the correlation as the dichotomized variable departs from a 50-50 split. Biserial correlations were derived by correcting point-biserial correlations using the formulas presented by Cohen, Cohen, West, and Aiken (2003).

Corrected biserial item-total correlations are shown in Table 8. *Corrected* item-total correlations are based on an adjustment of the total score by taking the mean of all item scores excluding the item entering into the item-total correlation. The average corrected item-total biserial correlation was .09 for the marginalization items (versus .01 for the remaining items with the marginalization total), .04 for the revocable commitment items (versus -.01 for the remaining items with the revocable commitment total), and .08 for the social injustice items (versus .01 for the remaining items with the social injustice total). This pattern of relationships indicates that the items tended to correlate somewhat more highly with their corresponding JM item-total scores than they do with the other JM item-total scores. Nevertheless, this evidence can be considered weak given that the average item-total correlations are generally low.

However, somewhat stronger evidence is available for a subset of items within each JM. In particular, positive and statistically significant ($p < .05$) item-total correlations were observed for items 1, 3, and 8 from the marginalization JM, items 1, 3, and 9 from the revocable commitment JM, and item 6 from the social injustice JM. In addition, positive item-total correlations ranging from .06 to .12 (i.e., weak positive evidence) were observed for the following five items: items 2 and 4 from the marginalization JM and items 1, 3, and 5 from the social injustice JM. The remaining items had very low positive or negative item-total correlations. In addition, one item – item 5 from the marginalization JM – had a significant negative item-total correlation. This item was flagged for further consideration.

Overall, these results provide tentative positive evidence for a subset of items that correlate at least moderately with their respective JM item-total scores. At the same time, this evidence generally falls short of typical recommendations for scale development (e.g., item-total correlations exceeding .30). In order to investigate the alternative possibility that items correlated more strongly with a total score across all items (i.e., rather than within JMs), corrected item-total biserial correlations were also computed based on a total score for the CRT-W. As shown in the CRT-W Total column of Table 8 (far right), this did not result in significantly improved item-total correlations. In total, two items exhibited statistically significant and positive item-total correlations, compared to seven when the item-total correlations were examined within JMs. As with the previous analyses, the magnitude of item-total correlations was generally small, typically falling in the -.15 to .15 range.

Item intercorrelations. The relationships between items were examined for collinearity or other unexpected patterns. For this purpose, tetrachoric correlations were estimated. Tetrachoric correlations are a special case of polychoric correlations involving two artificially dichotomized variables. Following the rationale described previously with respect to the biserial correlation, tetrachoric correlations adjust for attenuation due to artificial dichotomy and distributions on the observed variables that depart from a 50-50 split (Cohen et al., 2003). As with the biserial correlations examined previously, tetrachoric correlations assume the underlying variables are continuous and normally distributed.

Tetrachoric item intercorrelations are shown in Table 9. In general, positive intercorrelations were expected, particularly within the item sets comprising each JM. This general pattern held but with intercorrelations generally low across items. Within JMs, average correlations were weak albeit positive in direction. The average intercorrelation was .05 for marginalization, .03 for revocable commitment, and .06 for social injustice. These averages were slightly higher than the average intercorrelation of .01 across all CRT-W items, again suggesting that items within JM correlate somewhat more highly than items across JMs. Nevertheless, item intercorrelations were generally low, with no item pair exceeding a correlation of +.40 and only 8% of item pairs exceeding a correlation of +.20.

Item-criterion correlations. The direction and magnitude of item-criterion correlations was considered next. Due to the strong skew observed for behavioral withdrawal criteria, item-criterion correlations were examined separately for *log* (x + 1) transformed and un-transformed criterion variables. Transformation had no substantive impact on the relationships observed. The average difference between item-criterion correlations with the transformed and un-transformed criteria was less than .001 across all item-criteria pairs. Thus, item-criterion analyses ultimately focused on the un-transformed behavioral withdrawal criteria.

Biserial item-criterion correlations are shown in Table 10. Each item was evaluated in terms of the direction and statistical significance of its biserial correlation with each behavioral withdrawal criterion. Positive relationships were generally anticipated, indicating that endorsement of the JM response option is associated with a higher frequency of withdrawal behaviors. A pattern of positive correlations across behavioral withdrawal criteria was taken as further evidence of an item's predictive validity.

In total, 11 items demonstrated significant positive relationships with at least one withdrawal criterion: items 1 and 8 from the marginalization of withdrawal JM, items 1, 2, 3, 4, 5, 6, and 9 from the revocable commitment JM, and items 4 and 6 from the social injustice JM. Among these items, seven had significant positive relationships with two or more behavioral withdrawal criteria: items 1, 2, 4, and 9 from the revocable commitment JM, and items 4 and 6 from the social injustice JM. Two items clearly stood out as having the most consistent patterns of positive relationships across criteria. Item 2 from the revocable commitment JM was significantly correlated with permanent withdrawal (r = .16, p < .05), lecture absenteeism (r = .12, p < .05), lab absenteeism (r = .15, p < .05), and lab lateness (r = .21, p < .05). Item 6 from the social injustice JM was significantly correlated with permanent withdrawal (r = .20, p < .01), frequency of lecture lateness (r = .16, p < .01), and lab lateness (r = .42, p < .01).

In addition to the 11 items noted above, two items exhibited a positive pattern of relationships across criteria, despite the fact that no single correlation reached statistical significance. Item 8 from the marginalization JM had trend-level (p < .10) positive correlations with permanent withdrawal, lecture absenteeism, and lecture lateness. Item 3 from the social injustice JM had trend-level positive correlations with lecture lateness, lab lateness, and early departures. The overall pattern of relationships observed for these items suggest that they may contribute positively to the predictive validity of JMs despite having relatively weak predictive relationships on their own. Adding these two items to the 11 items described previously, 13 items (or just over half) showed some positive evidence of predictive validity by correlating with one or more behavioral withdrawal criteria in the expected direction.

In contrast, six items correlated negatively (i.e., in the opposite direction than anticipated) with one or more behavioral withdrawal criteria, including four items from the marginalization JM (items 3, 4, 6, and 7), one item from the revocable commitment JM (item 8), and one item from the social injustice JM (item 1). The remaining items had generally weak and inconsistent correlations with behavioral withdrawal criteria. No items exhibited a mixed pattern in which correlations were positive and significant with some criteria and negative and significant with others.

Overall, the item-criterion correlational analyses point to a subset of items – 13 of 25 in total – that demonstrated initial positive evidence, as well as two more general conclusions. First, there was slight advantage observed for the number of items that exhibited positive evidence of predictive validity compared to the number of items that exhibited negative or equivocal evidence (13 versus 12 in total). Likewise, considering all of the item-criterion correlations reported in Table 10, significant positive correlations outnumbered significant negative correlations 21 to 9 – recall that positive correlations were hypothesized. The magnitude of biserial correlations were nevertheless small in the majority of cases, suggesting that predictive validity, at least at the item level, was not strong. Second, the number and proportion of items demonstrating initial positive evidence was clearly strongest for the revocable commitment JM and weakest for the marginalization JM. Although the hypothesis tests reported later in this manuscript provide evidence regarding the criterion-related validity of JMs, these analyses provide initial support for the tenability of the revocable commitment and social injustice JMs (and several current items therein). In contrast, item analyses call into question the tenability of the marginalization JM as currently conceived and operationalized.

Scale revision. Table 9 summarizes the reliability and validity evidence that was considered when deciding whether to retain or drop items from the final scales. Items were retained if they contributed positively to scale reliability and predicted students' behavioral withdrawal in the course; however, the latter consideration was weighted more heavily in all decisions about items, and a positive item-total correlation was not viewed as sufficient evidence in isolation to justify an item's retention. This is consistent with the primary objective of the conditional reasoning test, which is to predict the criterion of interest. In addition, problematic items flagged previously were evaluated in light of all available information.

In total, 12 of 25 initial CRT-W items were eliminated for subsequent analytic purposes. As already described, item 10 from the revocable commitment JM and item 2 from the social injustice JM were eliminated due to low variance (i.e., extreme p-values). The following items exhibited negative correlations with one or more behavioral withdrawal criteria (i.e., opposite of the direction anticipated) and were therefore eliminated: items 3, 4, 6, and 7 from the marginalization JM, item 8 from the revocable commitment JM, and item 1 from the social injustice JM. The following items were eliminated due to exhibiting non-significant relationships with all behavioral withdrawal criteria and failing to demonstrate evidence of positive trend-level relationships: items 2 and 9 from the marginalization JM, item 7 from the revocable commitment JM, and item

5 from the social injustice JM. Item 4 from the marginalization JM, previously flagged due to eliciting a disproportionate number of illogical distractor responses, was eliminated.

In contrast, 13 items were retained for subsequent analytic purposes. All 13 demonstrated positive evidence of criterion-related validity with one or more types of withdrawal behaviors. Among these, a total of six items exhibited positive item-total correlations with their corresponding JM item-total scores. Organized around the proposed JMs, the final sub-scales used in hypothesis testing included a 3-item measure of marginalization (items 1, 5, and 8), a 7-item measure of revocable commitment (items 1, 2, 3, 4, 5, 6, and 9), and a 3-item measure of social injustice (items 3, 4, and 6).

Revised scale descriptive statistics and reliability. Scale scores for marginalization of withdrawal, revocable commitment, and social injustice were computed by taking the mean of corresponding items. The descriptive statistics and internal consistency reliability analyses reported in this section are based on the final sample used for hypothesis testing.

Scores for marginalization of withdrawal ranged from 26 participants with a mean score of 0 (i.e., no JM responses endorsed across three items) to 10 participants with a mean score of 1 (i.e., all JM responses endorsed). The overall mean score across participants for marginalization of withdrawal was equal to 0.41 (*SD* = 0.24). Skewness was equal to .36 with a corresponding standard error of .17 (standardized skewness = 2.16), indicating a moderate positive skewness. The histogram for marginalization of withdrawal is shown in Figure 9. Coefficient alpha for the 3-item scale was .07, indicating poor internal consistency reliability. This is somewhat reduced from the coefficient alpha of .14 observed for the initial 9-item scale.

Scores for revocable commitment ranged from 10 participants with a mean score of 0 (i.e., no JM responses endorsed across the 7 items) to 1 participant with a mean score of 0.86. The overall mean across participants was equal to 0.33 (*SD* = 0.18), indicating that, on average, participants endorsed just over two JM responses. Skewness was equal to 0.33 with a corresponding standard error of .17, indicating a moderate positive skew (standardized skewness = 1.96). The histogram for revocable commitment is shown in Figure 10. Coefficient alpha for the 7-item scale was .16, indicating poor internal consistency reliability. This is somewhat improved from the coefficient alpha of .07 observed for the initial 10-item scale.

Scores for social injustice ranged from 134 participants with a mean score of 0 (i.e., no JM responses endorsed across the 3 items) to one participant with a mean score of 1. The overall mean across participants was equal to 0.15 (*SD* = 0.21), indicating that, on average, participants endorsed less than one JM response. Skewness was equal to 1.21 with a corresponding standard error of .17, indicating a strong positive skew (standardized skewness = 7.22). However, the decision was made not to transform scores on social injustice since (a) the withdrawal criteria being predicted were similarly positively skewed such that JM's correlation with criterion variables may not be attenuated as a result, and (b) the strength of the positive skewness observed for both the JM and withdrawal criterion variables precluded the effectiveness of transformations to reduce or eliminate skewness. The histogram for social injustice is shown in Figure 11. Coefficient alpha for the 3-item scale was .09, indicating poor internal consistency reliability. This is somewhat reduced from the coefficient alpha of .17 observed for the initial 6-item scale.

A total score was also derived across all retained CRT-W items following scale refinement. Scores ranged from six participants with a mean score of 0.08 (i.e., just under one JM response across the 13 items) to one participant with a mean score of 0.75. The overall mean across participants was equal to 0.31 (*SD* = 0.12), indicating that, on average, participants endorsed just fewer than four JM responses across 13 items. Skewness was equal to 0.56 with a corresponding standard error of .17, indicating a moderate-to-strong positive skewness (standardized skewness = 3.38). For the reasons described previously with respect to the social injustice JM, the decision was made not use a transformation in subsequent analyses involving total scores on the CRT-W. The histogram for total scores on the CRT-W is shown in Figure 12. Coefficient alpha for the 13-item scale was .05, indicating poor internal consistency reliability. The estimated coefficient alpha was negative for the initial 25-item scale, a violation of the assumptions of reliability analysis due to negative average covariance among the full set of items.

The scale analyses described above lead to a few important insights and qualifications for subsequent hypothesis testing. The most apparent is the low internal consistency reliability observed based on JM sub-scales, as well as for an overall scale composed of all 13 retained items. This suggests that the items comprising the JMs and overall scale have substantial heterogeneity. From a pragmatic perspective, this indicates that the ability of JMs to correlate with external criterion variables will be greatly attenuated. From a construct validity standpoint, the low internal consistency reliability calls into question the tenability of the JMs as an organizing framework for the tendencies assessed within each item. However, at the same time, it should be noted that the consistently low inter-item correlations suggests that alternative item groupings likely would not lead to adequate levels of internal consistency reliability. Furthermore, the present study does not provide an adequate sample size for conducting a full-scale exploratory factor analysis (James & LeBreton, 2011).

In light of these observations, it is important to recognize with appropriate caution that strong evidence of construct validity (including factor analytic support for these particular JMs) awaits future investigations. There is nevertheless adequate rationale to proceed with testing the study's hypotheses and gathering further evidence of the JM's predictive relationships with behavioral withdrawal criteria. First, it is important to recognize that heterogeneous scales can have substantial utility as valid predictors of work-relevant behaviors. Biodata instruments provide one example of this approach to scale development and validation, where the primary focus is on predicting the external criterion rather than developing an internally reliable instrument. Although both objectives are ultimate goals within a line of research, the present study was focused primarily on testing predictive validity as an initial step toward an instrument that is both practically useful as a predictor of withdrawal behaviors and possesses adequate evidence of construct validity. Toward this purpose, the hypothesis testing is meaningful for helping to prioritize future work around the items and JM concepts with the greatest predictive potential.

Hypothesis Testing

The results of hypothesis testing are presented in three sections. The first section describes tests for hypotheses linking scores on the CRT-W to behavioral withdrawal criteria (*Hypotheses 1 – 10*). The second section describes tests for hypotheses linking personality traits and scores on the biodata variables to behavioral withdrawal criteria (*Hypotheses 11 – 22*). The third section describes tests for the hypothesized incremental validity of the CRT-W beyond personality and biodata measures (*Hypothesis 23*). Prior to reporting the results of hypothesis testing, preliminary analyses were undertaken to examine the role of grades in withdrawal behaviors and consider the appropriateness of including grades as a covariate in subsequent analyses.

Role of Student Grades in Withdrawal Behaviors

An indication of student grades was derived as the cumulative points earned by a student in the course. Cumulative point totals were computed three times per week across all weeks of the semester with the exception of Thanksgiving and final exam week. The three days included Tuesdays, Thursdays, and Fridays. On each of these days, all students had an opportunity to accumulate new points based on in-class opportunities (e.g., exams, quizzes, and participation points during Tuesday and Thursday lecture meetings) and labs and dated assignments occurring throughout the week. The latter point opportunities were added to cumulative point totals on Fridays to account for differences in timing across lab sections. In total, these days encompassed 38 different opportunities throughout the semester when an updated cumulative point total could be computed for all students in the course.

As a starting point for examining the role of grades in withdrawal behaviors, students' end-of-semester cumulative point total was correlated with each criterion variable. Results indicated that end-of-semester grades were strongly correlated with permanent withdrawal (r = -.82), moderate-to-strongly correlated with absenteeism behaviors (r = -.61 for lecture absenteeism; r = -.52 for lab absenteeism), and weak-to-moderately correlated with lateness behaviors (r = -.29 for lecture lateness; r = -.16 for lab lateness). These results suggest that course grades are an important factor in students' withdrawal behaviors, particularly their likelihood to permanently withdraw and their levels of absenteeism throughout the semester.

At the same time, it is important to point out that these correlations capture not only the prospective effect of grades on withdrawal behaviors but also the reverse effect of withdrawal behaviors on grades. Although this is plausible for all of the withdrawal behaviors examined, reverse effects are most directly apparent as a result of absenteeism (e.g., absence precludes a student from getting in-class points on the day of the absence episode) and permanent withdrawal (e.g., permanently withdrawing precludes the student from getting any subsequent points in the course). Therefore, to establish the appropriateness of controlling for student grades when examining hypothesized predictive effects, some indication should be provided that grades indeed account for prospective variance in withdrawal behaviors.

In follow-up analyses, cumulative points earned were correlated with subsequent withdrawal behaviors at several points in time excluding from consideration all prior episodes of withdrawal. For permanent withdrawal, this approach was repeated several times just prior to each student's date of withdrawal from the course. Correlations ranged from $r = -.48$ to $r = -.65$, with a median correlation of $r = -.60$. For absenteeism and lateness behaviors, the semester was divided into halves and cumulative points through the first half of the course were correlated with absenteeism and lateness behaviors in the second half. Correlations revealed moderate-to-strong prospective effects of cumulative points earned in the first half of the semester on absenteeism behaviors in the second half of the semester ($r = -.49$, $p < .01$, for lecture absenteeism; $r = -.39$, $p < .01$, for lab absenteeism). Results were mixed for correlations with lateness behaviors, with the effect holding for lab lateness ($r = -.15$, $p < .05$) but not for lecture lateness ($r = -.03$, $p = .32$). In follow-up to these analyses, withdrawal behaviors through the first half of the course were correlated with points earned in the second half. Correlations revealed moderate-to-strong effects of first-half absenteeism on second-half points ($r = -.53$, $p < .01$, for lecture absenteeism; $r = -.42$, $p < .01$, for lab absenteeism) and mixed results for first-half lateness correlated with second-half points ($r = -.25$, $p < .01$, for lecture absenteeism; $r = .00$, $p = .99$, for lab absenteeism).

These results suggest in a crude way that students' grades might be recursively related to withdrawal behaviors. Based on the pattern of correlations, poor performance in the course would be expected to increase subsequent withdrawal behaviors, which then leads to further decline in performance, and so on. This also suggests that, for analytic purposes, use of an end-of-semester point total as a covariate is likely to overcorrect for the *prospective* effects of grades on withdrawal behaviors (and perhaps more than slightly with respect to lecture lateness). There is nevertheless an important prospective effect that ought to be investigated when modeling the hypothesized effects of the conditional reasoning test. In other words, performance in the course appears to be a proximal factor in students' decision to continue or discontinue their participation and attendance in the course, and to a lesser degree, students' decisions related to lateness behaviors in lab and lecture. In the analyses that follow, students' end-of-semester point totals are examined as a covariate at two stages: (a) first, when examining the predictive validity of the CRT-W (Section I), and (b) second, when examining the incremental effects of the CRT-W through regression models that incorporate the explicit personality traits and biodata measures (Section III).

Section I: Hypotheses Linking CRT-W to Behavioral Withdrawal

Marginalization of withdrawal JM. Hypotheses involving absenteeism and lateness behaviors were evaluated using bivariate correlations. As described previously, behavioral withdrawal criteria and scores on the conditional reasoning test were not normally distributed. However, in all cases the direction and shape of skewness was generally similar – positive, ranging from moderate to strong in magnitude. As a result, the degree of attenuation and potential bias in parameters due to non-normality may be somewhat ameliorated. Beyond univariate normality, four additional assumptions of the general linear model were examined with respect to bivariate relationships, specifically absence of multivariate outliers, multivariate normality, linearity, and homogeneity of variance.

As a first step, bivariate scatterplots were examined for all pairwise combinations of marginalization with absenteeism and lateness variables. The panels in Figure 13 show bivariate scatterplots, including the fitted linear trend line as a visual point of reference. In general, visual inspection of the scatterplots did not signal the presence of extreme non-linearity or heteroscedasticity. Moderate heteroscedasticity was noted in the lab absenteeism scatterplot, such that there was greater variance through the moderate range of scores on the marginalization JM. Next, scatterplots of residuals were visually examined, with standardized residuals shown on the y-axis and standardized predicted values shown on the x-axis (see Figure 14). These scatterplots revealed substantial deviations from normality, evidenced by a higher proportion of residuals falling above the 0-residual reference line.

The residual plots also revealed a small number of cases with extreme residuals (e.g., standardized residuals greater than 6.0), signaling their disproportionate effect on fitted bivariate relationships. These included one case with a studentized deleted residual of 6.08 (marginalization-*lab absenteeism*), one case with a studentized deleted residual of 6.50 (marginalization-*lecture lateness*), and two cases with studentized deleted residuals of 6.23 and 5.38 (marginalization-*lab lateness*). In all cases, these individuals had extremely high levels of withdrawal behaviors that were considerably under-predicted by the estimated linear relationships. This same pattern held across all bivariate relationships involving lab absenteeism, lecture lateness, and lab lateness. Therefore, these four cases were eliminated from all subsequent correlational analyses on a pairwise basis (i.e., only from affected bivariate relationships). In contrast, no multivariate outliers were identified with respect to analyses involving lecture absenteeism and early departures.

While eliminating the four cases described above (on a pairwise basis) provides a more accurate estimate of the linear relationship between marginalization and withdrawal behaviors, the non-normality that was observed was not significantly reduced, suggesting that parametric estimates and accompanying statistical significance tests need to be interpreted with caution. For this reason, a sensitivity analysis approach was adopted.[3] Parametric and non-parametric correlations were estimated, as shown in Table 12. The discussion in text relies on Pearson's correlations unless a difference was observed in the statistical inferences produced by the parametric and non-parametric tests.

Hypothesis 1 proposed a positive relationship between the marginalization of withdrawal JM and frequency of lateness behaviors (i.e., more frequent lateness). Correlations with both lecture lateness (r = .01, p = .43, N = 212) and lab lateness (r = .00, p = .48, N = 132) failed to provide support for this hypothesis. Hypothesis 2 proposed a positive relationship between the marginalization of withdrawal JM and frequency of absenteeism behaviors (i.e., more frequent absence). Correlations with both lecture absenteeism (r = -.03, p = .32, N = 213) and lab absenteeism (r = .02, p = .38, N = 212) failed to provide support for this hypothesis.

Follow-up analyses examined whether controlling for course grades would impact the relationship between marginalization of withdrawal and lateness and absenteeism behaviors, for example, allowing for a stronger positive relationship to emerge. However, results did not change the substantive interpretation of these relationships offered above. Furthermore, scores on the marginalization JM were not significantly correlated with students' end-of-semester cumulative point totals (r = -.01, p = .87, N = 213).

Survival analysis. Hypothesis 3 proposed that participants with a greater tendency for marginalizing withdrawal exhibit a higher likelihood of permanent withdrawal. Traditional analytic approaches to studying turnover behavior have involved logistic regression with a dichotomous turnover variable. Several authors have advanced arguments against the use of logistic regression in the context of studying turnover, in favor of longitudinal analytic approaches, such as survival analysis (Morita, Lee, & Mowday, 1989; Singer & Willett, 1991; Somers, 1996; Somers & Birnbaum, 1999). Their arguments for the use of survival analysis over traditional static analyses can be summarized as follows: (a) dichotomizing turnover behavior is conceptually misleading and suggests that, "the duration of retention is not seen as important in understanding employee withdrawal" (p. 318; Somers, 1996), (b) survival analysis allows time to be included as an explicit variable in models of employee turnover, (c) the censoring feature of survival analysis is capable of dealing with missing data, which is common in longitudinal data sets, and (d) traditional approaches to identifying employees who turnover early (e.g., in the first 3 months) versus late (e.g., after 5years) and developing predictors separately for these groups severely decreases the detailed nature of information that can be yielded and reduces statistical power (Morita et al., 1989). In short, these authors have suggested that traditional static analyses are inappropriate and statistically disadvantaged for studying turnover and other longitudinal phenomena and that survival analysis is an appropriate alternative (see Singer & Willett, 1991). In line with these authors' recommendations, a survival analysis framework was adopted in this study to test all hypotheses involving permanent withdrawal from the course.

Within the survival analysis framework, Cox hazards regression provides a flexible statistical technique for modeling the predictive effect of variables on the timing and occurrence of the permanent withdrawal event. The procedure estimates the effect of time invariant and/or time-varying predictors on *hazard* or risk of experiencing the event, expressed as the log cumulative hazard function. The log cumulative hazard function expresses the conditional probability of the event, in this case permanent withdrawal, given survival up to each assessed point in time. Regression coefficients and corresponding confidence intervals provide a test of each predictor's effect on the log cumulative hazard function, with the degree of vertical shift of the hazard function relative to a theoretical baseline function in which all predictor values equal 0, reflecting the magnitude of the variable's impact (Singer & Willet, 2003).

There are three main data requirements or assumptions of Cox hazards regression. First, the time to event variable must represent a continuous measure of time. Although this is not strictly met in any instance – that is, because all measurement is discontinuous at some level – practically speaking, days can be considered a continuous measurement of time in the present study, when it is not feasible and unnecessary to capture permanent withdrawal events in terms of hours, minutes, or a smaller unit of measurement (Singer & Willet, 2003). In contrast, a discrete measurement of time in the present context could involve chunked units, such as weeks or months. Nevertheless, days provide a reasonably continuous assessment of time and allows for the use of Cox hazards regression in the present study. The timing variable was set as the number of days to the permanent withdrawal event. For participants who eventually withdrew from the course, this equaled the number of days from the first day of class to the date of withdrawal. For censored participants, or those who did not experience the event within the study window, this equaled the total number of days in the semester, 106.

Second, the number of *ties* or individuals who experience the event at the same time needs to be examined. A high proportion of ties – for example, more tied than unique event times – creates computational difficulties for Cox hazards regression (Singer & Willet, 2003). In the present data, there were a moderate number of ties. Withdrawal events occurred on a total of 16 days, 10 of which were unique to a particular individual. The other six days involved a tie, with more than one individual withdrawing on the same day. Although this is not ideal for computational purposes, the approximation techniques implemented by statistical software provide a reasonable solution to the problem. SPSS implements Breslow's (1974) approximation method, which performs comparably to other approximation methods such as Efron's (1977) method (Pelz & Klein, n.d.), when the number of ties is reasonably low. Breslow's method assumes a sequential ordering of the tied cases.

And third, the use of time-invariant predictors in Cox hazards regression assumes that the predictor's effect on the log hazard function is constant over time. The shape of the theoretical baseline hazard function – that is when all predictor values are equal to 0 – is irrelevant. Rather, the assumption applies to the proportionality of hazard functions based on different predictor values. This implies that, if plotted separately, the shape of the hazard functions would be identical for subgroups based on all levels of the predictor – that is, separated by the same vertical distance across all time points.

Although there is a log cumulative hazard function for all possible values of the predictor variables (and all combinations of predictor values when there are multiple predictors), a check of the assumption is made by visually examining the proportionality of separate cumulative log hazard functions based on a median split on the predictor. Kaplan-Meier estimation is the preferred method for plotting subgroup cumulative survival and hazard functions for continuous-time models (Singer & Willet, 2003).

Prior to testing Hypothesis 3, the baseline cumulative survival and hazard plots were examined for descriptive purposes in order to better understand the timing and distribution of permanent withdrawal events throughout the semester. Figure 15 shows the cumulative hazard function (top panel) and the cumulative survival function (bottom panel). The first estimate of cumulative hazard is based on the first observation of a permanent withdrawal event at day 26. Hazard appears to increase somewhat linearly between days 30 to 70, followed by a somewhat stronger increase in hazard between days 70 to 95, and finally a rapid spike between days 95 to 100, the days just preceding the final exam in the course. In other words, the risk of permanent withdrawal for those remaining in the course built at a somewhat steady rate up to day 70 and then increased more rapidly and finally spiking just prior to the final exam. It is also important to recognize, however, that the overall hazard is relatively low over time never exceeding a cumulative hazard of .20. The cumulative survival function shows this point more clearly, indicating that just over 85% of the original sample remained in the course to its conclusion.

Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split of the marginalization predictor. These plots are shown in the top two panels of Figure 16. Based on the median split and resulting separate lines that are plotted, it is possible to gauge both the potential effect of the marginalization predictor and the tenability of the assumed proportionality. The plots suggest that hazard is slightly higher across time for participants with above-median scores on the marginalization JM – opposite of the anticipated effect. The overall shape of the hazard functions is similar over time, and a plot of the partial residuals of the marginalization variable over time confirms that they are indeed uncorrelated. Taken together, this suggests that the assumption of proportional hazards is not violated.

Cox hazards regression models were estimated in two steps. The first model estimated the effect of the marginalization JM as a sole predictor of the log cumulative hazard function. This model was not significant, χ^2 (1, N = 213) = .46, p = .46, nor was the predictive effect of marginalization (b = .52, p = .50).[4] The second model estimated the effect of marginalization controlling for students' cumulative points, which was entered as a time-varying predictor based on 38 successive observations of this variable over time. This model was significant, χ2 (3, N = 213) = 8.43, p < .05, as was the predictive effect of time-varying cumulative points (b = -.01, p < .05); however, the effect of marginalization was unchanged from the prior model (b = .51, p = .50). Therefore, Hypothesis 3 was not supported.

The negative sign of the regression coefficient for the cumulative points variable confirms that an increase in cumulative points was associated with a decrease in hazard. The exponentiated regression coefficient (or hazard ratio) of .99 suggests that an increase of 1-point in the course was associated with a 1% decrease in risk of permanent withdrawal. This is indicative of a large effect given that the overall mean for cumulative points across all time observations was equal to 352.

Revocable Commitment JM. Hypothesis 4 proposed a positive relationship between revocable commitment and likelihood of permanent withdrawal. Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split of the revocable commitment predictor. These plots are shown in the top two panels of Figure 17. The plots indicate that hazard was similar over time for participants with above- and below-median scores on the revocable commitment JM up to 95 days, at which point the hazard is markedly higher for the above-median sample. The overall shape of the hazard functions appear to differ over time. A bivariate scatterplot of the partial residuals of the revocable commitment variable with days indicates that that there is a positive trend over time. Taken together, this suggests that the assumption of proportional hazards is violated and that results may be biased as a result.

With the assumption of proportional hazards violated, an alternative model was specified entering both the "main effect" revocable commitment variable and a linear interaction term with time in days. This model was significant overall, $\chi 2$ (2, N = 213) = 6.38, p < .05, with a significant interaction observed between revocable commitment and time (b = .11, p < .05). This indicates that the effect of revocable commitment on permanent withdrawal depends on time. The positive regression coefficient observed indicates that the hazard associated with higher scores on revocable commitment is observed to increase over time. The Kaplan-Meier plots shown in Figure 17 support this interpretation, demonstrating that the increased risk associated with above-median revocable commitment scores is not observed until after 80 days in the course. This provides partial support for Hypothesis 4, demonstrating a time-contingent effect of revocable commitment.

No hypotheses were offered for revocable commitment's relationship with temporary forms of withdrawal (i.e., lateness and absenteeism). Nevertheless, these relationships were examined on an exploratory basis. Bivariate scatterplots of scores on the revocable commitment JM with absenteeism and lateness variables are shown in Figure 18. The scatterplot for lecture lateness suggested a potential curvilinear relationship; however, exploratory iterative curve fitting procedures did not support the existence of a quadratic or cubic relationship. This same plot revealed a moderate degree of heteroscedasticity, with greater variance observed through the moderate range of scores on the revocable commitment JM. The remaining scatterplots did not point to obvious patterns of non-linearity or heteroscedasticity. Next, scatterplots of residuals were visually examined, with standardized residuals shown on the y-axis and standardized predicted values shown on the x-axis (see Figure 19). As with the residual plots for marginalization, these scatterplots revealed substantial deviations from normality, with a higher proportion of residuals falling above the 0-residual reference line.

Correlations with both lecture lateness (r = .06, p = .19, N = 212) and lab lateness (r = .12, p = .09, N = 132) failed to provide support for a positive relationship between the revocable commitment JM and frequency of lateness behaviors (i.e., more frequent lateness). On the other hand, results provided partial support for a positive relationship between the revocable commitment JM and frequency of absenteeism behaviors (i.e., more frequent absence). Revocable commitment had a small but statistically significant positive correlation with lab absenteeism (r = .15, p < .05, N = 212) but was uncorrelated with lecture absenteeism (r = .06, p = .19, N = 213). This evidence should be qualified as a relatively weak form of support given that revocable commitment accounted for less than 3% of the variance in lab absenteeism.

Follow-up analyses examined whether controlling for course grades would impact the relationship between revocable commitment and lateness and absenteeism behaviors, for example, allowing for a stronger positive relationship to emerge. However, partial correlations were markedly similar in direction and magnitude to the correlational results reported above. Furthermore, scores on the revocable commitment JM were not correlated with end-of-semester cumulative points (r = -.09, p = .21, N = 213).

Social Injustice JM. Bivariate scatterplots of scores on the social injustice JM with absenteeism and lateness variables are shown in Figure 20. The scatterplots for lecture absenteeism and lecture lateness suggested a potential curvilinear relationship; however, exploratory iterative curve fitting procedures did not support the existence of a quadratic or cubic relationship. Moderate heteroscedasticity was observed for the lab absenteeism plot, with less variance observed toward the high range of scores on the social injustice JM. The remaining scatterplots did not point to obvious patterns of non-linearity or heteroscedasticity. Next, scatterplots of residuals were visually examined, with standardized residuals shown on the y-axis and standardized predicted values shown on the x-axis (see Figure 21). As with the previous residuals plots, residuals deviated from the multivariate normal pattern, with a higher proportion of residuals falling above the 0-residual reference line.

Hypothesis 5 proposed a positive relationship between the social injustice JM and frequency of lateness behaviors (i.e., more frequent lateness). Results provided partial support for this hypothesis, with the results of parametric and non-parametric tests differing slightly (Spearman's rho = .15, $p < .05$, $N = 132$ versus Pearson's $r = .11$, $p = .09$, $N = 132$). Taken together with the small effect sizes for these relationships, this can be considered a weak form of positive evidence. The correlation with lecture lateness, on the other hand, was small and non-significant in both cases ($r = .10$, $p = .09$, $N = 212$). Hypothesis 6 proposed a positive relationship between the social injustice JM and frequency of absenteeism behaviors (i.e., more frequent absence). Correlations with both lecture absenteeism ($r = .01$, $p = .44$, $N = 213$) and lab absenteeism ($r = .00$, $p = .50$, $N = 212$) failed to provide support for this hypothesis.

Follow-up analyses examined whether controlling for course grades would impact the relationship between social injustice and lateness and absenteeism behaviors, for example, allowing for a stronger positive relationship to emerge. However, partial correlations were markedly similar in direction and magnitude to the correlational results reported above. Interestingly, the correlation between the social injustice JM and end-of-semester grades was significant and negative ($r = -.15$, $p < .05$, $N = 213$), indicating that a stronger social injustice bias was associated with lower performance in the course.

Hypothesis 7 proposed a positive relationship between social injustice and likelihood of permanent withdrawal. Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split of the social injustice predictor.

These plots are shown in the top two panels of Figure 22. As described previously with respect to the revocable commitment JM, these plots indicate that hazard was similar over time for participants with above- and below-median scores on the social injustice JM up to 95 days, at which point the hazard function diverges and is markedly higher for the above-median sample. The overall shape of the hazard functions appear to differ over time and cross around 80 days. A bivariate scatterplot of the partial residuals of the social injustice variable with days indicates that that there is a positive trend over time. Taken together, this suggests that the assumption of proportional hazards is violated; the social injustice variable's effect on hazard is not constant over time.

With the assumption of proportional hazards violated, an alternative model was specified entering both the "main effect" social injustice variable and a linear interaction term with time in days. The overall model was not significant, $\chi 2$ (2, *N* = 213) = 5.14, $p = .08$, as was the interaction term ($b = .08$, $p = .08$) and main effect of social injustice ($b = -5.45$, $p = .16$). Although the trend-level relationship with the interaction of time and scores on the social injustice JM are consistent with the general pattern of increasing hazard over time associated with above-median scores, the magnitude of this effect is small and not statistically significant. Therefore, hypothesis 7 was not supported.

Overall CRT-W Scores. Analyses in this section examined the relationships between total scores on the CRT-W, derived as a mean across all retained items from the three JMs, and behavioral withdrawal criteria. Bivariate scatterplots of total scores with each absenteeism and lateness variable are shown in the panels of Figure 23. Potential non-linearity was examined for the relationship between overall CRT-W scores and lecture lateness. Exploratory curve fitting did not support the incremental gain of a quadratic relationship (e.g., inverse u-shape) beyond a simple linear relationship. This plot nonetheless suggests moderate heteroscedasticity, with higher variance observed for moderate ranges of overall CRT-W scores. Next, scatterplots of residuals were visually examined, with standardized residuals shown on the y-axis and standardized predicted values shown on the x-axis (see Figure 24). As with the previous residuals plots, residuals deviated from the multivariate normal pattern, with a higher proportion of residuals falling above the 0-residual reference line. The non-normal pattern was most distinct for lab absenteeism, lecture lateness, and lab lateness.

Hypothesis 8 proposed a positive relationship between overall scores on the CRT-W and frequency of lateness behaviors (i.e., more frequent lateness). Results provided partial support for this hypothesis. Parametric and non-parametric results differed slightly for relationships with lab lateness, which reached statistical significance based on Spearman's rho (rho = .15, $p < .05$, $N = 132$) but not Pearson's correlation ($r = .12$, $p = .07$, $N = 132$). Taken together with the small effect sizes for these relationships, this can be considered a weak form of positive evidence. The correlation with lecture lateness, on the other hand, was small and non-significant in both cases ($r = .09$, $p = .09$, $N = 212$).

Hypothesis 9 proposed a positive relationship between overall scores on the CRT-W and frequency of absenteeism behaviors (i.e., more frequent absenteeism). Results provided partial support for this hypothesis. Parametric and non-parametric results differed slightly for relationships with lab absenteeism, which reached statistical significance based on Pearson's correlation ($r = .12$, $p < .05$, $N = 212$), but not Spearman's rho (rho = .10, $p = .07$, $N = 212$). Taken together with the small effect sizes for these relationships, this can be considered a weak form of positive evidence. The correlation with lecture absenteeism, on the other hand, was small and non-significant in both cases (r = .10, p < .09, N = 213).

Follow-up analyses examined whether controlling for course grades would impact the relationship between overall scores on the CRT-W and lateness and absenteeism behaviors, for example, allowing for a stronger positive relationship to emerge. However, partial correlations were markedly similar in direction and magnitude to the correlational results reported above. A trend-level (p < .10) relationship was noted for overall CRT-W scores and end-of-semester cumulative grades (r = -.12, N = 213).

Hypothesis 10 proposed a positive relationship between overall scores on the CRT-W and likelihood of permanent withdrawal. Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split. These plots are shown in the top two panels of Figure 25. The plots indicate that hazard was similar over time for participants with above- and below-median overall CRT-W scores up to 80 days, at which point the hazard functions diverge and become higher for the above-median sample. The overall shape of the hazard functions appear to differ over time. A bivariate scatterplot of the partial residuals indicates that that there is a positive trend over time. Taken together, this suggests that the assumption of proportional hazards is violated.

With the assumption of proportional hazards violated, an alternative model was specified entering both the “main effect” variable and a linear interaction term with time in days. This model was significant overall, $\chi 2$ (2, N = 213) = 9.58, $p < .01$, with a significant interaction observed between overall CRT-W scores and time (b = .19, $p < .05$). This indicates that the effect of overall implicit dispositional tendency to withdraw on permanent withdrawal depends on time. The positive regression coefficient observed indicates that the hazard associated with higher overall CRT-W scores is observed to increase over time. The Kaplan-Meier plots shown in Figure 25 support this interpretation, demonstrating that the increased risk associated with above-median CRT-W scores is not observed until after 80 days in the course. This provides partial support for Hypothesis 10, demonstrating a time-contingent effect of overall implicit tendencies to withdraw.

CRT-W and early departures. In addition to the hypothesized relationships with permanent withdrawal, lateness, and absenteeism, the relationships between JMs and early departures were examined on an exploratory basis. The panels in Figure 26 show bivariate scatterplots with early departure. No obvious departures from linearity were observed. However, scatterplots with social injustice and overall scores on the CRT-W revealed moderate heteroscedasticity. Specifically, there appears to be decreasing variance in early departures as scores on social injustice increase and increasing variance in early departures as overall scores on the CRT-W increase. Residuals plots for these relationships were examined next and are shown in the panels of Figure 27. Residual plots revealed substantial non-normality, as evidenced by a higher proportion of residuals falling above the 0-residual reference line. Consequently, both parametric and non-parametric correlations were examined to gauge the extent to which parametric statistics may be biased due to non-normality. Results were highly consistent across both correlational analyses.

Correlational analyses revealed that all predictive relationships with early departures were small and non-significant, including marginalization of withdrawal (r = .09, p = .09, N = 211), revocable commitment (r = .05, p = .23, N = 211), social injustice (r = .06, p = .18, N = 211), and overall scores on the CRT-W (r = .11, p = .06, N = 211).

Regression analyses with JMs predicting behavioral withdrawal. A series of exploratory analyses were undertaken to examine the predictive effects of JMs in combination within a multiple regression framework. Models were developed and tested separately for each absenteeism and lateness variable, in addition to early departures. All three JMs were entered into the models as predictors simultaneously within a single step. Residual plots (not shown) revealed moderate heterogeneity of variance and larger deviations from multivariate normality, particularly for lab absenteeism, lecture lateness, and lab lateness variables. In contrast, deviation from multivariate normality was less pronounced for lecture absenteeism and early departures. Therefore, results should be interpreted cautiously, as it is possible that parameter estimates are biased due to non-normality. Outlier screening led to the identification of no multivariate outliers. However, the four extreme cases previously identified at the bivariate level were removed using pairwise deletion.

For lecture absenteeism, the overall model was not significant, F (3,209) = .319, p = .81; nor were the JMs as individual predictors within the model (marginalization, b = -.02, p = .69; revocable commitment, b = .05, p = .40; social injustice, b = .01, p = .86). For lab absenteeism, the overall model was not significant, F (3, 208) = 1.64, p = .18;

however, the revocable commitment JM did have a significant positive relationship with the criterion (revocable commitment, $b = .11$, $p < .05$), indicating that higher scores on this JM was associated with higher levels of lab absenteeism while controlling for scores on marginalization ($b = .02$, $p = .61$) and social injustice ($b = .00$, p = .95).

For lecture lateness, the overall model was not significant, F (3, 207) = .93, p = .43; nor were the JMs as individual predictors within the model (marginalization, $b = .00$, $p = .88$; revocable commitment, $b = .03$, $p = .39$; social injustice, $b = .04$, $p = .16$). For lab lateness, the overall model was not significant, F (3, 128) = 1.20, $p = .31$; nor were the JMs as individual predictors within the model (marginalization, $b = -.01$, $p = .86$; revocable commitment, $b = .05$, $p = .19$; social injustice, $b = .04$, $p = .19$). Likewise, for early departures, the overall model was not significant, F (3, 207) = 1.08, $p = .36$, N = 213; nor were the JMs as individual predictors within the model (marginalization, b = .06, $p = .18$; revocable commitment, $b = .05$, $p = .38$; social injustice, $b = .03$, $p = .44$).

Section II: Hypotheses Linking Personality Traits and Biodata Measures to Behavioral Withdrawal

This section examines relationships between previously validated predictors of withdrawal and the behavioral withdrawal variables measured in the present investigation. Conscientiousness and emotional stability are the focus of analysis when examining the effects of personality on withdrawal, though exploratory analyses were undertaken to investigate the predictive effects of the remaining three traits within the Five Factor Model (i.e., agreeableness, extraversion, and openness to experience). Likewise, previously validated themes that were adapted for the present study's context were the main focus of analysis when examining the predictive effects of biodata.

Several additional themes were investigated on an exploratory basis. Because the personality and biodata variables represented a secondary focus of the present study and also because these variables have been studied extensively in prior research (e.g., Barrick & Zimmerman, 2005; 2009), discussions of data screening and preliminary descriptive analyses are streamlined.

Conscientiousness. Figures 28 and 29 show bivariate scatterplots and residuals plots for conscientiousness with the absenteeism and lateness variables. The plots reveal moderate heteroscedasticity, with greater variance observed in lecture absenteeism, lab absenteeism, and lecture lateness as scores on conscientiousness increase. Residuals were non-normally distributed about the 0-residual reference line, with a higher proportion of residuals having a positive value. This indicates conditions of multivariate normality were not met. There was no indication of non-linearity in the plots, and exploratory curve fitting procedures indicated that the simple linear model provided the best fit to the data.

Hypothesis 11 proposed a negative relationship between conscientiousness and frequency of lateness behaviors (i.e., fewer lateness episodes associated with higher levels of conscientiousness). This hypothesis was not supported. Correlations were in the expected direction but were small in magnitude and failed to reach statistical significance for both lecture lateness (r = -.09, p = .11, N = 211) and lab lateness (r = -.02, p = .41, N = 132). Hypothesis 12 proposed a negative relationship between conscientiousness and frequency of absenteeism behaviors. This hypothesis was supported. Conscientiousness was negatively correlated with both lecture absenteeism (r = -.22, p < .01, N = 213) and lab absenteeism (r = -.16, p < .05, N = 212). The magnitude of the correlations observed suggests that conscientiousness accounted for between 3 and 5% of the variance in absenteeism behaviors.

Hypothesis 13 proposed a negative relationship between conscientiousness and permanent withdrawal. Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split. These plots are shown in the top two panels of Figure 30. The plots suggest that hazard is slightly higher across time for participants with below-median scores on conscientiousness. The overall shape of the hazard functions is roughly similar over time, and a plot of the partial residuals of the marginalization variable over time suggests that there is a slight positive trend. Taken together, this suggests that the assumption of proportional hazards is not clearly violated and analyses proceeded based on the time invariant predictor.

Cox hazards regression models were estimated in two steps. The first model estimated the effect of conscientiousness as a sole predictor of the log cumulative hazard function. This model was not significant, $\chi 2$ (1, N = 213) = .19, p = .66, nor was the predictive effect of conscientiousness (b = .09, p = .66). The second model estimated the effect of conscientiousness controlling for students' cumulative points, which was entered as a time-varying predictor. This model was significant, $\chi 2$ (3, N = 213) = 8.20, $p < .05$, as was the predictive effect of time-varying cumulative points (b = -.01, $p < .05$); however, the effect of conscientiousness was unchanged from the prior model (b = .51, p = .50). Therefore, hypothesis 13 was not supported.

Emotional Stability. Figures 31 and 32 show bivariate scatterplots and residuals plots for emotional stability with the absenteeism and lateness variables. The plots reveal moderate heteroscedasticity, with lower variance observed in lecture absenteeism and lab absenteeism as scores on conscientiousness increase. Residuals were non-normally distributed about the 0-residual reference line, with a higher proportion of residuals having a positive value. This indicates conditions of multivariate normality were not met. There was no indication of non-linearity in the plots, and exploratory curve fitting procedures indicated that a simple linear model provided the best fit to the data.

Hypothesis 14 proposed a negative relationship between emotional stability and frequency of lateness behaviors (i.e., fewer lateness episodes associated with higher levels of conscientiousness). This hypothesis was not supported. Correlations were in the expected direction but were small in magnitude and failed to reach statistical significance for both lecture lateness (r = -.11, p = .06, N = 211) and lab lateness (r = -.03, p = .39, N = 132). Hypothesis 15 proposed a negative relationship between emotional stability and frequency of absenteeism behaviors. This hypothesis was supported. Emotional stability was negatively correlated with both lecture absenteeism (r = -.28, p < .01, N = 213) and lab absenteeism (r = -.16, p < .05, N = 212). The magnitude of the correlations observed suggests that emotional stability accounted for between 3 and 8% of the variance in absenteeism behaviors.

Hypothesis 16 proposed a negative relationship between emotional stability and permanent withdrawal. Prior to estimating the Cox hazards regression model, Kaplan-Meier estimates of the cumulative survival function and cumulative hazard function were examined based on a median split of the emotional stability predictor. These plots are shown in the top two panels of Figure 33. The plots suggest that hazard is slightly higher across time for participants with below-median scores on emotional stability. The overall shape of the hazard functions is roughly similar over time, and a plot of the partial residuals of the marginalization variable over time suggests that there is a slight positive trend. Taken together, this suggests that the assumption of proportional hazards is not clearly violated and analyses proceeded based on the time invariant predictor.

Cox hazards regression models were estimated in two steps. The first model estimated the effect of emotional stability as a sole predictor of the log cumulative hazard function. This model was not significant $\chi2$ (1, N = 213) = .35, p = .55, nor was the predictive effect of emotional stability (b = -.09, p = .55). The second model estimated the effect of emotional stability controlling for students' cumulative points, which was entered as a time-varying predictor. This model was significant, $\chi2$ (3, N = 213) = 8.30, $p < .05$, as was the predictive effect of time-varying cumulative points (b = -.01, $p < .05$); however, the effect of emotional stability was unchanged from the prior model (b = -.09, p = .58). Therefore, hypothesis 16 was not supported.

Personality traits and early departures. In addition to the hypothesized relationships with permanent withdrawal, lateness, and absenteeism, the relationships between conscientiousness and early departures were examined on an exploratory basis. Figures 34 and 35 show bivariate scatterplots and residuals plots for conscientiousness and emotional stability with the early departures variable. No obvious deviations from linearity or heteroscedasticity were observed. However, residuals plots revealed substantial non-normality, as evidenced by a higher proportion of residuals falling above the 0-residual reference line.

Correlational analyses revealed that emotional stability (*r* = -.14, *p* < .05, *N* = 211) but not conscientiousness (*r* = -.11, *p* = .06, *N* = 211) was significantly correlated with the frequency of early departures from lecture. Although statistically significant, the effect size for emotional stability was small, accounting for less than 3% of the variance in early departure behavior.

Exploratory analyses with agreeableness, extraversion, and openness to experience. The analytic process described above was repeated on an exploratory basis for the remaining three personality traits within the Five Factor Model. Bivariate scatterplots and residuals plots (not shown) supported similar conclusions as the prior analyses involving conscientiousness and emotional stability. Specifically, the assumptions of homogeneity of variance and multivariate normality were not met. As a result, parametric and non-parametric correlations with lateness and absenteeism behaviors are shown in Table 12. Statistical inferences were based on two-tailed significance tests.

Predictive relationships with agreeableness were non-significant in all cases, with no correlations exceeding ± .10. Extraversion was significantly correlated with early departures (*r* = .16, *p* < .05, *N* = 211) but was not correlated with any other absenteeism or lateness behaviors including early departures. The observed relationship indicates that participants with higher levels of extraversion departed lectures early with greater frequency; however, the magnitude of the effect was relatively small. Openness to experience was significantly correlated with lab absenteeism (*r* = .22, *p* < .01, *N* = 212) but was not correlated with any other absenteeism or lateness behaviors. The positive relationship observed suggests that participants with higher openness were absent from labs more frequently than participants with lower openness. Interestingly, both observed relationships suggest a negative effect of traits typically viewed as positive.

Finally, the relationship between these traits and permanent withdrawal was examined vis-à-vis separate Cox proportional hazards analyses within survival analysis. Examination of the Kaplan-Meier cumulative hazard and cumulative survival functions (plots not shown), as well as the scatterplots of residuals over time, indicated that the assumption of proportional hazards was met for each trait's relationship with hazard. However, separate Cox regression models indicated that these effects were small and non-significant (agreeableness: b = -.30, p = .22; extraversion: b = .02, p = .89; openness: b = -.10, p = .69), as well as after controlling for students' cumulative points (agreeableness: b = -.30, p = .21; extraversion: b = .02, p = .90; openness: b = -.10, p = .68).

Biodata Measures of Prior Withdrawal Behaviors. Prior withdrawal behaviors were assessed by multiple items that were mapped to six variables, each of which was tested separately as a predictor of behavioral withdrawal. The number of relationships examined in this subsection was prohibitive of reporting all bivariate scatterplots and residuals plots. Moreover, the visual tests of assumptions led to consistent conclusions and patterns which are described below in the aggregate. As with prior analyses, examination of the bivariate scatterplots and residuals plots indicated that the assumptions of homogeneity of variance and multivariate normality were generally violated. No instances of non-linearity were observed. Consequently, all parametric and non-parametric correlations with lateness and absenteeism behaviors are shown in Table 12.

Hypothesis 17 proposed a positive relationship between prior withdrawal behavior and the frequency of students' lateness behavior. Overall, the pattern of correlations observed partially supported this hypothesis. Lecture lateness was positively correlated with prior attendance problems (r = .14, p = .02, N = 211) but not with job tenure (r = -.02, p = .40, N = 205), the number of jobs held (r = .10, p = .08, N = 210), the number of universities attended (r = -.06, p = .21, N = 211), and the number of times plans (i.e., careers/majors) have changed (r = .03, p = .32, N = 210). Mixed evidence was observed for a positive relationship with the number of courses dropped (Pearson's r = .10, p = .08, N = 210 versus Spearman's rho = .14, p < .05, N = 210); however, the magnitude of each relationship suggests that its effect is relatively weak. On the other hand, lab lateness was positively correlated with the number of courses dropped (r = .16, p < .04, N = 131) and prior attendance problems (r = .29, p < .001, N = 132) but not with current and prior job tenure (r = -.08, p = .19, N = 126), number of jobs previously held (r = .02, p = .40, N = 131), the number of universities attended (r = .09, p = .14, N = 132), and the number of times plans (i.e., careers/majors) have changed (r = .03, p = .38, N = 131).

In summary, these findings partially supported biodata measures of prior withdrawal as predicting lateness behaviors in the course, particularly those facets of prior withdrawal focused on prior attendance problems and the number of courses students' have dropped. The magnitude of correlations observed suggests that, as separate predictors, these variables accounted for between 2 and 8% of the variance in students' lateness behavior.

Hypothesis 18 proposed a positive relationship between prior withdrawal behavior and the frequency of students' absenteeism behavior. Overall, the pattern of correlations obtained with lecture and lab absenteeism variables provides support for this hypothesis. Lecture absenteeism was positively correlated with the number of jobs previously held (r = .18, p < .01, N = 212), the number of courses previously dropped (r = .36, p < .001, N = 212), and prior attendance problems (r = .12, p < .05, N = 213). Alternatively, non-significant correlations were observed for current and prior job tenure (r = -.09, p = .10, N = 207) and the number of changes to career plans (r = .11, p = .06, N = 212). Mixed evidence for a negative relationship between the number of universities attended and lecture absenteeism was observed (Pearson's r = .09, p = .11, N = 213 versus Spearman's rho = -.13, p < .05, N = 213); however, the magnitude of both correlations suggests that this should be considered weak evidence that the number of universities attended has the opposite of anticipated effect on absenteeism (i.e., more universities attended associated with lower absenteeism). On the other hand, lab absenteeism was positively correlated with the number of jobs held (r = .17, p < .01, N = 211), the number of courses dropped (r = .29, p < .001, N = 211), the number of times plans (i.e., careers/majors) have changed (r = .16, p < .01, N = 211), and prior attendance problems (r = .21, p < .001, N = 212). Alternatively, non-significant correlations were observed for job tenure and the number of universities attended.

In summary, these findings supported biodata measures of prior withdrawal as predicting absenteeism behaviors in the course, particularly facets of prior withdrawal focused on the number of jobs students have held, the number of courses they have dropped, and the frequency of their past attendance problems. The magnitude of correlations observed suggests that, as separate predictors, these variables accounted for between 1 and 13% of the variance in students' absenteeism behavior.

Hypothesis 19 proposed a positive relationship between biodata measures of prior withdrawal and permanent withdrawal from the course. Examination of the Kaplan-Meier cumulative hazard and cumulative survival functions (plots not shown), as well as the scatterplots of residuals over time, indicated that the assumption of proportional hazards was met for the following biodata variables: number of jobs held, courses dropped, plans changed, and prior attendance problems. Therefore, Cox hazards regression proceeded for separately for each of these variables based on the simple time-invariant predictor models. Results were non-significant for the number of jobs held ($b = .09$, $p = .45$), the number of courses dropped ($b = .05$, $p = .55$), and the number of times plans (i.e., career/majors) have changed in the past ($b = .06$, $p = .44$), as well as after controlling for students' cumulative points (number of jobs: $b = .08$, $p = .45$; courses dropped: $b = .05$, $p = .59$; plans changed: $b = .06$, $p = .45$).

On the other hand, prior attendance problems was found to have a significant predictive relationship with withdrawal ($b = .07$, $p < .05$), as well as after controlling for students' cumulative points ($b = .07$, $p < .05$). Kaplan-Meier plots of the cumulative survival and cumulative hazard functions, as well as the log cumulative hazard plot based on Cox hazards estimates (bottom panel) are shown in Figure 36. As illustrated by the plots, greater attendance problems in the past were associated with an increased risk of permanent withdrawal over time, controlling for performance in the course.

Two additional biodata measures included current and prior job tenure and the number of universities students have attended on a full- or part-time basis. For these predictors, the assumption of proportional hazards was violated, as indicated by a strong positive trend in residuals over time. With the assumption of proportional hazards violated, alternative models were specified entering the main effects of these variables and their linear interaction term with time in days. In both cases, the overall models were non-significant (job tenure: $\chi 2$ (2, N = 213) = 3.44, p = .18; universities attended: $\chi 2$ (2, N = 213) = 4.45, p = .11), as were the interaction terms (job tenure: b = .01, p = .11; universities attended: b = .03, p = .11) and main effects (job tenure: b = -.71, p = .11; universities attended: b = -2.41, p = .10).

In summary, these analyses indicate that attendance problems had a significant predictive effect on permanent withdrawal in the course, whereas the remaining biodata variables failed to receive support. Therefore, the overall pattern of predictive relationships was not consistent, providing partial support for hypothesis 19.

Biodata measures of social embeddedness. Social embeddedness represented the total number of friends and relatives participants had in the introductory psychology course. Bivariate scatterplots and residuals plots for social embeddedness with absenteeism and lateness variables are shown in Figures 37 and 38. Scatterplots revealed moderate heteroscedasticity, with variance in each absenteeism and lateness behavior generally decreasing as the number of friends and relatives in the course (i.e., social embeddedness) increased. No obvious examples of non-linearity were observed. Residuals plots further emphasized heteroscedasticity and showed a similar pattern of non-normality as described in previous sections – that is, a disproportionate number of positive residuals, rather than equal distribution about the 0-residual line.

Hypothesis 20 proposed a negative relationship between social embeddedness and the frequency of lateness behaviors. This hypothesis was not supported, as evidenced by small and non-significant correlations with lecture absenteeism (r = .09, p = .09, N = 213) and lab absenteeism (r = .05, p = .23, N = 212). Hypothesis 21 proposed a negative relationship between social embeddedness and the frequency of absenteeism behaviors. This hypothesis was not supported, as evidenced by small and non-significant correlations with lecture lateness (r = .09, p = .09, N = 211) and lab lateness (r = -.08, p = .18, N = 132).

Hypothesis 22 proposed a negative relationship between social embeddedness and permanent withdrawal from the course. Examination of the Kaplan-Meier cumulative hazard and cumulative survival functions (plots not shown), as well as the scatterplots of residuals over time, indicated that the assumption of proportional hazards was upheld. Results of the Cox hazards regression analysis were non-significant (b = -.004, p = .95), indicating that social embeddedness did not impact withdrawal from the course. Thus, hypothesis 22 was not supported.

Biodata measures and early departures. In addition to the hypothesized relationships with permanent withdrawal, lateness, and absenteeism, the relationships between prior withdrawal behaviors, social embeddedness, and early departures were examined on an exploratory basis. Bivariate scatterplots and residuals plots displayed no obvious departures from linearity or heteroscedasticity. However, residuals plots revealed substantial non-normality, as evidenced by a higher proportion of residuals falling above the 0-residual reference line.

Correlational analyses supported a positive relationship between the number of jobs held and the frequency of early departures (*r* = .19, $p < .01$, *N* = 210) and provided mixed support for the predictive effects of job tenure (Pearson's *r* = -.16, $p < .05$, *N* = 205 versus Spearman's rho = -.07, *p* = .15, *N* = 205) and the number of courses dropped (Pearson's *r* = 11, *p* = .05, *N* = 210 versus Spearman's rho = .15, $p < .05$, *N* = 210). Together these results indicate that a higher number of jobs held, shorter job tenure, and a higher number of courses dropped are each associated with a higher frequency of early departures from class. However, the effects of job tenure and number of courses dropped should be qualified as relatively weak given the small magnitude of correlations observed (accounting for less than 3% of the variance in early departures) and mixed parametric and non-parametric findings. No other indicators of prior withdrawal demonstrated a statistically significant correlation with early departures including prior attendance problems (*r* = -.08, $p < .12$, *N* = 211).

Exploratory analyses with biodata measures of commute and intentions. Exploratory analyses examined the role of commuting method and students' intentions to withdraw (i.e., to be absent, late, or drop) in observed withdrawal behaviors. Bivariate scatterplots and residuals plots (not shown) supported similar conclusions as the prior analyses involving biodata measures of prior withdrawal and social embeddedness. Specifically, the assumptions of homogeneity of variance and multivariate normality were generally not upheld. Consequently, all parametric and non-parametric correlations with lateness and absenteeism behaviors are shown in Table 12. Statistical inferences were based on two-tailed significance tests.

Results indicated that commute distance, commute difficulty, and commute method-weather were uncorrelated with absenteeism or lateness behaviors across the lecture and lab. Commute method-independence exhibited mixed evidence with lab absenteeism, having a statistically significant negative Pearson's r (-.14, $p < .05$, N = 212) correlation but a non-significant Spearman's rho (-.08, $p = .22$, N = 212). These results tentatively suggest that students with more independent modes of commuting to campus (e.g., walking, biking, and driving as compared to riding with a friend) had slightly lower frequencies of lab absenteeism; however, the small magnitude of the correlation, accounting for less than 3% of the variance in lab absenteeism, in addition to the inconsistent parametric and non-parametric findings, suggest that this relationship should be qualified as relatively weak.

Students' intentions for daily withdrawal (i.e., judged likelihood of missing or being late for class) was positively correlated with lecture absenteeism ($r = .36$, $p < .001$, $N = 211$), lab absenteeism ($r = .32$, $p < .001$, $N = 210$), lecture lateness ($r = .26$, $p < .001$, $N = 209$), and early departures ($r = .26$, $p < .001$, $N = 209$) but not with lab lateness ($r = .14$, $p = .12$, $N = 130$). These relationships confirm that students' stated likelihood of missing or being late in the course consistently predicted their actual frequency of daily withdrawal. The magnitude of correlations observed suggest that intentions for daily withdrawal accounted for between 7 and 13% of the variance in withdrawal outcomes.

Intention to drop, on the other hand, was positively correlated with early departures ($r = .22$, $p < .01$, $N = 209$), exhibited a mixed pattern with lecture absenteeism (Pearson's $r = .17$, $p < .01$, $N = 211$ versus Spearman's rho = .13, $p = .06$, *N* = 211) and was uncorrelated with lab absenteeism (*r* = -.01, *p* = .90, *N* = 210), lecture lateness (*r* = .05, *p* = .44, *N* = 209), and lab lateness (*r* = -.01, *p* = .88, *N* = 130). Thus, as might be expected, students' daily withdrawal intentions were generally better predictors of actual daily withdrawal behaviors than were students' intentions to drop the course.

Finally, the predictive effects of commute method and intentions on the permanent withdrawal process were examined vis-à-vis separate Cox proportional hazards analyses. All variables involving students' commute had small and non-significant effects. In contrast, intentions for daily withdrawal was significantly associated with permanent withdrawal behavior over time, and intentions to drop exhibited a trend-level effect that is described below given the intuitive relevance of this variable. For daily withdrawal intentions, the assumption of proportional hazards was met. As anticipated, higher withdrawal intentions predicted greater hazard over time (*b* = .55, *p* < .01), as well as after controlling for students' cumulative points in the course (*b* = .53, *p* < .01). The observed hazard ratio of 1.73 indicates that there was a 73% increase in risk of permanent withdrawal associated with an increase of 1 point on the 5-point scale used to assess daily withdrawal intentions.

For intentions to drop, the assumption of proportional hazards was violated, as evidenced by a strong positive trend in residuals plotted over time, as well as the Kaplan-Meier plots shown in Figure 39. The cumulative survival and cumulative hazard plots indicate that intentions to drop had little predictive effect on hazard up to approximately 70 days, after which a marked increase in risk was associated with higher intentions to drop the course. The hazard functions further widened in the final days of the course. With the assumption of proportional hazards violated, an alternative model was specified entering both the "main effect" variable and a linear interaction term with time in days. This model was significant overall, $\chi2$ (2, N = 213) = 10.97, $p <$ .01; however, the interaction effect did not reach statistical significance (b = .02, p = .09), nor did the main effect of intentions to drop (b = -.96, p = .38). This suggests that, although the hazard rate for participants with above-median intentions to drop does increase toward the latter part of the semester, the magnitude of the effect in the overall sample is marginal.

Section III: Incremental Validity of CRT-W

This section examines the incremental validity of the CRT-W (and corresponding JMs) above and beyond personality and biodata as predictors of withdrawal. These analyses were undertaken using ordinary least square multiple regression. Models were developed and tested separately for each behavioral withdrawal variable, adding predictors in a hierarchical manner in three steps. Personality and biodata variables were entered in step 1, followed by the JMs for withdrawal in step 2. In step 3, students' end-of-semester cumulative point total was added. To limit the number of predictors entered at step 1, only those predictors with a significant bivariate correlation were entered into the model. In all cases, the ratio of sample size to predictors exceeded the minimum recommendations described by Tabachnick and Fidell (2007) and others (e.g., Green, 1991).

Assumptions of multivariate normality, linearity, and homogeneity of variance were examined based on the full model – that is, including all predictors included in the step 3 model. Bivariate scatterplots and residuals plots for the full models are shown in

Figures 40 and 41. Scatterplots reveal consistent heteroscedasticity, with variance in withdrawal behaviors increasing with increasing predictor values. The same pattern of heteroscedasticity is also evidence by residuals plots. Although the proportion of positive and negative residuals is more evenly distributed than in prior analyses, the residuals plots nevertheless point to moderate deviations from multivariate normality. Outlier screening led to the identification of no multivariate outliers. In addition, no instances of multicollinearity due to high predictor intercorrelations were identified, as evidenced by consistently high tolerance values (i.e., greater than .10). Intercorrelations among all study variables are shown in Table 13.

Lecture Lateness. The results of regression analyses for lecture lateness are show in Table 14. The omnibus test of model 1, which contained three biodata predictors (attendance problems, courses dropped, and intentions for daily withdrawal) was significant, F (3, 205) = 9.45, $p < .001$, accounting for a total of 12% of the variance in lecture lateness (adjusted $R^2 = .11$). The addition of the withdrawal JMs in model 2 resulted in a non-significant gain of 1% of the variance accounted for in lecture lateness, ΔF (3, 202) = 0.95, $p = .42$, $R^2 = .13$. The addition of students' cumulative point total in model 3 resulted in a statistically significant gain of 3% of the variance explained, ΔF (1, 201) = 7.23, $p < .01$, $R^2 = .16$. The adjusted R^2 for the final model of .14 suggests that there is minimal shrinkage, and that the predictors as a set account for approximately 15% of the variance in lecture lateness.

Based on the final model (i.e., model 3), two variables emerged as having a unique predictive effect: intentions for daily withdrawal and students' cumulative point total. Comparison of the standardized regression coefficients and semi-partial correlations revealed that intentions for daily withdrawal had the largest unique effect (□ = .25; semi-partial r^2 = .05) followed by students' cumulative points (□ = -.19; semi-partial r^2 = .03).

Lab Lateness. The results of regression analyses for lab lateness are show in Table 15. The omnibus test of model 1, which contained two biodata predictors (attendance problems and courses dropped), was significant, F (2, 130) = 7.33, $p < .01$, accounting for a total of 10% of the variance in lab lateness (adjusted R^2 = .09). The addition of the withdrawal JMs in model 2 resulted in a non-significant gain of 3% of the variance accounted for in lab lateness, ΔF (3, 127) = 1.46, p = .23, R^2 = .13. Although the overall gain was non-significant, the additional 3% of the variance accounted for in lab lateness might be indicative of a practically significant effect. Moreover, it should be pointed out that this effect's non-significance is likely due to the smaller sample size available for analyses involving lab absenteeism in comparison to the remaining criterion variables examined (n = 133 versus N = 212). Finally, the addition of students' cumulative point total in model 3 resulted in a non-significant gain of 1% of the variance explained, ΔF (1, 126) = 1.20, p = .28, R^2 = .14. The adjusted R^2 for model 2 of .10 suggests that there is minimal shrinkage, and that the predictors as a set account for approximately 10% of the variance in lab lateness.

Based on the final model (i.e., model 2), two variables emerged as having a unique predictive effect: prior attendance problems and the social injustice JM. Comparison of the standardized regression coefficients and semi-partial correlations revealed that prior attendance problems had the largest unique effect (□ = . 28; semi-partial r^2 = .07) followed by the social injustice JM (□ = .17; semi-partial r^2 = .03).

Most importantly, social injustice JM was found to have an incremental effect on lab lateness behaviors after controlling for biodata measures of prior withdrawal behaviors. The direction of the relationship suggests that, as predicted, higher scores on the social injustice JM were associated with higher frequencies of lateness in labs.

Lecture absenteeism. The results of regression analyses for lecture absenteeism are shown in Table 16. The omnibus test of model 1, which contained the personality and biodata predictors was significant, F (7, 201) = 11.41, p < .001, accounting for a total of 28% of the variance in lecture absenteeism (adjusted R^2 = .26). The addition of the withdrawal JMs in model 2 resulted in a non-significant gain of 1% of the variance accounted for in lecture absenteeism, ΔF (3, 198) = 0.66, p = .58, R^2 = .29. The addition of students' cumulative point total in model 3 resulted in a statistically significant gain of 23% of the variance explained, Δ F (1, 197) = 97.18, p < .001, R^2 = .53. The adjusted R^2 for the final model of .50 suggests that there is minimal shrinkage, and that the predictors as a set account for approximately half of the variance in lecture absenteeism. As described previously this might, however, be a slight overestimate of the prospective effect of these variables on subsequent withdrawal behavior, given the reverse effect of absenteeism on students' performance in the course.

Based on the final model (i.e., model 3), five variables emerged as having a unique predictive effect: conscientiousness, emotional stability, the number of courses dropped, intentions for daily withdrawal, and students' cumulative point total. Comparison of the standardized regression coefficients and semi-partial correlations revealed that cumulative points had the largest unique effect (□ = -.54; semi-partial r^2 = .23) followed by the number of courses dropped □ =(.23; semi-partial r^2 = .05), intentions for daily withdrawal (□ = .17; semi -partial r^2 = .02), emotional stability (□ = -.13; semi-partial r^2 = .02), and conscientiousness (□ = -.10; semi-partial r^2 = .01). In general, this suggests that course performance and previous withdrawal behaviors carried the largest effects followed by smaller unique effects of explicit personality traits.

Lab absenteeism. The results of regression analyses for lab absenteeism are show in Table 17. The omnibus test of model 1, which contained the personality and biodata predictors was significant, F (8, 200) = 9.34, p < .001, accounting for a total of 27% of the variance in lab absenteeism (adjusted R^2 = .24). The addition of the withdrawal JMs in model 2 resulted in a marginally significant gain of 3% of the variance accounted for in lab absenteeism, ΔF (3, 197) = 2.60, p = .05, R^2 = .30. The addition of students' cumulative point total in model 3 resulted in a statistically significant gain of 15% of the variance explained, ΔF (1, 196) = 55.33, p < .001, R^2 = .45. The adjusted R^2 for the final model of .42 suggests that there is minimal shrinkage, and that the predictors as a set account for approximately 40% of the variance in lab absenteeism. As described previously this might, however, be a slight overestimate of the prospective effect of these variables on subsequent withdrawal behavior, given the reverse effect of absenteeism on students' performance in the course.

Based on the final model (i.e., model 3), five variables emerged as having a unique predictive effect: openness to experience, the number of courses dropped, intentions for daily withdrawal, the revocable commitment JM, and students' cumulative point total. Comparison of the standardized regression coefficients and semi-partial correlations revealed that cumulative points had the largest unique effect (□ = -.42; semi-partial r^2 = .15) followed by openness to experience (□ = .21; semi-partial r^2 = .04), the number of courses dropped□(= . 17; semi-partial r^2 = .02), intentions for daily withdrawal (□ = .15; semi-partial r^2 = .02), and the revocable commitment JM (□ = .14; semi-partial r^2 = .02). In general, this suggests that course performance carried the largest unique effect on lab absenteeism, with personality traits and previous withdrawal behaviors carrying smaller but statistically reliable effects.

Most importantly, revocable commitment was found to have an incremental effect on lab absenteeism after controlling for the other variables in the model, including course performance and explicit personality traits. The direction of the relationship suggests that, as predicted, higher scores on the revocable commitment JM were associated with higher frequencies of absenteeism in labs.

Early departures. The results of regression analyses for early departures from lecture are show in Table 18. The omnibus test of model 1, which contained the personality and biodata predictors, was significant, F (7, 194) = 6.70, p < .001, accounting for a total of 20% of the variance in early departures (adjusted R^2 = .17). The addition of the withdrawal JMs in model 2 resulted in a non-significant gain of 2% of the variance accounted for in early departures, ΔF (3, 191) = 1.47, p = .22, R^2 = .21. The addition of students' cumulative point total in model 3 resulted in a statistically significant gain of 4% of the variance explained, ΔF (1, 190) = 10.80, p < .01, R^2 = .26. The adjusted R^2 for the final model of .21 suggests that there is minimal shrinkage, and that the predictors as a set account for just over 20% of the variance in early departure behavior.

Based on the final model (i.e., model 3), four variables emerged as having a unique predictive effect: extraversion, the number of jobs held, job tenure, and students' cumulative point total. Comparison of the standardized regression coefficients and semi-partial correlations revealed that cumulative points had the largest unique effect (□ = -.23; semi-partial r^2 = .04) followed by the number of jobs previously held (□ = .19; semi-partial r^2 = .03), job tenure (□ = -.18; semi-partial r^2 = .03), and extraversion (□ = .13; semi-partial r^2 = .02). In general, this suggests that course performance and previous withdrawal behaviors carried the largest effects followed by smaller unique effects of biodata measures of prior withdrawal and extraversion.

Permanent withdrawal. Analyses in this section examined whether the interaction effects observed revocable commitment and overall CRT-W scores with time held after controlling for biodata predictors of permanent withdrawal. As with the preceding incremental tests, models were developed starting with the variables that were supported by prior "bivariate" tests with permanent withdrawal. Therefore, prior attendance problems and daily withdrawal intentions were entered in a step 1 Cox hazards regression model. Step 2 models assessed the incremental effects of adding the main effect and linear interaction terms for revocable commitment and overall CRT-W scores.

Results indicated that the step 1 model was significant overall, $\chi 2$ (2, N = 213) = 9.55, p < .01. Intentions for daily withdrawal was significant (b = .52, p < .01), indicating an increase in hazard associated with increasing daily withdrawal intentions. Attendance problems exhibited a trend-level effect (b = .07, p = .05), providing tentative support for an increased risk associated with greater attendance problems in the past. Adding the revocable commitment variables to the step 2 model did not produce a significant change, $\Delta\chi 2$ (2, N = 213) = 4.34, p = .11. The interaction term exhibited a trend-level effect (b = .10, p = .06), suggesting that the increasing impact of revocable commitment on permanent withdrawal over time is diminished when controlling for biodata predictors of withdrawal.

In contrast, an incremental effect was observed for overall CRT-W scores in the step 2 model. Specifically, the addition of the main effect and linear interaction term variables led to a significant change, $\Delta\chi 2$ (2, N = 213) = 7.01, $p < .05$. Likewise, the linear interaction term of overall CRT-W scores by time was statistically significant (b = .19, $p < .05$). This indicates that overall CRT-W scores have a time contingent effect on permanent withdrawal controlling for biodata measures of past attendance problems and intentions for daily withdrawal.

Reflecting on the set of analyses examining the incremental effects of conditional reasoning variables beyond personality and biodata predictors, partial support was obtained for hypothesis 23. While it is clear that the biodata indicators of prior withdrawal behaviors and intentions to withdraw accounted for the largest percentage of variance explained in criteria, some positive evidence was gained for the incremental contribution of the revised conditional reasoning test to lab lateness (social injustice), lab absenteeism (revocable commitment), and permanent withdrawal (overall CRT-W scores).

CHAPTER 4

DISCUSSION

Development of a conditional reasoning instrument that satisfies each of several considerations related to construct validity is an iterative process that will involve multiple validation and cross-validation studies. The reviews by James et al. (2004, 2005) highlight how the published CRT-A was preceded by several developmental test versions and many items that were modified or ultimately discarded. In total, the final test version was a product of nearly a decade of research and many independent validation studies with a variety of samples and criteria. As such, it is important to view the evidence presented in this study as a first step toward evaluating a developmental version of a conditional reasoning test of withdrawal. Consistent with the organization of the study's Results, the Discussion is organized around two main topics, insights into the development of a conditional reasoning measure and insights surrounding the substantive relationships between predictor variables and behavioral withdrawal criteria.

Insights into the Development of a CRT-W

A large component of the scale evaluation and refinement portion of this study was focused on testing the criterion-related validity of conditional reasoning items. Results highlighted a subset of 13 items that demonstrated positive evidence of criterion-related validity with one or more specific forms of withdrawal behavior. It is possible to view the predictive potential of these items at several different degrees of scrutiny. For the purpose of the present study, the evaluative criteria for dropping items was conservative, or stated alternatively, favoring item retention. This was done for two main reasons: (a) to allow for potential additive effects of items to emerge in combination and (b) to manage the level of capitalization due to chance built into subsequent hypothesis testing. In deciding which items to retain for subsequent research, however, it may be beneficial to adopt a more stringent criterion. For example, if one focuses on the subset of items that exhibited significant positive item-criterion correlations with two or more behavioral withdrawal criteria, seven items would be identified for retention, including five items from the revocable commitment JM and two items from the social injustice JM. Results suggest that these items provide a stronger baseline for subsequent iterations than the full 13 items retained here, having demonstrated some evidence of consistency across different indicators and types of withdrawal behaviors. It is further telling – and consistent with the findings of hypothesis tests (see below) – that no items from the marginalization of withdrawal JM satisfied this modified criterion. At the other end of the spectrum, there was a subset of items that exhibited no evidence of predictive potential and can thus be confidently discarded from further developmental iterations.

Moreover, the items that tended to exhibit the highest predictive validities also tended to have relatively low p-values (e.g., < .40). Following the logic of James and colleagues with the CRT-A (James & LeBreton, 2011), there is a conceptual basis for this observation. These authors have described implicit aggression JMs as occurring at a low overall base rate in the population in order to account for aggressive and harmful behaviors that also occur at a relatively low base rate in the population. Thus, the predictor and criterion have similar distributional properties. A similar, albeit less extreme, situation may hold in reference to withdrawal behaviors which are often positively skewed (Harrison, 2002). For the classroom withdrawal behaviors followed in the present study, the vast majority of students had no observed episodes of withdrawal. In contrast, it was a small proportion of students that filled out the positive tail of severely skewed distributions for each withdrawal behavior. Therefore, future development efforts focused on shifting items toward more stringent thresholds for JM endorsement may help to improve the scale's predictive validity.

A clear shortcoming of the present instrument was the low internal consistency reliability of the initial and refined instrument. Alternative item groupings did not improve scale reliability, including exploratory tests of the most highly intercorrelated item subsets. Rather, across the board, items generally exhibited low inter-item and item-total correlations, pointing to the fact that item content was largely heterogeneous. This was somewhat surprising given that item generation was guided by the conceptualization of the JMs and reviewed by SMEs for face validity. Nevertheless, these results indicate that future developmental efforts should concentrate on: (a) writing item sets that are somewhat narrower in focus and (b) understanding respondents' thought process when reading the premises and arriving at an inductive solution, for example, by undertaking a verbal protocol analysis.

Insights into Substantive Relationships with Withdrawal

Table 19 summarizes the results of hypothesis testing. Overall, results were mixed for hypotheses linking the conditional reasoning variables to behavioral withdrawal criteria. One clear outcome was that hypotheses involving the marginalization of withdrawal JM were not supported. This was evident in both the results of the item-analyses, in which no marginalization items displayed a consistent pattern of predictive relationships across criteria, and subsequently in the null results of hypothesis testing. These results suggest that this JM ought to be re-conceptualized prior to future research, or perhaps, abandoned in favor of an intensified focus on the social injustice and revocable commitment concepts that received greater support in the present study.

Reflecting on the content of the JMs, it could be suggested that the marginalization JM had the disadvantage of being too broadly defined. In general, the marginalization of withdrawal bias was designed to capture the subtle tendency to rationalize withdrawal behaviors, such as when under-estimating the seriousness and frequency of withdrawal. However, it might be argued that both the revocable commitment and social injustice JMs provide a more focused investigation of these rationalization processes – for revocable commitment, in specific ways related to malleable beliefs about obligation and reciprocity, and for social injustice, in specific ways related to the use of fairness (and unfairness) as a route for justifying withdrawal. This insight suggests that future iterations of the JMs and accompanying measure might "learn" from the social injustice and revocable commitment concepts by specifying narrower rather than broader content domains.

Among the three JMs that were studied, revocable commitment demonstrated the strongest positive evidence, receiving partial support for hypotheses linking this JM to permanent withdrawal and demonstrating incremental variance beyond explicit personality traits and biodata predictors for absenteeism in labs. In addition, this JM had the highest proportion of items that exhibited positive criterion-related evidence (6 of 10 overall). Nevertheless, it is important to reiterate that the positive forms of evidence obtained were relatively modest, including that the relationships observed had small effect sizes and did not hold consistently across criterion variables (hence, the qualified “partial support”). For example, higher scores on revocable commitment were associated with higher frequencies of lab absenteeism but not lecture absenteeism.

The evidence linking revocable commitment to permanent withdrawal is similarly qualified by the linear interaction with time. As illustrated through Kaplan-Meier plots of the cumulative hazard functions, revocable commitment’s impact on withdrawal emerged only in the final portion of the semester. Through the first 80 days of the semester, participants with above- and below-median scores on the revocable commitment JM experienced very similar levels of risk. However, through the final 20 days of the semester, the above-median participants experienced a level of risk that was roughly two times as high as below-median participants. One interpretation is that the effects of revocable commitment on the withdrawal process take time to unfold and best differentiate the “later quitters” from “late stayers” as opposed to “early quitters” from “early stayers.” At the same time, the robustness of this finding is challenged by the small number of quitters that were in the study’s sample. For example, the diverging hazard functions in the last 20 days are driven largely by the 15 students that withdrew during this time period, underscoring the need to replicate this pattern in future studies.

Whereas revocable commitment was correlated with absenteeism and turnover processes, social injustice’s predictive relationships were observed for lateness behaviors, and specifically the frequency of lateness to the lab. This relationship indicated that a stronger tendency toward a social injustice rationalization process was associated with a higher frequency of lateness behaviors. The effect observed was small, albeit statistically significant, accounting for approximately 2% of the variance in lab lateness. Moreover, this relationship held after controlling for personality and biodata predictors as well as students' cumulative points in the course, thereby *partially* supporting the social injustice JM's incremental validity in predicting lateness behaviors. The indication of "partial" support in this case reflects the fact that a similar positive relationship was not observed for lecture lateness. Interestingly, a negative relationship was noted with performance in the course, such that higher scores on the social injustice JM predicted lower end-of-semester point totals. However, the magnitude of the effect was again modest, accounting for approximately 2% of the variance. Nonetheless, a future direction for research linking the JMs to withdrawal criteria involves examining potential direct and indirect (through withdrawal behaviors) effects on performance criteria.

Inclusion of the tests linking overall scores on the conditional reasoning measure to withdrawal behaviors was designed to address a practical question: What is the predictive validity of the measure as a whole? The partial support obtained for hypotheses with lateness, absenteeism, and permanent withdrawal reflects the combined strengths of the revocable commitment and social injustice JMs as predictors of lab absenteeism and permanent withdrawal (revocable commitment) and lab lateness (social injustice). However, as with the JM-level relationships, the predictive relationships based on the combined scale score were not strong overall, accounting for between 1 and 2% of the variance in withdrawal behaviors. Though statistically significant, these effects might not connote practical significance. Future iterations of the measure can lead to improved practical significance by retaining the items with the highest predictive potential from this study and developing new items that better balance predicting the criteria of interest while improving the measure's reliability.

Personality and Biodata Predictors of Withdrawal

Beyond the conditional reasoning test, insights were gained into personality and biodata predictors of withdrawal in an academic context. For explicit personality traits from the Five Factor Model, conscientiousness and emotional stability were negatively correlated with the frequency of absenteeism behaviors both in lecture and in the lab. Higher levels of emotional stability (but not conscientiousness) were also associated with a slightly lower frequency of early departures from the lecture. In addition, exploratory analyses linked higher levels of extraversion and openness to experience – traditionally viewed as positive traits – with higher frequencies of early departure and lab absenteeism, respectively. Overall, these findings add to prior studies that have reported predictive relationships for trait-based assessments with absenteeism and lateness behaviors (e.g., Ones et al., 2003). However, results failed to replicate prior studies' findings that personality traits, particularly conscientiousness and emotional stability, are similarly related to permanent forms of withdrawal (Barrick & Zimmerman, 2005; 2009). Specifically, none of the traits from the Five Factor Model were linked to the occurrence and timing of permanent withdrawal from the course.

In comparison to all other predictors investigated (aside from students' cumulative points in the course), biodata measures of past withdrawal behaviors and intentions to withdraw demonstrated the strongest and most consistent predictive relationships with withdrawal criteria. Among biodata measures, the two that clearly stood out as most predictive were *prior attendance problems* (assessed by two items asking about the frequency of being disciplined for missing or being late to work and school) and the *number of courses students have dropped in the past*. Prior attendance problems exhibited small-to-moderate relationships with lateness behavior in labs and lecture, absenteeism behaviors in labs and lecture, and a time-contingent effect with permanent withdrawal (increasing hazard over time associated with greater past problems). The number of courses students have dropped was correlated – also in the small-to-moderate range – with lateness behaviors in the lab and absenteeism behaviors in both lecture and labs. Unexpectedly, the number of courses dropped was not associated with the timing and occurrence of permanent withdrawal in the present study. Nevertheless, the overall pattern of results supports prior authors' conclusion that withdrawal behaviors are relatively stable within-persons over time and across situations; prior attendance problems correlated with current attendance problems and did so across the separated contexts of labs and lectures. This speaks to the potential of stable individual differences to contribute to a heightened readiness for withdrawal, or as Frogatt (1970) initially described, a "withdrawal proneness."

Of course, the question at the heart of the present study is whether such individual difference factors take on a more explicit or implicit nature. Although this question will ultimately be answered only through the programmatic study of implicit withdrawal tendencies in tandem with improved methods and instruments for their measurement, the fact that stated intentions to withdraw (an explicit measure) does indeed correlate with actual withdrawal behaviors – consistently across all criteria for daily withdrawal intentions – points to a certain level of awareness of withdrawal tendencies and a degree of calculation as opposed to spontaneity in the manifestation of the behaviors. Of course, this does not diminish the potential importance of underlying implicit dispositional mechanisms, but it does suggest that overt intentions are an important component that ought to be understood alongside potential covert dispositional sources. It also bears mentioning that the practical utility of a biodata indicator of daily withdrawal intentions in an employment selection context may be substantially reduced due to respondents' self-enhancement motivations (Barrick & Zimmerman, 2005). This emphasizes the very practical, as well as academic, impetus for continuing the search for implicit dispositional sources of behavior, particularly those with clear detrimental consequences.

This research builds on prior studies investigating the conditional reasoning test of aggression. That line of research demonstrates that withdrawal behaviors, specifically absenteeism and turnover, can serve as one manifestation of an implicit motive to aggress (James et al., 2004, 2005). In other words, withdrawal can occur as part of an underlying hostility or retribution motive that is held and targeted toward the organization or other individuals in the workplace. Taken together with the present study, this research not only supports an implicit dispositional source of withdrawal behaviors but also points to the potentially diverse domains of the implicit personality that are relevant.

Limitations

It is important to recognize several limitations of the present research, two of which are specific to the methods and results surrounding the conditional reasoning instrument and two that have broader implications for the generalizability of the present study's results. Consistent with the former, the first limitation has to do with the use of a single same sample for scale development and validation purposes and the related concern of inflation due to capitalization on chance. Ideally, two completely independent samples would be used for these purposes to establish a pure test of cross-validation for item and scale validity. Though this wasn’t possible within the scope of the present study, several noteworthy factors mitigate against the potential for relationships to be inflated due to the use of single sample. The most obvious is that the pattern of predictive relationships themselves. In other words, although the conditional reasoning items with the strongest positive evidence of criterion-related validity were retained, their combination into scales for the purpose of hypothesis testing did not result in stronger correlations in the vast majority of cases. Instead, many of the predictive relationships observed at the item level were slightly reduced when items were combined into scales.

A related issue involves the use of multiple criteria to test predictive relationships. The potential for inflation would have been much greater had the revised sub-scales been tailored to particular criteria where the observed item-level relationships were consistently strong. Rather, the more general approach adopted for retaining items based on positive evidence with one or more criterion variables can be considered a conservative approach to scale refinement. Alternatively, a more aggressive approach could involve tailoring sub-scales to the prediction of a single criterion-type; however, the related disadvantage would be a much clearer potential for inflated scale validity coefficients with respect to those same criteria.

Along the same lines, it was the author’s judgment that use of a hold-out sample to separate the development and validation phases of the study was a less optimal strategy than taking the steps described above in tandem with examining statistical criterion for potential over-fitting of relationships to the present sample. Monte Carlo studies have found holdout or single sample cross-validation methods to approximate formula-based shrinkage estimations while using up valuable degrees of freedom (i.e., because of the divided sample), such that shrinkage formulas are a generally more favorable approach to assessing the effect of random sampling error (Murphy, 1983). In the present study, shrinkage estimates were examined and generally did not point to a significant loss in predictive validity due to the potential for over-fitting. Nevertheless, Murphy notes that neither shrinkage corrections nor holdout methodologies are sensitive to non-random sampling errors and reminds us that true cross-validation implies use of a second *independent* sample. Thus, cross-validation of the items that are retained is an obvious concern and one that can be best satisfied by subsequent research within an ongoing program.

Another limitation associated with the conditional reasoning measure has to do with the low internal reliability of the JM variables. Though this point has been discussed already in the discussion of scale revision considerations, it represents a limitation in the context of hypothesis testing as well. The ability of the JMs to correlate with an external criterion was severely attenuated due to the low reliability of underlying measures. Taking an optimistic perspective on this problem, the relationships that emerged are perhaps all the more impressive, having overcome a strong psychometric disadvantage. From a different perspective, the low reliability raises questions about construct validity that need to be addressed by future iterations of the measure and reinforces that the results of the present study ought to be interpreted cautiously.

Two additional limitations are related to the design and sample used in the present research. An ideal design for a predictive study in this context would involve collecting all predictor measures prior to the start of the semester. Although predictors were collected for the majority of the sample early on in the semester, the precise timing varied across participants with a subset completing the predictor measures in the latter half of the semester. This introduces two related concerns. The first is that the participants in the present sample who permanently withdrew may not be representative of all the students in the course that withdrew, particularly those that withdrew in the early portion of the course. Obviously, it was not possible for a student who withdrew in the first few weeks of the semester to enroll in the study thereafter. A second concern is that differences in timing of the predictor collection introduced time as a potential confound for some of the variables assessed in the present study, most notably those that are likely to change over the course of a semester. Fortunately, the variables of greatest substantive interest in the present study are not likely to change over the course of a semester, including explicit measures of personality based on the Five Factor Model, implicit measures of personality based on the conditional reasoning test, and biodata measures assessing prior histories of withdrawal-related behaviors.

In contrast, the variables most likely impacted by the timing of predictor collection are social embeddedness (e.g., number of friends in the course) and withdrawal intentions. For example, social embeddedness may increase as students have more time to interact with their classmates, and intentions to withdraw may change as students receive feedback about their performance. Beyond shifts in mean levels, the more important issue is how timing may impact the relationship between social embeddedness (intentions to withdraw) and criterion variables. It seems further plausible that these variables' impact on withdrawal would be enhanced over time. For example, intentions may be a stronger predictor as the semester progresses and students have a better sense of what to expect in the classroom, advanced information about their performance in the course, and so on. By extension, these predictors may have an "unfair advantage" in comparison to predictors that are not expected to benefit from the passage of time (e.g., the personality and biodata variables). However, no upward bias was evident based on the predictive relationships observed for social embeddedness, which was uncorrelated with all criteria in the present study. The intentions variables on the other hand may have a slight upward bias built into their predictive validities, which would have made it more difficult for the conditional reasoning variables to demonstrate incremental validity.

One final limitation involves the generalizability of a student sample and the academic context examined in the present study. As previously described, there are a number of similarities between academic and work contexts that make the classroom a useful context for studying antecedent factors in work-related processes. At the same time, there are a few unique features of academic contexts that deserve mention and have implications for the types of contextual features that could be incorporated in future evaluations of the implicit JMs underpinning withdrawal behavior. The most obvious is that the semester course is an inherently short-term obligation with a clearly defined end. In this context, withdrawal prone individuals may be able to "hang on" in order to complete the course and receive academic credit. In contrast, work arrangements and other longer-term obligations may not have a clear end-in-sight. In those “weaker” contexts individuals would have (a) a longer period of time to manifest implicit withdrawal tendencies, and (b) a greater opportunity to define the terms of the commitment as revocable. Given these considerations, future studies should build on the present research by examining the predictive effects of implicit dispositional tendencies to withdraw in scenarios involving longer and more fluid commitments, such as mobility across jobs or organizations. On the other hand, it is worth pointing out that these differences seem less readily apparent with respect to lateness and absenteeism behaviors for the reasons described previously (e.g., both contexts provide an opportunity for these behaviors on a daily basis, both have norms and consequences surrounding attendance behaviors, and so on).

ENDNOTES

[1] It should be noted that recent unpublished studies have focused on developing and validating alternative item formats, including differential framing and reading comprehension items.

[2] Standardized skewness is derived as the skewness statistic divided by its standard error. Standardized skewness values greater than 3.29 or less than -3.29 indicated a statistically significant skew ($p < .001$).

[3] The term sensitivity analysis is used broadly to describe methods that examine the consequence of statistical and methodological decisions. In particular, sensitivity analysis has gained popularity in meta-analysis (e.g., see Greenhouse & Iyengar, 2009) but can be applied in various methodological domains.

[4] It should be noted that the positive direction of the regression coefficient is in contradiction to the visual depiction of marginalization's effect on the hazard function using Kaplan-Meier estimated cumulative survival and hazard plots. This discrepancy is an artifact of the dichotomization of the marginalization variable for plotting purposes. In other words, the dichotomized variable has a slight, non-significant negative direction, whereas the continuous variable has a slight, non-significant positive direction. Nevertheless, both results point to the fact that marginalization's effect on the hazard function is small and not statistically significant. The visual plots using the dichotomized variable are informative for descriptive purposes; however, their potential to conflict with the results of Cox hazards regression (using the continuously defined variable) when the modeled effects are small should be recognized.

Table 1

Justification Mechanisms (JMs) Supporting a Dispositional Tendency to Withdraw

JM	Theoretical Background	Description
Marginalization of Withdrawal Bias	Social psychological literatures on dissonance and cognitive rationalization processes, which suggest that people will seek to minimize, normalize, and rationalize their negative behaviors in an attempt to protect ego.	Presence of JM: • Absenteeism and tardiness are not serious problems; lenient consequences are appropriate; control is external (i.e., extraneous factors are to blame); frequency of others' absenteeism and tardiness is over-estimated. Absence of the JM: • Absenteeism and tardiness are serious problems; discipline is an appropriate consequence; accountability is ultimately within-person; absenteeism and tardiness are rare.
Revocable Commitment Bias	Commitment as a central construct in cognitive models of withdrawal. Normative and continuance commitment as guides for beliefs about obligation and reciprocity.	Presence of JM: • Decisions (e.g., to accept a job) are revocable; commitments are short-term and/or evolving; conception of contractual obligation is loose. Absence of the JM: • Decisions (e.g., to accept a job) are consequential; commitments are relatively binding; fulfillment of contractual obligations is responsible.
Social Injustice Bias	Equity and fairness theories, which suggest that individuals interpret organizational events relative to the experience of others and differ in their preferences and sensitivity to equity.	Presence of JM: • Referent cognitions are relied upon heavily; predisposed to feel inequity (i.e., the grass is always greener elsewhere); others' behavior (e.g., absenteeism) is gauge for appropriate self-behavior; sensitivity to injustices and use of injustice to justify bad behavior. Absence of the JM: • Referent cognitions are downplayed or infrequent; less sensitive to feelings of inequity; unlikely to view inequity as justification for withdrawal; reliance on internalized moral reasoning.

Table 2

Demographics for the Scale Development (*N* = 243) *and Hypothesis Testing Subsamples* (*N* = 213)

Variable	Scale Development *N* (% of total)	Hypothesis Testing *N* (% of total sample)
Sex [a]		
female	168 (66.4)	142 (66.7)
male	83 (32.8)	69 (32.4)
Race/ethnicity [b]		
black / African American	97 (38.3)	71 (33.3)
white / Caucasian	93 (36.8)	86 (40.4)
mixed ethnicity	15 (5.9)	10 (4.7)
Asian / Asian American	14 (5.5)	13 (6.1)
Arab American	13 (5.1)	13 (6.1)
Hispanic / Latino	12 (4.7)	11 (5.2)
Native American	1 (0.4)	1 (0.5)
other	6 (2.4)	6 (2.8)
Year [c]		
freshman	110 (43.5)	95 (44.6)
sophomore	90 (35.6)	75 (35.2)
junior	40 (15.8)	34 (16.0)
senior	9 (3.6)	7 (3.2)
Status		
full-time	230 (90.9)	196 (92.0)
part-time	19 (7.5)	15 (7.0)
Major field of study		
nursing	43 (17.0)	35 (16.4)
psychology	38 (15.0)	32 (15.0)
undecided	26 (10.3)	24 (11.3)
business	22 (8.7)	15 (7.0)
education	18 (7.1)	16 (7.5)
physical therapy	9 (3.6)	7 (3.3)
accounting	8 (3.2)	8 (3.8)
biology	7 (2.8)	6 (2.8)
social work	7 (2.8)	5 (2.3)
criminal justice	7 (2.8)	4 (1.9)
other	68 (26.9)	59 (28.0)
Age [d]	20.54 (4.44)	20.63 (4.74)
Total credits earned [d]	24.94 (25.8)	24.80 (26.53)
Semester credits [d]	13.16 (1.82)	13.20 (1.81)

Notes. [a]Information about sex was not available for 2 individuals in the total sample. [b]Information about race/ethnicity was not available for 2 individuals in the total sample. [c]Information about year in undergraduate education was not available for 3 individuals in the total sample. [d]Mean and standard deviation are shown.

Table 3

Weekly Timeline for Withdrawal Behaviors from Course Lecture and Laboratory Meetings

	Wk-1	Wk-2		Wk-3		Wk-4		Wk-5		Wk-6		Wk-7		Wk-8	
Criteria	9-2	9-7	9-9*	9-14*	9-16	9-21	9-23	9-28	9-30	10-5	10-7	10-12	10-14	10-19	10-21
Lecture															
Absenteeism			X	X		X	X	X	X	X	X	X	X	X	X
Lateness			X	X		X				X	X	X		X	
Early depart				X		X				X		X		X	
Lab [a]															
Absenteeism		X		X		X		X		X		X		X	
Lateness		X		X		X		X		X		X		X	

	Wk-9		Wk-10		Wk-11		Wk-12		Wk-13	Wk-15		Wk-16		Wk-17
Criteria	10-26	10-28	11-2	11-4	11-9	11-11	11-16	11-18	11-23	11-30	12-2	12-7	12-9	12-9
Lecture														
Absenteeism	X	X	X	X	X	X	X	X	X	X	X	X	X	
Lateness	X		X	X	X		X	X	X	X		X		
Early depart	X				X		X		X	X		X		
Lab [a]														
Absenteeism	X		X		X		X			X				
Lateness	X		X		X		X			X				

Notes. [a]Labs were held between Tuesday and Friday of designated weeks. Lateness criteria for labs were available for 9 of 15 lab sections (N = 137).
Xs indicate the availability of withdrawal criteria by week and day.
* indicates consent days for student clicker data.

Table 4

Full CRT-W Test Items and Scoring Key

Items	Test order [a]	JM	Non-JM
Marginalization of Withdrawal			
1. Many universities are using a delayed schedule…*	4	c	a
2. High school attendance policies are usually strict…	6	d	a
3. It is common for airlines to overbook…	9	a	c
4. The harshness of punishment is associated with its frequency…	13	d	a
5. There is a popular story about an employee who made a mistake…*	16	b	d
6. Company policies are always changing…	18	a	c
7. Replacing a good employee that quits can be costly…	20	b	c
8. The daily commute has ranked among the most stressful events…*	26	a	d
9. A quote by a famous movie director…	28	c	b
Revocable Commitment			
1. Online college classes are not recommended for all students…*	3	a	b
2. The old saying, "there are a lot of fish in the sea…*	5	b	c
3. More first-time marriages are occurring later in adulthood…*	7	a	c
4. New Year's resolutions are rarely effective…*	11	c	d
5. Shaking hands is important in business…*	14	d	a
6. A pre-nup specifies how a couple's wealth is divided…*	19	c	a
7. Losing a job can be one of life's most stressful events…	22	c	d
8. Being labeled a "flip-flopper" during an election…	25	b	d
9. Successful law firms are known for their lawyer pipeline…*	27	a	d
10. Current versus past generation parenting styles…	29	a	c
Social Injustice			
1. Scheduling shifts for doctors on holidays and weekends…	10	b	d
2. A current issue being debated in many universities…	12	a	c
3. Some have argued that labor unions are unnecessary…*	17	c	d
4. It is well known that many high school students skip school…*	21	d	b
5. The old saying, "If you can't beat em, then join em"…	23	d	b
6. An old story involving a man and his dying wife…*	30	a	d

Notes. [a]Test order indicates the question number from the CRT-W test booklet (see Appendix A).
JM indicates the withdrawal justification response. Non-JM indicates the oppositely valenced or non-withdrawal response. The two multiple-choice options not shown for each item are the distractor responses.
* denotes items that were retained for hypothesis testing.

Table 5

Mapping of Items to Biodata Themes and Corresponding Scaling Details

Biodata Theme Variable name	Location in Questionnaire [a]	Scaling Details
Prior Job Changing		
Job tenure	C – 1, 2	number of years (interval/ratio)
Number of jobs held	C – 3	number (interval/ratio)
Prior University and Course Changing		
Universities attended	C – 5, 6	number (interval/ratio)
Courses dropped	C – 7, 8	number (interval/ratio)
Plans changed	C – 4, 11	number (interval/ratio)
Social Embeddedness		
Friends and relatives	C – 9, 10	number (interval/ratio)
Prior Attendance		
Attendance problems	C – 12, 13	number (interval/ratio)
Method and Difficulty of Commute		
Commute distance	C – 14	number of miles (interval/ratio)
Commute difficulty	C – 16	Likert-type (ordinal/interval)
Commute method-independence	C – 15	multiple-choice (nominal)
Commute method-weather	C – 15	multiple-choice (nominal)
Withdrawal Intentions		
Intentions-daily withdrawal	D – 7, 8	Likert-type (ordinal/interval)
Intention-drop	D – 9	Likert-type (ordinal/interval)

Notes. [a] Appendix and item number are shown.

Table 6

Pearson Correlations among Behavioral Withdrawal Criterion Variables

Behavioral Withdrawal Variables	1	2	3	4	5	6
1. Permanent withdrawal						
2. Lecture absenteeism	.39**					
3. Lab absenteeism	.41**	.57**				
4. Lecture lateness frequency	.24*	.26**	.25**			
5. Lab lateness frequency [a]	.08	.10	.27**	.33**		
6. Lecture lateness time lost	.04	.04	.10	.61**	.25**	
7. Lecture early departures	.22**	.43**	.13	.21**	.14	.06

Notes. [a] Correlations with the lab lateness criterion were based on $N = 134$. Sample sizes for the other variables ranged from 210 to 213.
* $p < .05$, ** $p < .01$

Table 7

Item Response Characteristics for the Full CRT-W

Variables	p-value [a]	SD	Skewness [b]	Distractor (%) [c]	Sample Size
1. M-1	.86	.34	-2.14	2.9	236
2. M-2	.52	.50	-0.09	2.1	237
3. M-3	.78	.41	-1.37	0.4	242
4. M-4	.62	.49	-0.51	14.8	207
5. M-5	.20	.20	1.54	7.8	224
6. M-6	.39	.39	0.45	0.8	241
7. M-7	.50	.50	-0.01	0.4	241
8. M-8	.15	.36	1.95	2.5	237
9. M-9	.83	.38	-1.76	6.2	228
10. RC-1	.18	.38	1.68	1.6	239
11. RC-2	.17	.37	1.81	0.4	242
12. RC-3	.40	.49	0.41	0.4	242
13. RC-4	.58	.49	-0.32	1.2	240
14. RC-5	.12	.33	2.35	0.8	241
15. RC-6	.42	.50	0.31	0.0	243
16. RC-7	.24	.43	1.21	1.6	239
17. RC-8	.76	.43	-1.20	2.1	238
18. RC-9	.46	.50	0.18	6.2	228
19. RC-10	.98	.14	-6.87	5.3	198
20. SI-1	.39	.49	0.44	7.0	226
21. SI-2	.99	.07	-15.26	0.4	233
22. SI-3	.27	.45	1.02	7.0	226
23. SI-4	.08	.27	3.21	2.9	236
24. SI-5	.36	.48	0.57	2.5	237
25. SI-6	.13	.33	2.24	0.4	242

Notes. [a]p-value indicates the base rate or proportion of individuals who endorsed the JM response option. [b]The standard error of skewness ranged from 0.156 to 0.173. [c]Distractor % indicates the percentage of respondents who endorsed an illogical distractor response option.

Items are numbered in the order displayed in Table 4. *RC* denotes items from the revocable commitment JM. *M* denotes items from the marginalization JM. *SI* denotes items from the social injustice JM.

Table 8

Corrected Item-Total Biserial Correlations for the Full CRT-W

Item	Marginalization of Withdrawal	Revocable Commitment	Social Injustice Bias	CRT-W Total
M-1	**.30****	.01	-.01	-.03
M-2	**.06**	-.12*	-.04	-.09
M-3	**.28****	.05	-.13*	.11
M-4	**.11**	.00	.06	.03
M-5	**-.14***	.03	.07	-.20
M-6	**.05**	.11*	.00	.08
M-7	**-.01**	.12*	-.14*	-.22
M-8	**.15***	-.11*	-.06	.06
M-9	**-.03**	-.12*	.16*	.05
RC-1	.01	**.17****	.06	.24**
RC-2	-.04	**-.10**	.18**	-.21**
RC-3	-.02	**.13***	-.09	.06
RC-4	-.14*	**.04**	-.13	-.04
RC-5	-.02	**.04**	.23**	-.10
RC-6	-.04	**.01**	-.01	-.05
RC-7	.03	**.02**	-.03	-.09
RC-8	.17*	**-.10**	-.07	-.05
RC-9	-.01	**.12***	.06	.03
SI-1	-.09	-.17	**.09**	.07
SI-3	.10	-.01	**.11**	.01
SI-4	.10	-.11*	**-.04**	-.19**
SI-5	-.03	.03	**.09**	-.02
SI-6	-.11	.15*	**.15***	.16

Notes. Item correlations with corresponding JM total scores are shown in bold. Items are numbered in the order displayed in Table 4. *M* denotes items from the marginalization JM. *RC* denotes items from the revocable commitment JM. *SI* denotes items from the social injustice JM. Significance tests are based on the confidence intervals for point-biserial correlations (see Cohen, Cohen, West, & Aiken, 2003). Sample sizes ranged from 198 to 243, as specified per item in Table 5.

* $p < .05$, ** $p < .01$

Table 9

Tetrachoric Item Correlations for the Full CRT-W

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. M-1																	
2. M-2	.19																
3. M-3	.25	.19															
4. M-4	.19	.04	.25														
5. M-5	.08	-.22	-.07	-.16													
6. M-6	.08	-.05	.06	.02	.14												
7. M-7	.13	.13	.12	.01	-.08	-.03											
8. M-8	.08	.00	.19	.17	-.04	-.04	-.14										
9. M-9	.12	.00	.08	.00	-.02	-.03	-.07	.31									
10. RC-1	.13	-.05	-.09	-.26	.24	.06	-.03	.06	.06								
11. RC-2	-.06	-.08	-.02	-.11	.01	.00	.06	.14	-.14	.06							
12. RC-3	-.11	-.18	.07	-.15	-.01	.15	.16	.07	-.04	.20	.00						
13. RC-4	-.10	-.19	-.05	.10	-.11	-.08	.16	-.31	-.18	-.08	-.21	.17					
14. RC-5	.21	.20	-.04	-.13	.12	-.17	-.08	-.04	-.24	.35	.23	-.09	-.15				
15. RC-6	-.04	-.13	.01	.00	.09	-.03	.03	-.12	.15	-.10	-.09	.11	.05	-.07			
16. RC-7	-.09	-.16	.07	.02	.00	.25	.01	-.04	.04	.07	-.08	.18	.08	-.30	.08		
17. RC-8	.12	.18	.26	.17	-.11	.08	.08	.04	-.06	.08	-.13	.05	.14	-.07	-.04	-.15	
18. RC-9	.01	.00	.00	.10	-.04	.05	.00	-.04	-.17	.07	.00	.14	.01	.32	.05	-.02	.04
19. SI-1	-.15	-.05	-.21	-.04	-.06	-.01	-.20	.08	.30	-.13	-.14	-.21	-.06	.06	.05	-.05	-.26
20. SI-3	.19	.00	.05	.01	.05	.10	-.10	.04	.13	-.06	.18	.10	-.05	.20	-.34	-.05	.09
21. SI-4	.24	.35	-.10	.04	-.13	-.16	.08	.03	-.02	-.02	.22	.00	-.17	.23	.11	-.15	-.30
22. SI-5	-.01	-.19	-.01	.16	.12	.03	-.05	-.24	.05	.21	.04	-.11	-.13	.13	.06	-.04	.02
23. SI-6	-.12	-.08	-.15	-.12	.21	.04	-.11	-.05	-.02	.12	.38	-.11	-.09	.27	.04	.22	.07

Table 9 continued…

Variables	18	19	20	21	22
1. M-1					
2. M-2					
3. M-3					
4. M-4					
5. M-5					
6. M-6					
7. M-7					
8. M-8					
9. M-9					
10. RC-1					
11. RC-2					
12. RC-3					
13. RC-4					
14. RC-5					
15. RC-6					
16. RC-7					
17. RC-8					
18. RC-9					
19. SI-1	.01				
20. SI-3	.15	.10			
21. SI-4	-.25	.14	-.27		
22. SI-5	.04	.06	.06	.10	
23. SI-6	.09	.05	.37	-.22	.18

Notes. Items are numbered in the order displayed in Table 4. *M* denotes items from the marginalization JM. *RC* denotes items from the revocable commitment JM. *SI* denotes items from the social injustice JM.

Table 10

Item-Criterion Biserial Correlation Coefficients for the Full CRT-W (*N* = 213) *with Criterion Variables*

CRT-W Item	Permanent Withdrawal	Lecture Absenteeism	Lab Absenteeism	Lecture Lateness	Lab Lateness [a]	Early Departure
M-1	.04	.08	.17*	.08	.02	. 07
M-2	.06	-.02	.09	-.10	.07	-.02
M-3	.04	-.01	.05	-.16*	-.14	-.10
M-4	-.12*	-.09	.06	-.01	-.04	.05
M-5	.01	-.03	-.04	.00	.11	.14*
M-6	-.14*	.03	-.13*	-.04	-.01	-.07
M-7	-.21**	-.10	-.06	-.03	.03	-.13*
M-8	.09	.11	.04	.10	-.05	.05
M-9	-.02	-.04	.08	.02	-.05	-.08
RC-1	.12*	.12*	.02	.10	.09	.09
RC-2	.16*	.12*	.15*	.04	.21*	.10
RC-3	.00	.04	.01	.16*	.20*	-.11
RC-4	-.04	.13*	.12*	.06	.10	.09
RC-5	.18**	-.10	-.06	-.06	.09	-.07
RC-6	.04	-.04	.17**	.00	.09	-.05
RC-7	-.06	-.04	-.08	.02	-.05	.01
RC-8	-.12*	.05	.04	-.04	.00	.03
RC-9	.06	-.04	-.04	-.06	.27*	.17*
SI-1	-.09	-.21**	-.09	-.02	.05	-.15*
SI-3	.02	-.07	.00	.10	.09	.07
SI-4	.18**	.05	.19**	.04	-.05	-.03
SI-5	-.02	.01	-.06	.05	.06	.08
SI-6	.20**	-.02	.03	.16*	.42*	.08

Notes. [a]Item correlations with the lab lateness criterion were based on *N* = 134. Significance tests are based on the confidence intervals for point-biserial correlations (see Cohen, Cohen, West, & Aiken, 2003).
Items are numbered in the order displayed in Table 4. *M* denotes items from the marginalization JM. *RC* denotes items from the revocable commitment JM. *SI* denotes items from the social injustice JM.
* *p* < .05, ** *p* < .01

Table 11

Summary of Scale Revision Considerations

Item	Item response characteristics	Item-total correlation	Item-criterion relationships	Decision
M-1	--	Positive	Positive correlations with *lab absenteeism*	Retained
M-2	--	--	--	Dropped
M-3	--	Positive	Negative correlation with *lecture lateness*	Dropped
M-4	High% of distractor responses	--	Negative correlation with *permanent withdrawal*	Dropped
M-5	--	Negative	Positive correlation with *early departure*	Retained
M-6	--	--	Negative correlation with *permanent withdrawal* and *lab absenteeism*	Dropped
M-7	--	--	Negative correlation with *permanent withdrawal* and *early departure*	Dropped
M-8	--	Positive	Trending positive correlations with *lecture absenteeism* and *lecture lateness*	Retained
M-9	--	--	--	Dropped
RC-1	--	Positive	Positive correlation with *permanent withdrawal* and *lecture absenteeism*	Retained
RC-2	--	--	Positive correlations with *permanent withdrawal*, *lecture absenteeism*, *lab absenteeism*, and *lab lateness*	Retained
RC-3	--	Positive	Positive correlations with *lecture lateness* and *lab lateness*	Retained
RC-4	--	--	Positive correlations with *lecture absenteeism* and *lab absenteeism*	Retained
RC-5	--	--	Positive correlation with *permanent withdrawal*	Retained
RC-6	--	--	Positive correlation with *lab absenteeism*	Retained
RC-7	--	--	--	Dropped
RC-8	--	--	Negative correlation with *permanent withdrawal*	Dropped
RC-9	--	Positive	Positive correlation with *lab lateness* and *early departure*	Retained
RC-10	Low item variance	--	--	Dropped
SI-1	--	--	Negative correlations with *lecture absenteeism* and *early departures*	Dropped
SI-2	Low item variance	--	--	Dropped
SI-3	--	--	Trending positive correlations with *lecture lateness* and *lab lateness*	Retained
SI-4	--	--	Positive correlation with *permanent withdrawal*, *lab absenteeism*, and *lecture lateness*	Retained
SI-5	--	--	--	Dropped
SI-6	--	Positive	Positive correlation with *permanent withdrawal*, *lecture lateness*, and *lab lateness*	Retained

Notes. Positive evidence favoring retention of the item is bolded. Items are numbered in the order displayed in Table 4. *M* denotes items from the marginalization JM. *RC* denotes items from the revocable commitment JM. *SI* denotes items from the social injustice JM. Evidence that was equivocal in reference to the item's retention is denoted '--'.

Table 12

Pearson (r) and Spearman (rho) Correlations for Hypothesis Tests and Exploratory Analyses

	Absenteeism Behaviors				Lateness Behaviors			
Variable	Lecture Absenteeism		Lab Absenteeism		Lecture Lateness		Lab Lateness[a]	
	r	*rho*	*r*	*rho*	*r*	*rho*	*r*	*rho*
CRT-W								
Marginalization JM	-.03	-.06	.02	-.07	.01	.00	.00	-.01
Revocable Commitment JM	.06	.04	.15*	.14*	.06	.04	.12	.10
Social Injustice JM	.01	.01	.00	.09	.10	.07	.12	.15*
Overall CRT-W Score	.04	-.01	.12*	.10	.09	.08	.13	.15*
Personality Traits								
Conscientiousness	-.22**	-.20**	-.16*	-.12*	-.09	-.09	-.02	-.06
Emotional Stability	-.28**	-.27**	-.16**	-.16**	-.11	-.01	-.03	-.04
Prior Withdrawal Behaviors								
Job tenure (years)	-.09	-.03	-.03	-.04	-.02	-.07	-.08	-.13
Number of jobs held	.18*	.20**	.17**	.16**	.10	.07	.02	.00
Universities attended	-.09	-.13*	.02	.01	-.06	-.09	.09	-.01
Courses dropped	.36**	.31**	.29**	.25**	.10	.14*	.16*	.13
Plans changed	.11	.10	.16**	.19**	.03	-.01	.03	-.04
Attendance problems	.12*	.20**	.21**	.25**	.14*	.17**	.29**	.16*
Social Embeddedness								
Friends and relatives	.09	.07	.05	.01	.09	.04	-.08	-.04
Exploratory Analyses								
Agreeableness	-.08	-.08	-.06	-.05	-.10	-.02	.07	.13
Extraversion	.10	.07	.04	.04	.05	.03	-.04	-.01
Openness to Experience	.09	.04	.22**	.17**	.03	.04	.08	.05
Commute distance (miles)	-.06	-.05	.04	-.02	.03	.02	-.09	-.13
Commute difficulty	-.07	-.05	.11	.12	.00	.03	.15	.11
Commute method-	-.05	-.04	-.14*	-.08	-.03	-.07	-.06	-.03
Commute method-weather	-.05	-.05	-.06	-.08	-.07	-.07	-.13	-.17
Intentions daily withdrawal	.36**	.32**	.32**	.29**	.26**	.22**	.14	.13
Intention drop	.17**	.13	-.01	-.02	.05	.03	-.01	.00

Notes. [a]Correlations with lab lateness were based on $N = 134$.
Sample sizes vary from 210 to 213, as specified in text.
* $p < .05$, ** $p < .01$

Table 13

Pearson Correlations among Study Variables (*N* = 213)

Variables	1	2	3	4	5	6	7	8	9	10
1. Permanent withdrawal										
2. Lecture absenteeism	.39**									
3. Lab absenteeism	.41**	.57**								
4. Lecture lateness	.24**	.26**	.25**							
5. Lab lateness [a]	.08	.10	.27**	.33**						
6. Lecture early departures	.22**	.43**	.13*	.22**	.14					
7. CRT-W total (initial)	-.07	-.03	.11	.04	.05	.01				
8. CRT-W marginalization (initial)	-.14*	-.10	.02	-.06	-.04	-.03	.63**			
9. CRT-W revocable commitment (initial)	.04	.08	.18**	.05	.03	.06	.61**	-.03		
10. CRT-W social injustice (initial)	.01	-.04	-.04	.08	.13	-.01	.41**	-.01	-.07	
11. CRT-W total (revised)	.11	.10	.17*	.12*	.12	.11	.67**	.08	.73**	.30**
12. CRT-W marginalization (revised)	.05	-.03	.01	.04	-.02	.09	.26**	.50**	-.09	-.01
13. CRT-W revocable commitment (revised)	.10	.06	.17*	.07	.07	.05	.51**	-.11	.89**	-.01
14. CRT-W social injustice (revised)	.09	.01	.02	.11*	.16*	.06	.34**	.05	-.03	.72**
15. Agreeableness	-.09	-.08	-.06	-.11	.00	-.02	-.01	.02	-.04	.00
16. Conscientiousness	.04	-.22**	-.13*	-.12*	-.03	-.11	.03	.08	-.01	-.03
17. Emotional stability	-.04	-.28**	-.17**	-.11*	.00	-.15*	.05	.05	.01	.06
18. Extraversion	.01	.10	.06	.10	.02	.16*	-.03	.02	-.10	.03
19. Openness to experience	-.03	.09	.23**	.01	.10	-.06	.13*	.13*	.07	-.01
20. Job tenure (years)	-.04	-.09	-.03	-.01	.03	-.16*	-.02	.00	-.04	.01
21. Number of jobs held	.05	.18*	.21**	.11	.13	.19**	.00	-.07	.06	.01
22. Universities attended	-.07	-.09	.05	-.07	.06	-.11	-.05	.00	-.02	-.07
23. Courses dropped	.04	.36**	.29**	.08	.15*	.11	-.11	-.06	-.08	-.04
24. Plans changed	.05	.11	.18**	.02	.04	-.07	-.03	-.08	-.07	.12*
25. Friends and relatives	.00	.09	.07	.10	-.03	.02	-.05	-.02	-.07	-.01
26. Attendance problems	.15*	.12*	.24**	.19**	.29**	-.08	.06	.05	.04	.01
27. Commute distance (miles)	-.09	-.06	.01	.02	.03	-.18**	.06	.00	.12*	-.04
28. Commute difficulty	-.06	-.07	.07	-.03	.24**	-.07	.15*	.06	.09	.11
29. Commute method-independence	.01	-.05	-.11*	-.01	-.10	.00	-.13*	-.09	-.14*	.04
30. Commute method-weather	-.12*	-.05	-.10	-.05	-.05	-.16*	.06	.04	.04	.03
31. Intentions daily withdrawal	.20**	.36**	.34**	.32**	.28**	.26**	.11	.09	.05	.04
32. Intention drop	.23**	.17**	.02	.03	.01	.22**	-.10	-.10	-.09	.08

Table 13 continued…

Variables	11	12	13	14	15	16	17	18	19	20	21
1. Permanent withdrawal											
2. Lecture absenteeism											
3. Lab absenteeism											
4. Lecture lateness											
5. Lab lateness [a]											
6. Lecture early departures											
7. CRT-W total (initial)											
8. CRT-W marginalization (initial)											
9. CRT-W revocable commitment (initial)											
10. CRT-W social injustice (initial)											
11. CRT-W total (revised)											
12. CRT-W marginalization (revised)	.38**										
13. CRT-W revocable commitment (revised)	.83**	-.09									
14. CRT-W social injustice (revised)	.45**	.08	.02								
15. Agreeableness	-.02	.05	-.03	-.04							
16. Conscientiousness	-.03	.03	-.02	-.01	.15*						
17. Emotional stability	.04	.00	.05	.00	.10	.22**					
18. Extraversion	-.03	.11	-.09	-.01	.24**	-.04	.03				
19. Openness to experience	.07	.08	.03	-.02	.16*	.07	.00	.21**			
20. Job tenure (years)	.00	.08	-.02	.02	.07	.11	.01	-.03	-.08		
21. Number of jobs held	.00	-.16**	.06	.00	-.03	-.07	-.03	.02	.03	.09	
22. Universities attended	-.10	-.02	-.05	-.01	-.01	.18**	.02	-.20**	-.02	.19**	.08
23. Courses dropped	-.06	-.06	-.07	-.03	.11	-.12	-.15*	-.02	.01	.09	.15*
24. Plans changed	.04	-.10	-.05	.18**	-.05	-.11	.00	.04	.10	.17*	.15*
25. Friends and relatives	-.04	-.03	-.04	.02	.10	.01	.04	.26**	.13*	-.05	.01
26. Attendance problems	.04	.03	.04	.03	-.11	-.08	-.08	.05	.05	.11	.05
27. Commute distance (miles)	.07	.02	.11	-.10	.06	-.02	.07	-.04	.08	.27**	-.02
28. Commute difficulty	.08	-.08	.08	.05	.04	-.09	-.03	-.06	-.07	.04	.03
29. Commute method-independence	-.10	-.02	-.13*	.01	-.05	.02	-.12	.15*	-.08	-.01	.05
30. Commute method-weather	.04	.06	.02	-.03	.02	-.01	-.01	.03	.00	.20**	.06
31. Intentions daily withdrawal	-.01	.16**	.01	.07	-.03	-.24**	-.13*	.07	.09	-.02	.09
32. Intention drop	-.07	.06	-.05	.09	-.08	-.06	-.09	-.07	-.18**	.01	-.05

Table 13 continued…

Variables	22	23	24	25	26	27	28	29	30	31
1. Permanent withdrawal										
2. Lecture absenteeism										
3. Lab absenteeism										
4. Lecture lateness										
5. Lab lateness [a]										
6. Lecture early departures										
7. CRT-W total (initial)										
8. CRT-W marginalization (initial)										
9. CRT-W revocable commitment (initial)										
10. CRT-W social injustice (initial)										
11. CRT-W total (revised)										
12. CRT-W marginalization (revised)										
13. CRT-W revocable commitment (revised)										
14. CRT-W social injustice (revised)										
15. Agreeableness										
16. Conscientiousness										
17. Emotional stability										
18. Extraversion										
19. Openness to experience										
20. Job tenure (years)										
21. Number of jobs held										
22. Universities attended										
23. Courses dropped	.31**									
24. Plans changed	.09	.32**								
25. Friends and relatives	-.18**	-.12	.05							
26. Attendance problems	-.09	.06	.15*	.00						
27. Commute distance (miles)	.03	.02	-.06	-.12	-.07					
28. Commute difficulty	.03	.06	.05	-.07	.00	.39**				
29. Commute method-independence	.08	-.02	-.07	-.04	-.10	-.18**	-.20**			
30. Commute method-weather	.12	.03	-.01	-.10	-.05	.56**	.24**	-.06		
31. Intentions daily withdrawal	-.05	.14*	-.04	.12	.19**	-.07	.10	.09	-.12	
32. Intention drop	.01	.06	-.08	-.05	-.09	-.11	.01	.13	-.11	.28*

Notes. [a] Item correlations with the lab lateness criterion were based on *N* = 134.
* *p* < .05, ***p* < .01

Table 14

Results of Regression Analyses for Lecture Lateness

	Model 1		Model 2		Model 3	
Predictors	b	*95% C. I.*	b	*95% C. I.*	b	*95% C. I.*
Attendance problems	.002	-.01, .01	.002	-.01, .01	.001	-.01, .01
Courses dropped	.003	.000, .01	.003	.000, .01	.002	.000, .01
Intentions daily withdrawal	.04**	.02, .05	.04**	.02, .05	.03**	.01, .05
Model 1 R^2	.12**					
F (3, 205)	9.45**					
Marginalization			-.003	-.05, .05	-.002	-.05, .05
Revocable commitment			.03	-.03, .10	.02	-.04, .09
Social Injustice			.04	-.02, .09	.03	-.03, .08
Model 2 R^2			.13**			
Δ F (3, 202)			0.95			
Cumulative point total					.000**	.000, .000 [a]
Model 3 R^2					.16**	
Δ F (1, 201)					7.23**	

Notes. [a] The standardized regression coefficient for cumulative point total was -.19 (p = .008).
Un-standardized regression coefficients and corresponding confidence intervals are shown.
* p < .05, ** p < .01

Table 15

Results of Regression Analyses for Lab Lateness

	Model 1		Model 2		Model 3	
Predictors	b	*95% C. I.*	b	*95% C. I.*	b	*95% C. I.*
Attendance problems	.01	-.002, .04	.01	-.002, .02	.01**	-.002, .02
Courses dropped	.01**	.003, .01	.01**	.003, .01	.01	.002, .01
Model 1 R^2	.10**					
F (2, 130)	7.33**					
Marginalization			-.03	-.09, .04	-.03	-.09, .04
Revocable commitment			.02	-.07, .10	.01	-.07, .10
Social Injustice			.08*	.001, .15	.07	-.01, .15
Model 2 R^2			.13**			
Δ F (3, 127)			1.46			
Cumulative point total					.000	.000, .000[a]
Model 3 R^2					.14**	
Δ F (1, 126)					1.20	

Notes. [a] The standardized regression coefficient for cumulative point total was -.09 (p = .28).
Un-standardized regression coefficients and corresponding confidence intervals are shown.
* $p < .05$, ** $p < .01$

Table 16

Results of Regression Analyses for Lecture Absenteeism

	Model 1		Model 2		Model 3	
Predictors	b	*95% C. I.*	b	*95% C. I.*	b	*95% C. I.*
Conscientiousness	-.02	-.04, .01	-.02	-.04, .01	-.02*	-.04, .00
Emotional stability	-.03*	-.04, -.01	-.03**	-.05, -.01	-.02*	-.04, -.004
Number of jobs	.01	.00, .03	.01	-.002, .03	.01	-.01, .02
Courses dropped	.03*	.01, .04	.03**	.01, .04	.02**	.01, .03
Attendance problems	.00	-.003, .01	.00	-.003, .01	-.001	-.01, .00
Intentions daily withdrawal	.06**	.03, .09	.06**	.03, .09	.04**	.01, .07
Intention drop	.02	-.01, .05	.02	-.01, .05	-.02	-.04, .01
Model 1 R^2	.28**					
F (7, 201)	11.41**					
Marginalization			-.03	-.11, .06	-.02	-.09, .05
Revocable commitment			.07	-.04, .18	.03	-.07, .12
Social Injustice			-.01	-.10, .09	-.05	-.13, .03
Model 2 R^2			.29**			
Δ F (3, 198)			0.66			
Cumulative point total					.000**	.000, .000[a]
Model 3 R^2					.53**	
Δ F (1, 197)					97.18**	

Notes. [a] The standardized regression coefficient for cumulative point total was -.54 (p < .001).
Un-standardized regression coefficients and corresponding confidence intervals are shown.
* p < .05, ** p < .01

Table 17

Results of Regression Analyses for Lab Absenteeism

	Model 1		Model 2		Model 3	
Predictors	b	*95% C. I.*	b	*95% C. I.*	b	*95% C. I.*
Conscientiousness	-.004	-.02, .02	-.002	-.02, .02	-.01	-.02, .01
Emotional stability	-.01	-.03, .01	-.01	-.03, .003	-.01	-.02, .01
Openness to experience	.03**	.01, .06	.03**	.01, .05	.04**	.02, .06
Number of jobs	.01*	.001, .02	.01	.000, .02	.01	-.002, .02
Courses dropped	.02*	.01, .03	.02**	.01, .03	.01**	.004, .02
Plans changed	.004	-.004, .01	.004	-.003, .01	.004	-.002, .01
Attendance problems	.01*	.001, .01	.01*	.001, .01	.003	.000, .01
Intentions daily withdrawal	.05**	.02, .07	.05**	.02, .07	.03*	.01, .05
Model 1 R^2	.27**					
F (8, 200)	9.34**					
Marginalization			.01	-.07, .08	.01	-.06, .07
Revocable commitment			.13**	.04, .22	.11*	.02, .19
Social Injustice			-.01	-.09, .07	-.04	-.11, .03
Model 2 R^2			.30**			
Δ F (3, 197)			2.60[a]			
Cumulative point total					.000**	.000, .000[b]
Model 3 R^2					.30**	
Δ F (1, 196)					55.33**	

Notes. [a] Δ F = 2.598 (*p* = .054). [b]The standardized regression coefficient for cumulative point total was -.42 (*p* < .001).
Un-standardized regression coefficients and corresponding confidence intervals are shown.
* *p* < .05, ** *p* < .01

Table 18

Results of Regression Analyses for Early Departures

	Model 1		Model 2		Model 3	
Predictors	b	*95% C. I.*	b	*95% C. I.*	b	*95% C. I.*
Emotional stability	-.01	-.03, .01	-.01	-.03, .01	-.01	-.02, .01
Extraversion	.02*	.003, .03	.02*	.002, .03	.02*	.001, .03
Number of jobs	.02**	.01, .03	.02**	.01, .03	.02**	.01, .03
Job tenure	-.01**	-.02, -.003	-.01**	-.02, -.003	-.01**	-.02, -.003
Courses dropped	.01	-.004, .02	.01	-.003, .02	.01	-.01, .02
Intentions daily withdrawal	.03*	.01, .06	.03*	.001, .06	.02	-.004, .05
Intention drop	.04**	.01, .06	.04**	.01, .06	.03	-.002, .05
Model 1 R^2	.20**					
F (7, 194)	6.70**					
Marginalization			.06	-.02, .14	.06	-.02, .13
Revocable commitment			.07	-.03, .17	.06	-.05, .16
Social Injustice			.03	-.05, .11	.01	-.07, .09
Model 2 R^2			.21**			
Δ F (3, 191)			1.47			
Cumulative point total					.000**	.000, .000[a]
Model 3 R^2					.26**	
Δ F (1, 190)					10.80**	

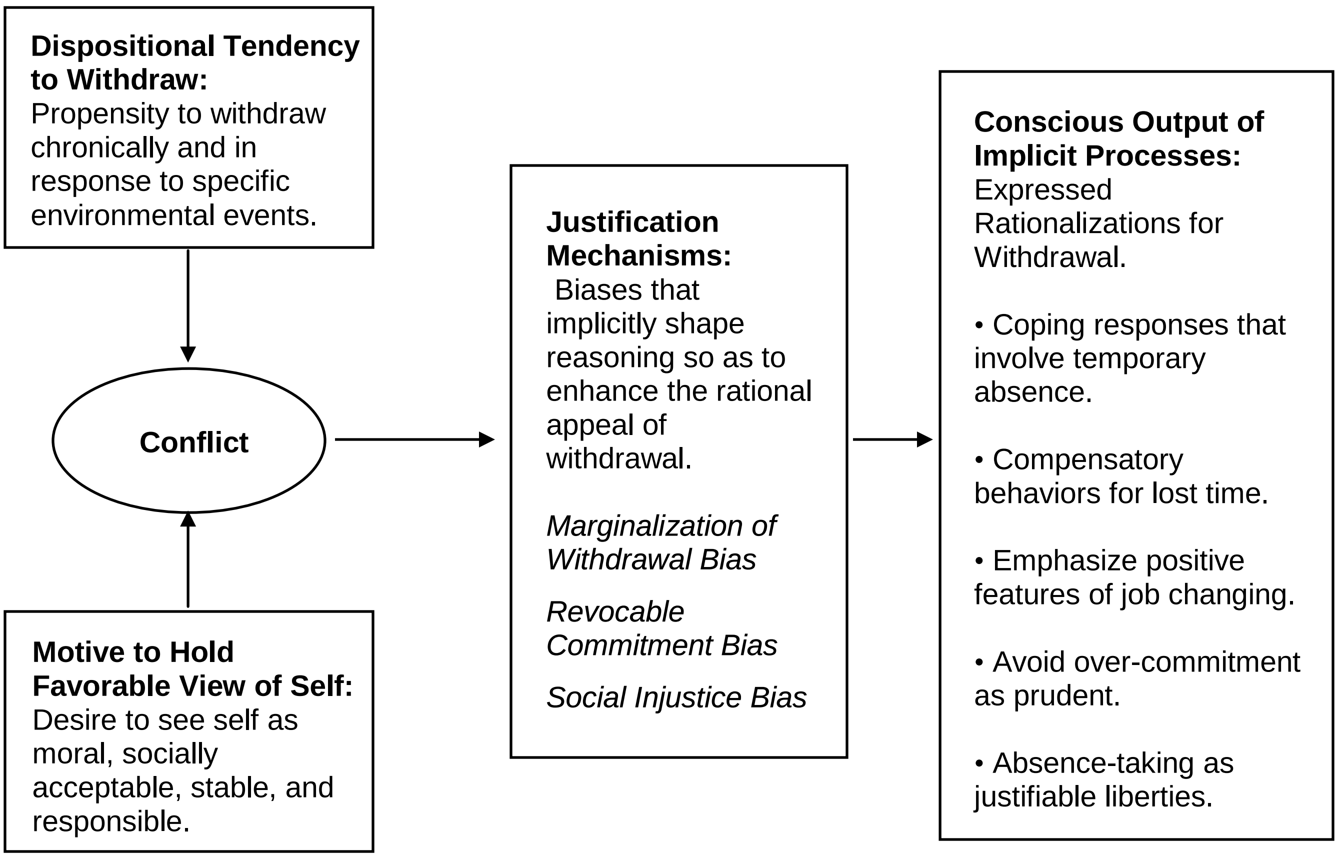
Notes. [a] The standardized regression coefficient for cumulative point total was -.23 (p = .001).
Un-standardized regression coefficients and corresponding confidence intervals are shown.
* $p < .05$, ** $p < .01$

Table 19

Summary of Hypothesis Testing Results

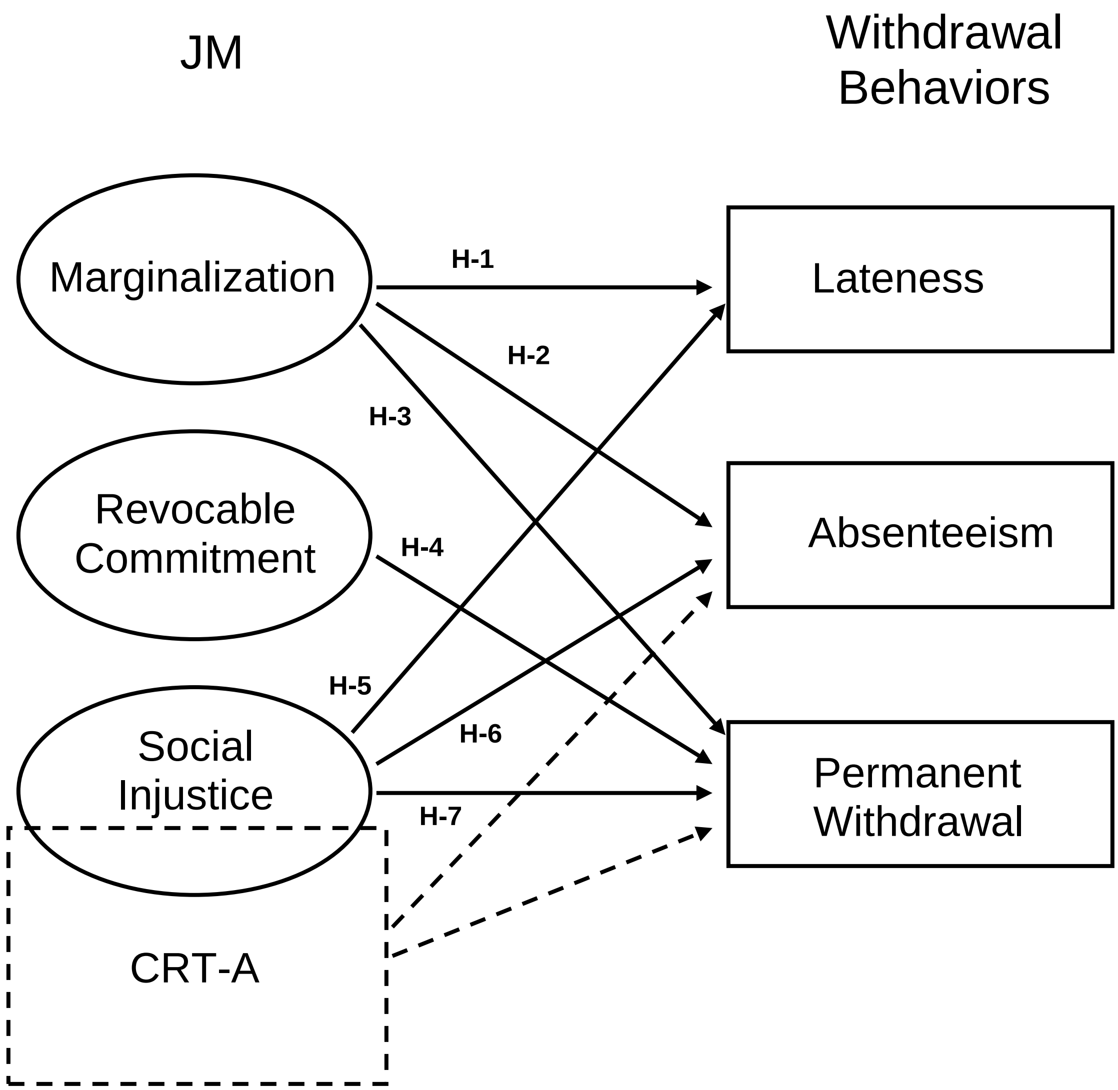
Hypothesis	Results
H1: Individuals with a stronger marginalization of withdrawal bias exhibit a higher frequency of lateness behaviors.	Not supported
H2: Individuals with a stronger marginalization of withdrawal bias exhibit a higher frequency of absenteeism behaviors.	Not supported
H3: Individuals with a stronger marginalization of withdrawal bias have a higher likelihood of permanent withdraw.	Not supported
H4: Individuals with a stronger revocable commitment bias have a higher likelihood of permanent withdraw.	Partial support
H5: Individuals with a stronger social injustice bias exhibit a higher frequency of lateness behaviors.	Partial support
H6: Individuals with a stronger social injustice bias exhibit a higher frequency of absenteeism behaviors.	Not supported
H7: Individuals with a stronger social injustice bias have a higher likelihood of permanent withdraw.	Not supported
H8: Individuals with a stronger dispositional tendency to withdraw exhibit a higher frequency of lateness behaviors.	Partial support
H9: Individuals with a stronger dispositional tendency to withdraw exhibit a higher frequency of absenteeism behaviors.	Partial support
H10: Individuals with a stronger dispositional tendency to withdraw have a higher likelihood of permanent withdraw.	Partial support
H11: Individuals with higher levels of conscientiousness exhibit a lower frequency of lateness behaviors.	Not supported
H12: Individuals with higher levels of conscientiousness exhibit a lower frequency of absenteeism behaviors.	Supported
H13: Individuals with higher levels of conscientiousness have a lower likelihood of permanent withdrawal.	Not supported
H14: Individuals with higher levels of emotional stability exhibit a lower frequency of lateness behaviors.	Not supported
H15: Individuals with higher levels of emotional stability exhibit a lower frequency of absenteeism behaviors.	Supported
H16: Individuals with higher levels of emotional stability have a lower likelihood of permanent withdrawal.	Not supported
H17: Individuals who have higher levels of prior withdrawal exhibit a higher frequency of lateness behaviors.	Partial support
H18: Individuals who have higher levels of prior withdrawal exhibit a lower frequency of absenteeism behaviors.	Supported
H19: Individuals who have higher levels of prior withdrawal have a higher likelihood of permanent withdrawal.	Partial support
H20: Individuals who are more socially embedded exhibit a lower frequency of lateness behaviors.	Not supported
H21: Individuals who are more socially embedded exhibit a lower frequency of absenteeism behaviors.	Not supported
H22: Individuals who are more socially embedded have a lower likelihood of permanent withdrawal.	Not supported
H23: A dispositional tendency to withdraw explains variance in behavioral withdrawal criteria above and beyond explicit measures of personality and biodata measures of prior withdrawal behaviors and social embeddedness.	Partial support

Figure 1. Model for rationalization process underlying withdrawal behaviors.



Note: Figure adapted from James et al. (2005).

Figure 2. Model relating withdrawal justification mechanisms (JMs) to types of withdrawal behaviors.



Notes: Hypotheses are labeled H-1, H-2, etc. All hypothesized relationships are positive, indicating that individuals who hold marginalization, revocable commitment, and social injustice biases are more likely to exhibit withdrawal behaviors. The predictive relationships for the CRT-A (shown with dashed arrows) will not be examined in the present study but are illustrated to point out the conceptual overlap with social injustice bias.

Figure 3. Histograms for lecture absenteeism before and after logarithmic transformation (N = 213).

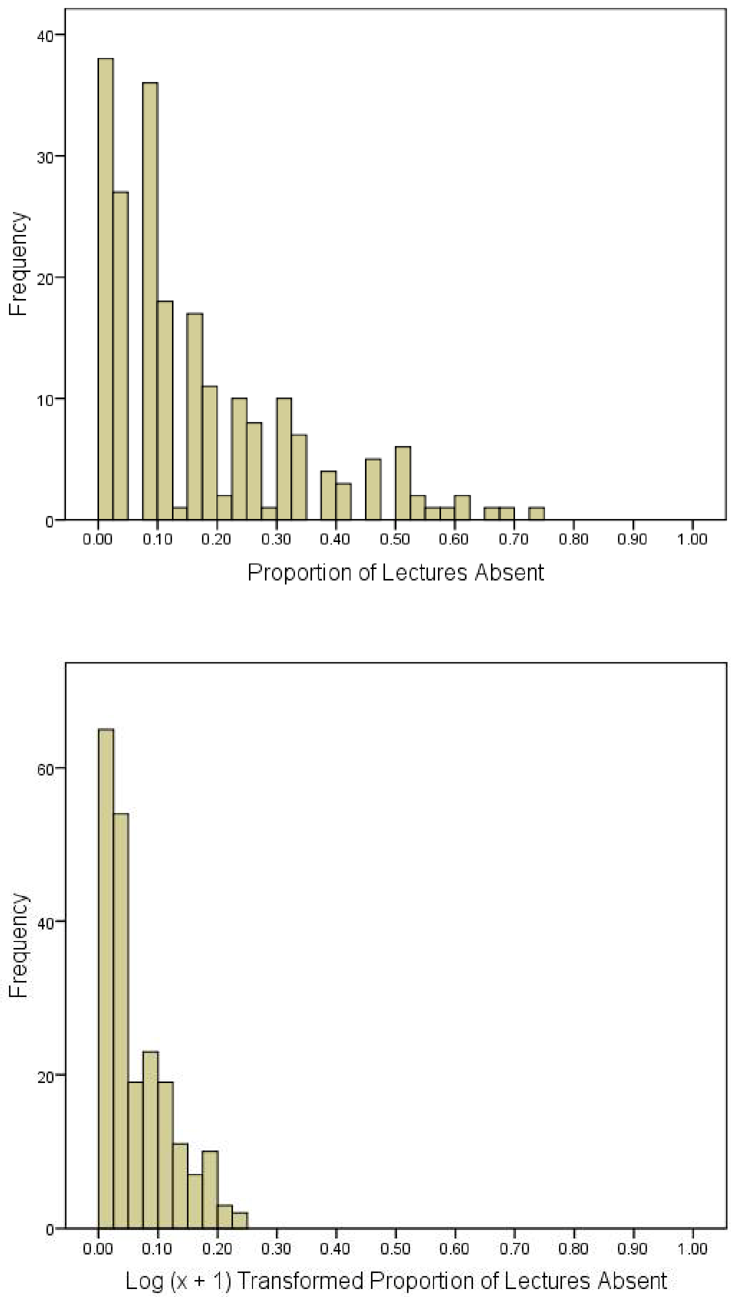


Figure 4. Histograms for lab absenteeism before and after logarithmic transformation (*N* = 213).

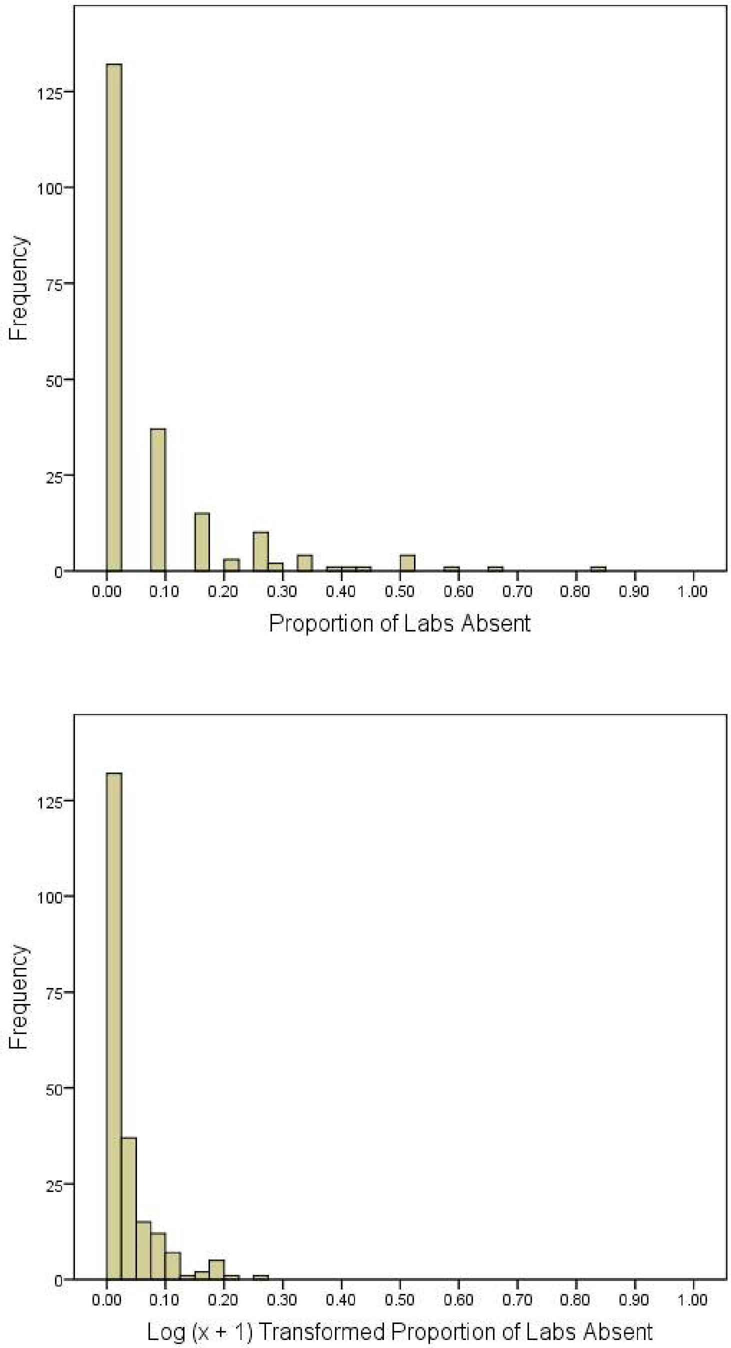


Figure 5. Histograms for lecture lateness (frequency) before and after logarithmic transformation (*N* = 212).

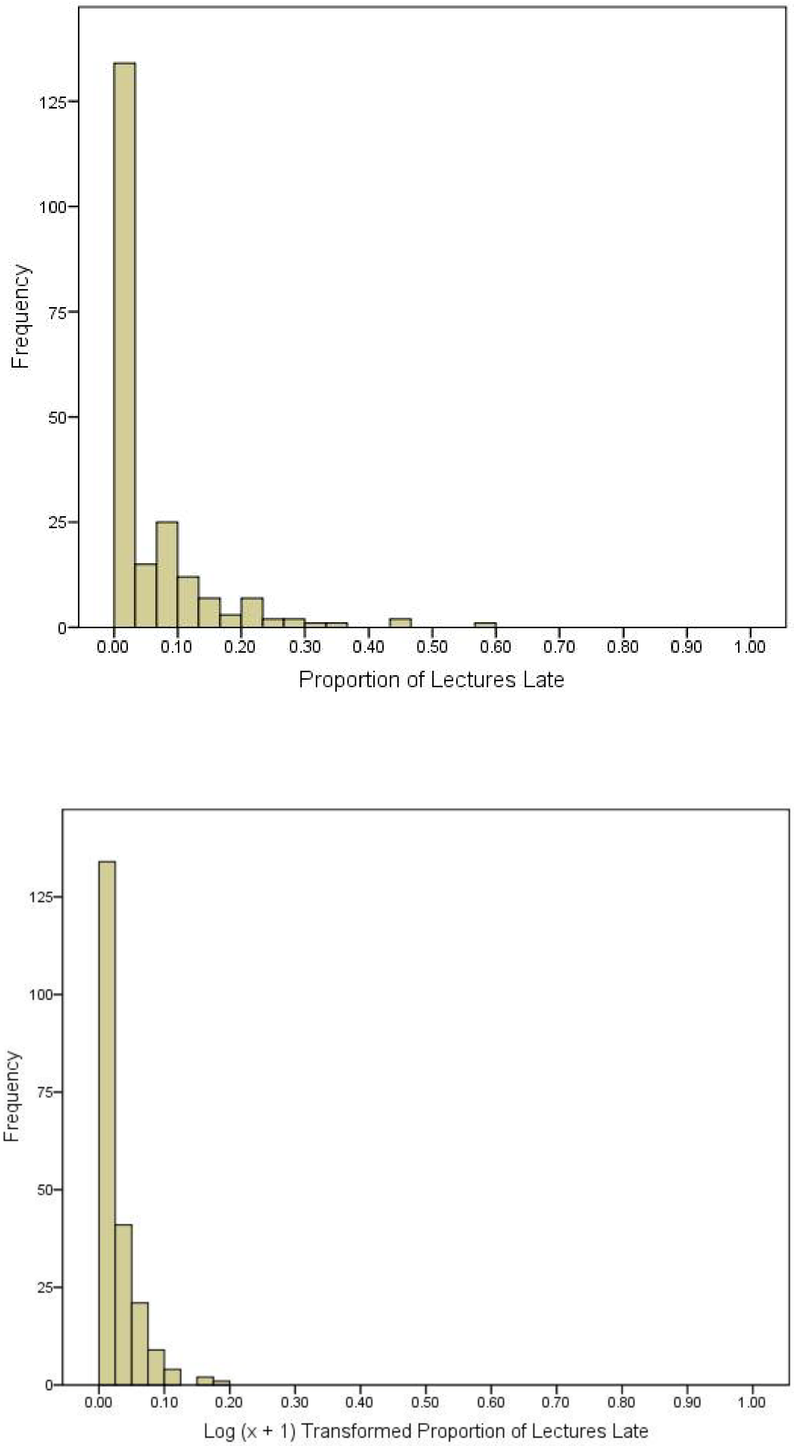
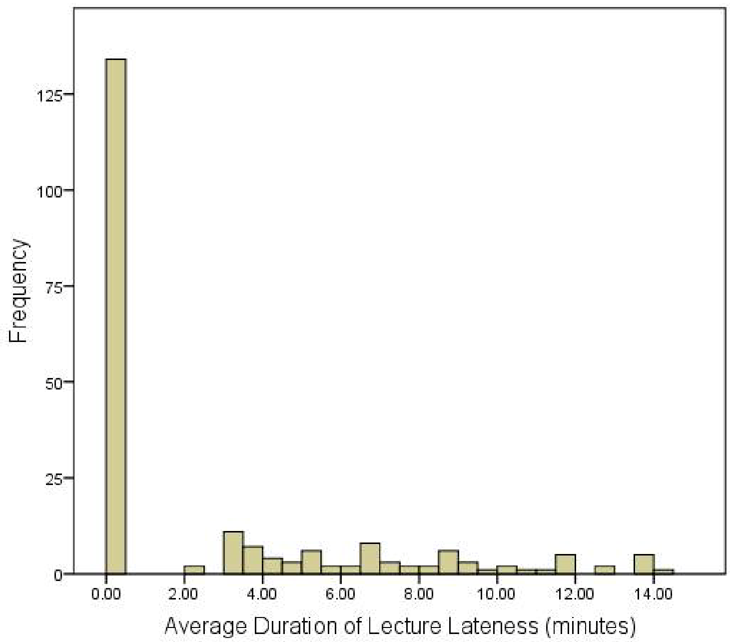


Figure 6. Histograms for lecture lateness (time lost) before and after logarithmic transformation (N = 213).



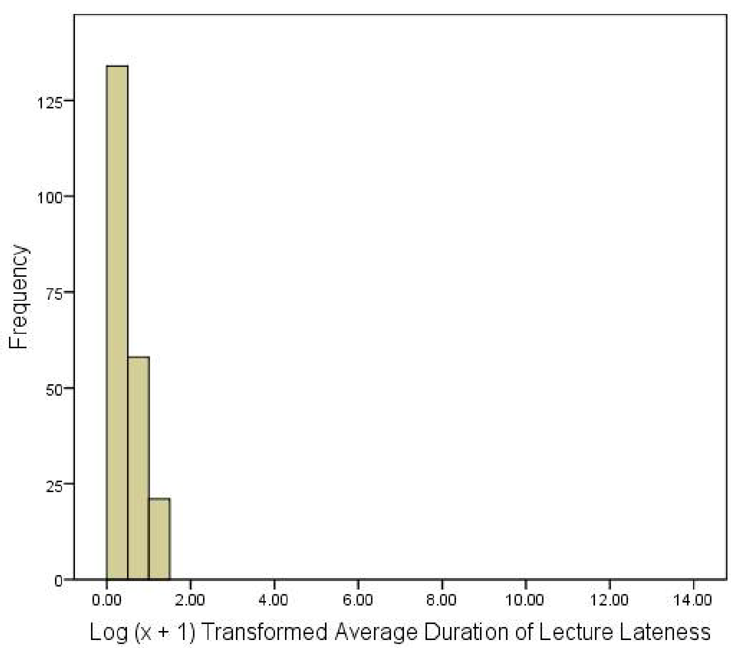


Figure 7. Histograms for lab lateness before and after logarithmic transformation (N = 134).

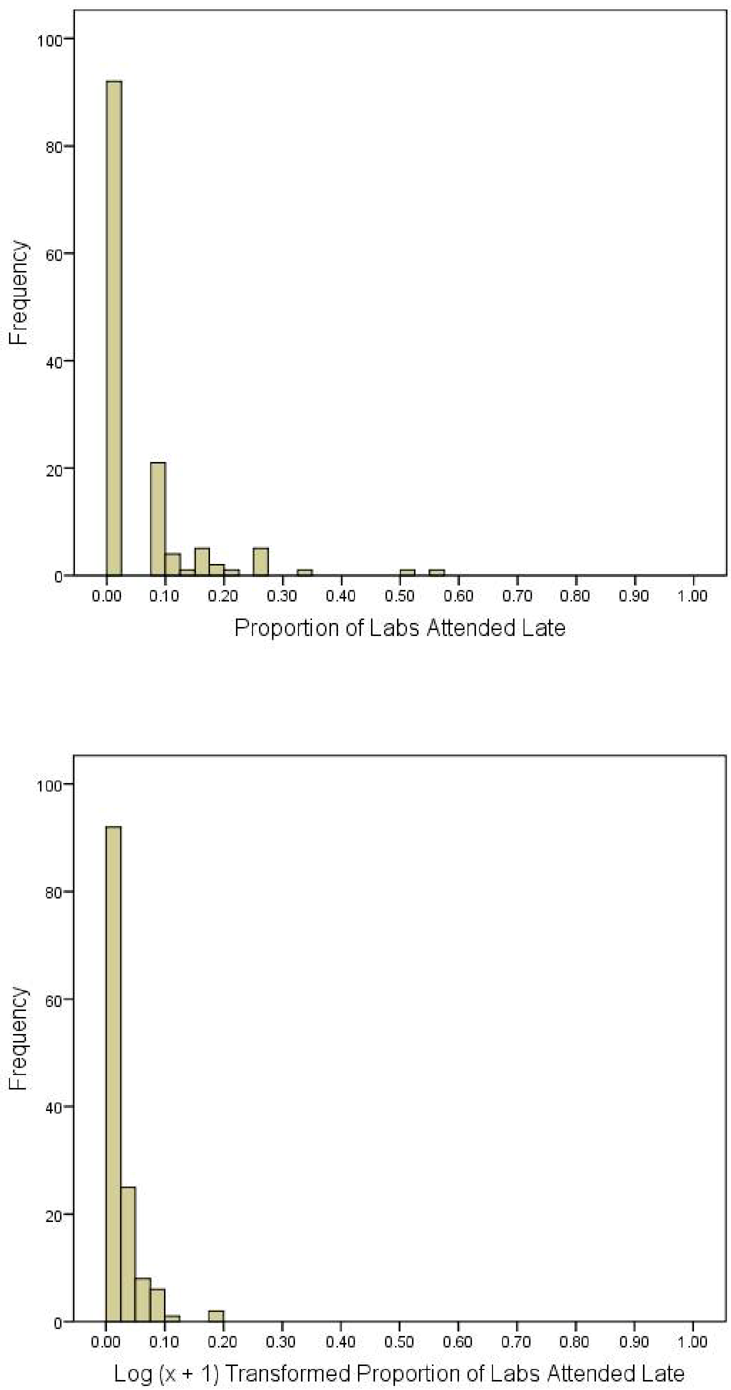
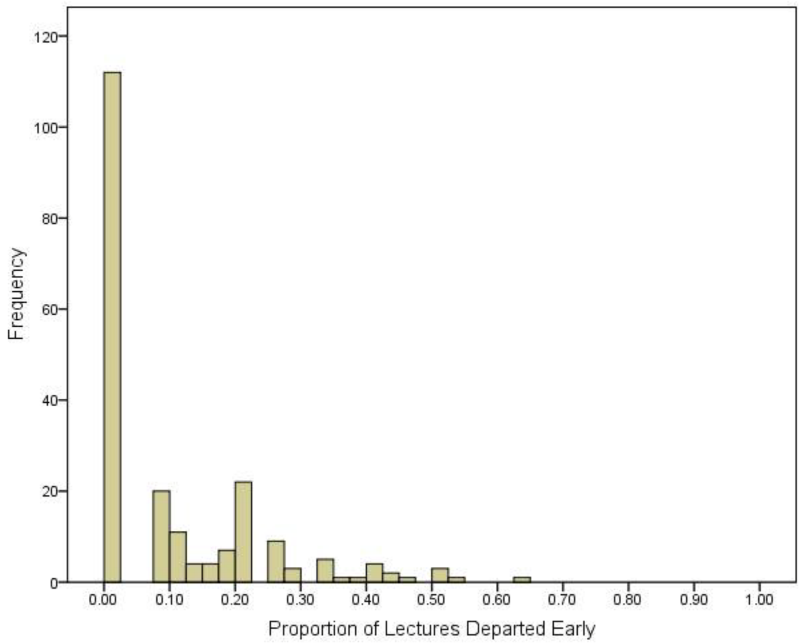


Figure 8. Histograms for early departures from lecture before and after logarithmic transformation (N = 211).



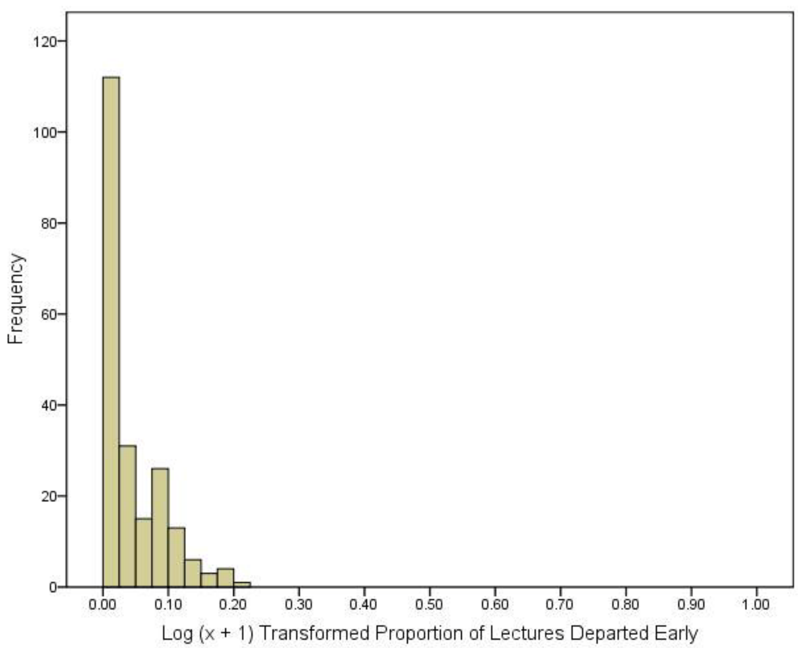


Figure 9. Histogram for the marginalization of withdrawal JM (N = 213).

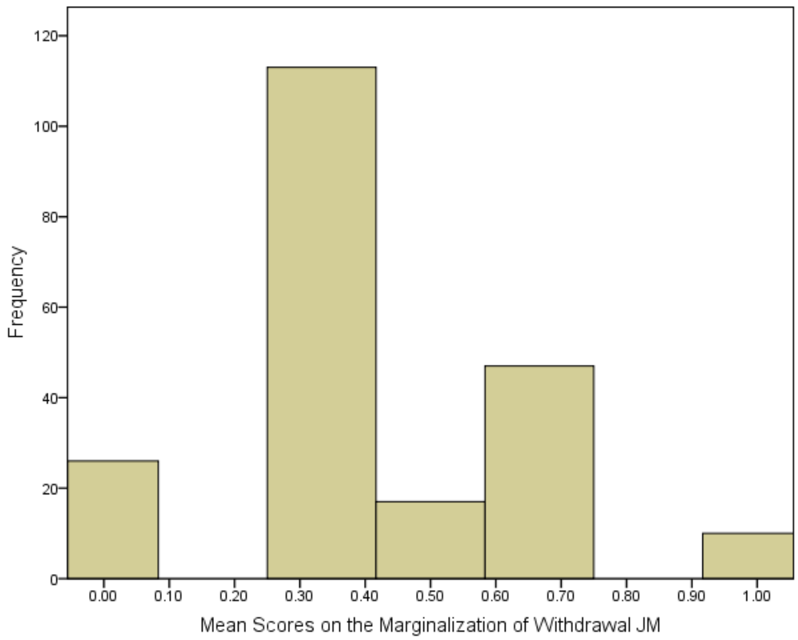


Figure 10. Histogram for the revocable commitment JM (N = 213).

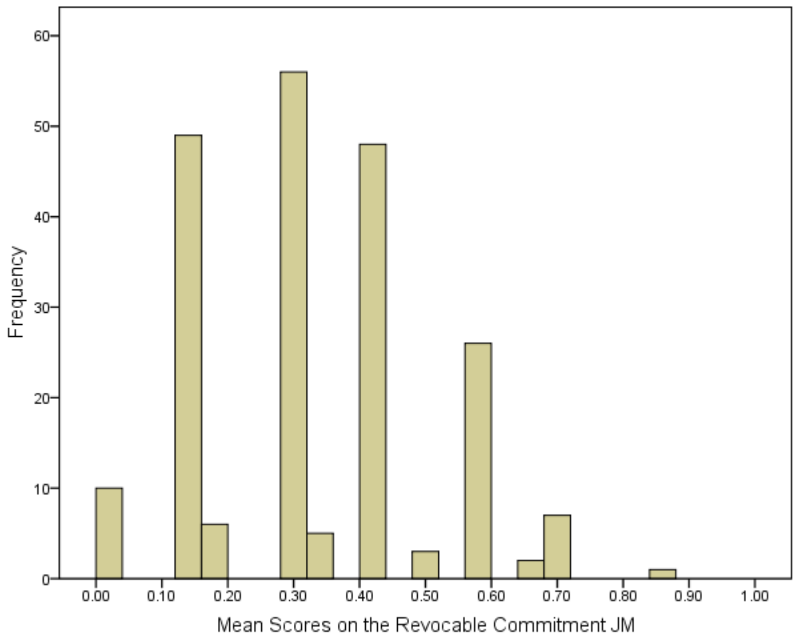


Figure 11. Histogram for the social injustice JM (*N* = 213).

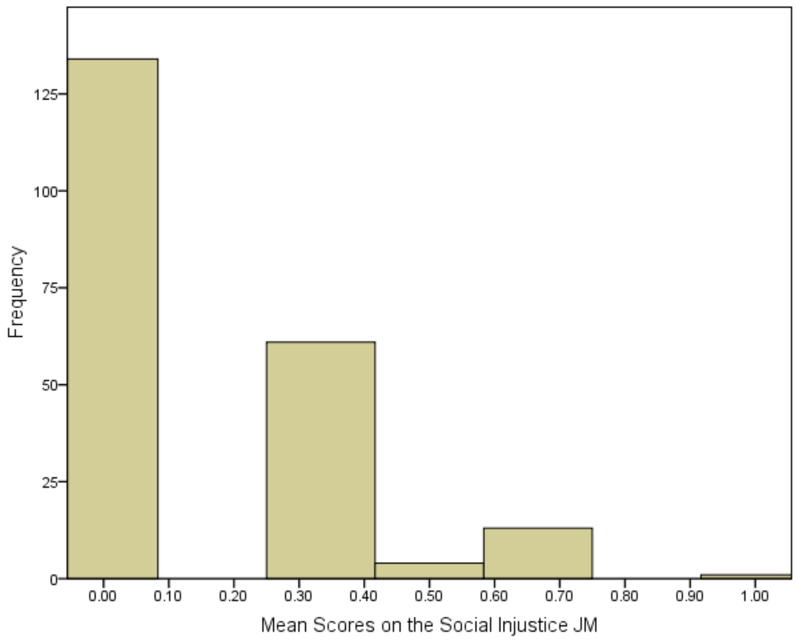


Figure 12. Histogram for CRT-W total scores (*N* = 213).

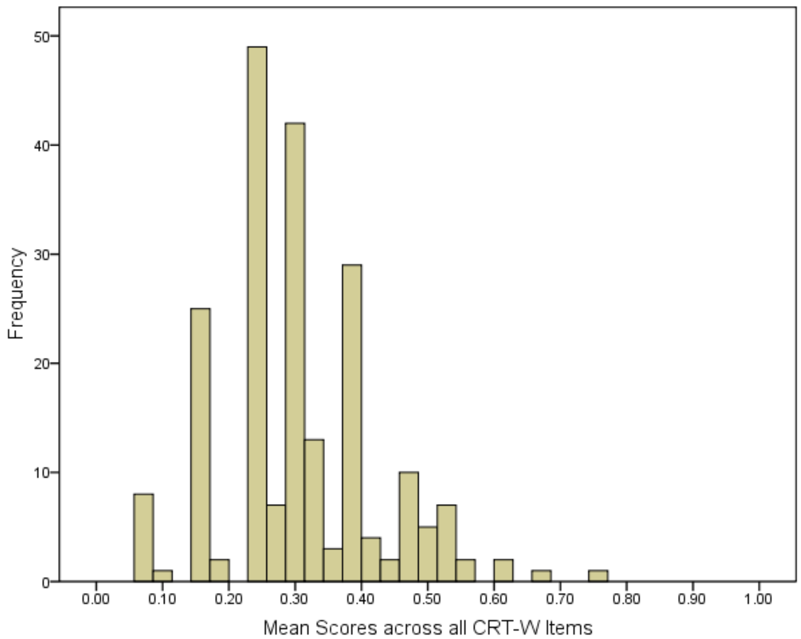


Figure 13. Bivariate scatterplots for marginalization of withdrawal with absenteeism and lateness variables.

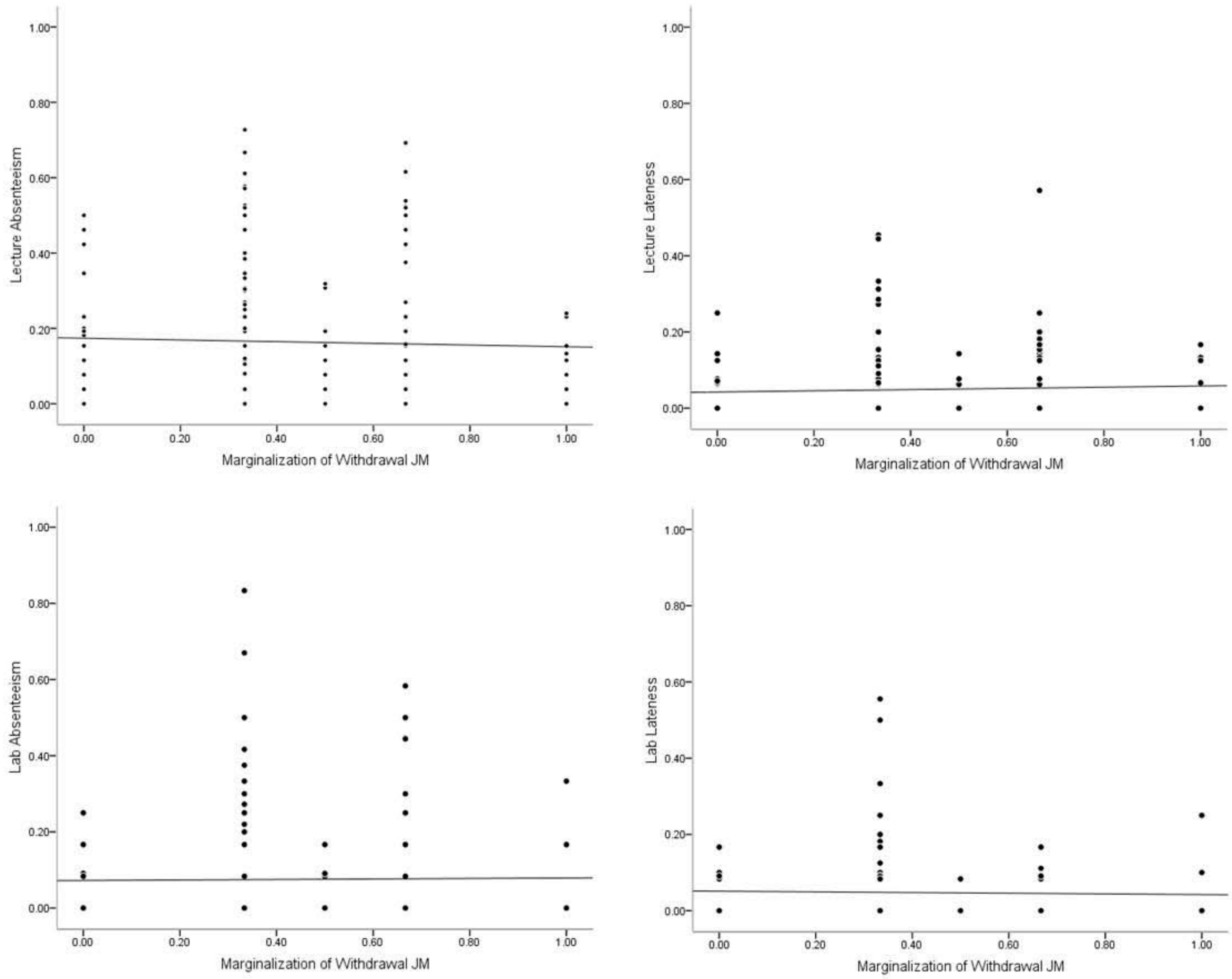


Figure 14. Residuals plots for marginalization of withdrawal with absenteeism and lateness variables.

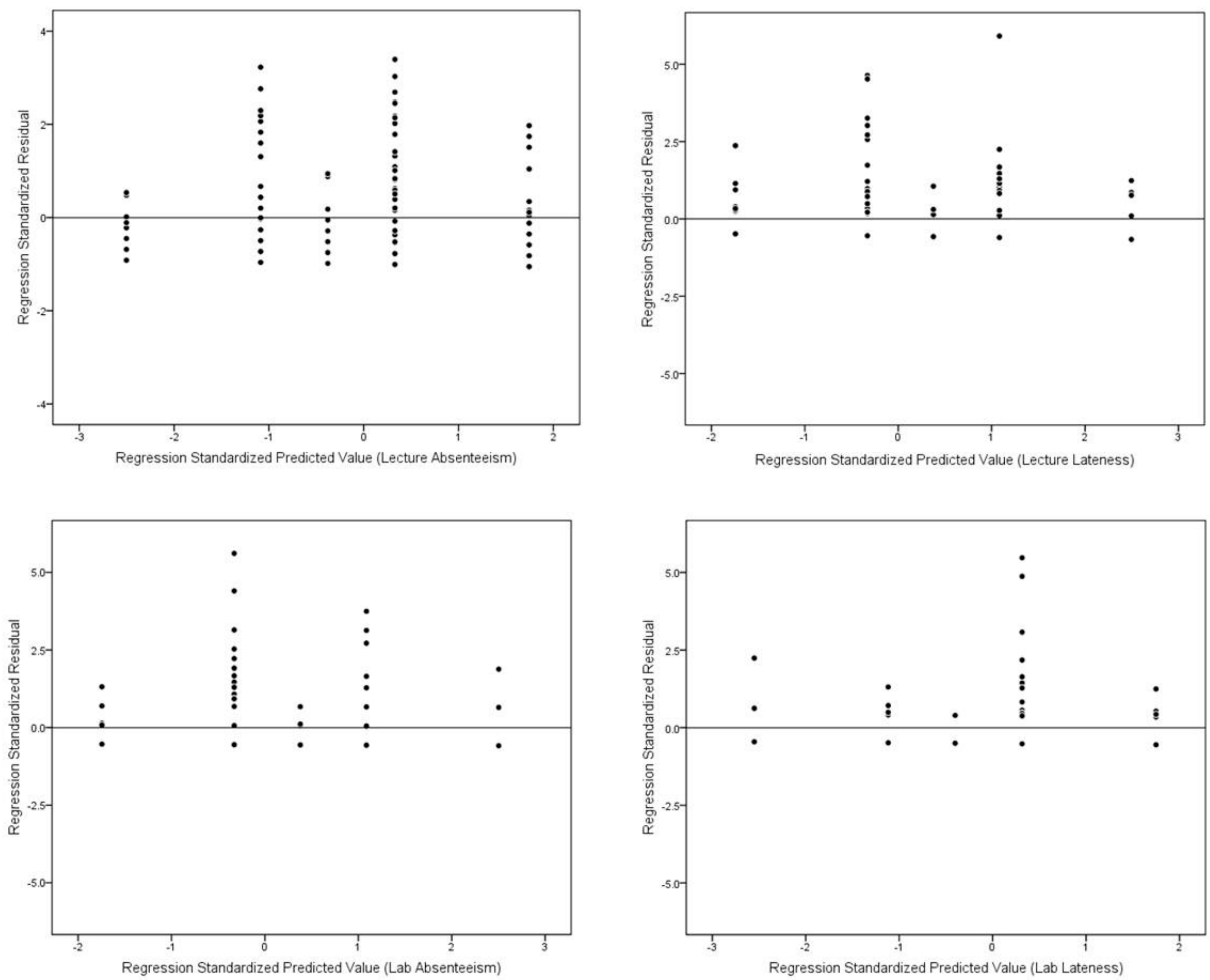


Figure 15. Cumulative hazard (top) and survival functions (bottom) for time-to-permanent withdrawal.

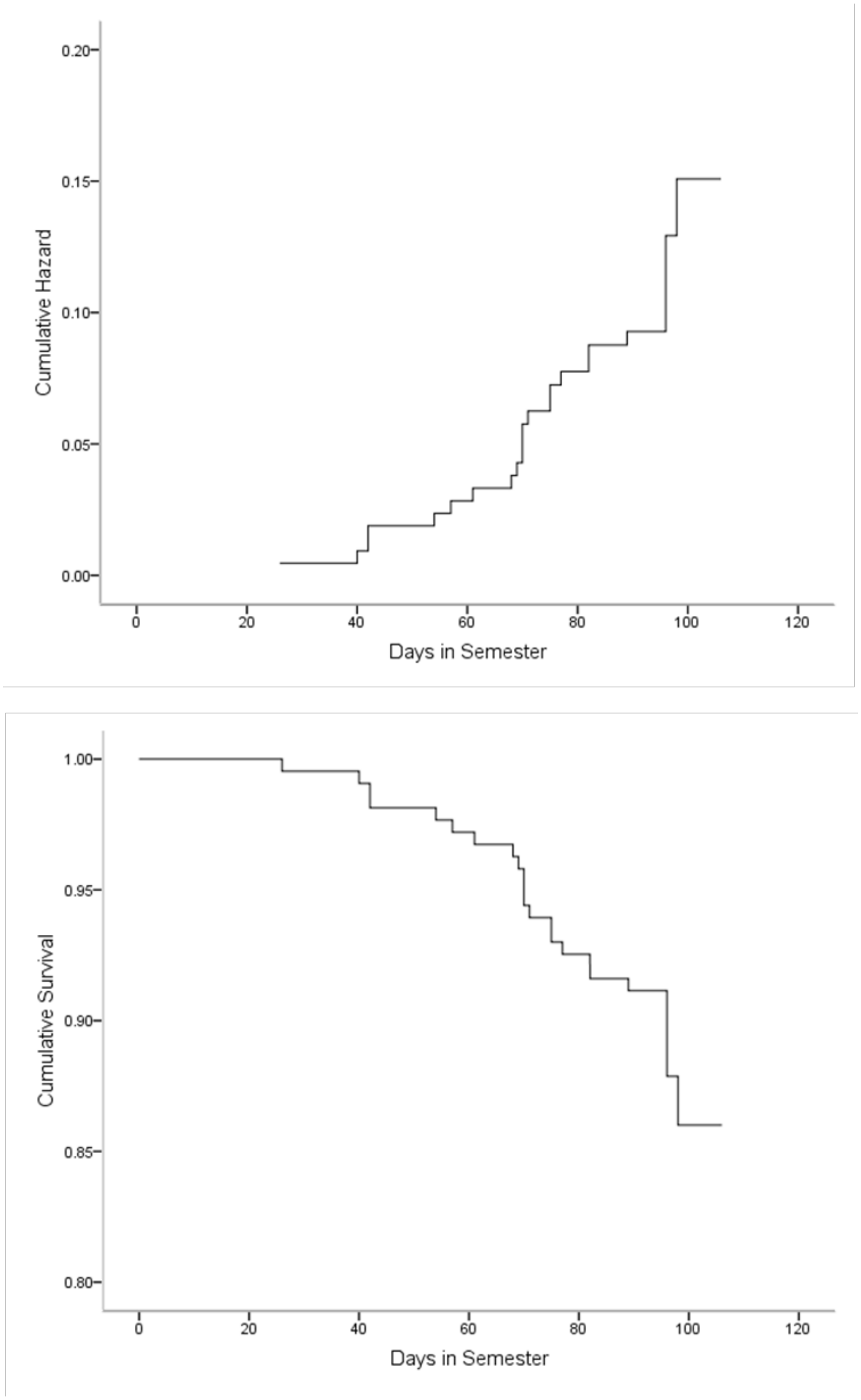
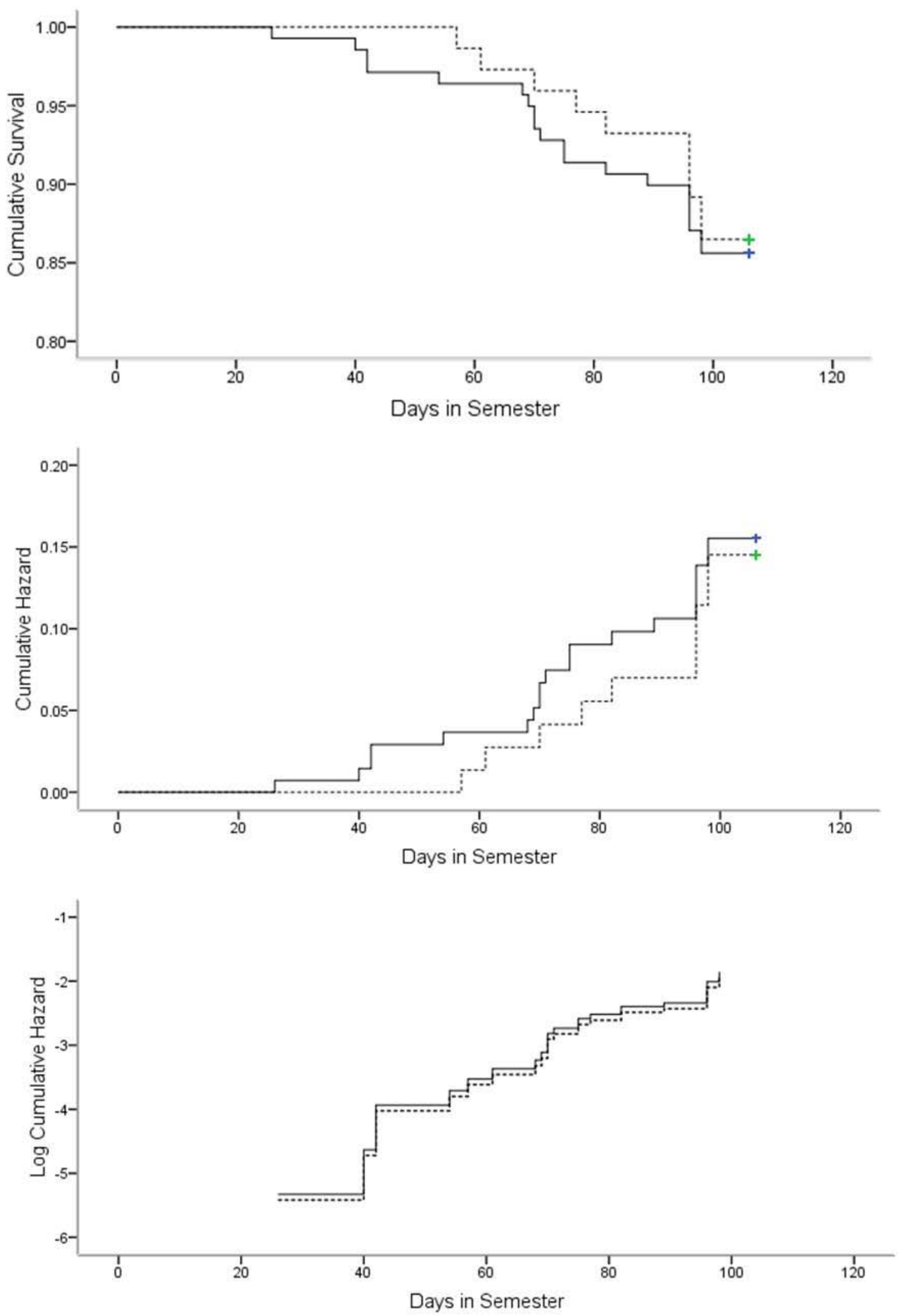
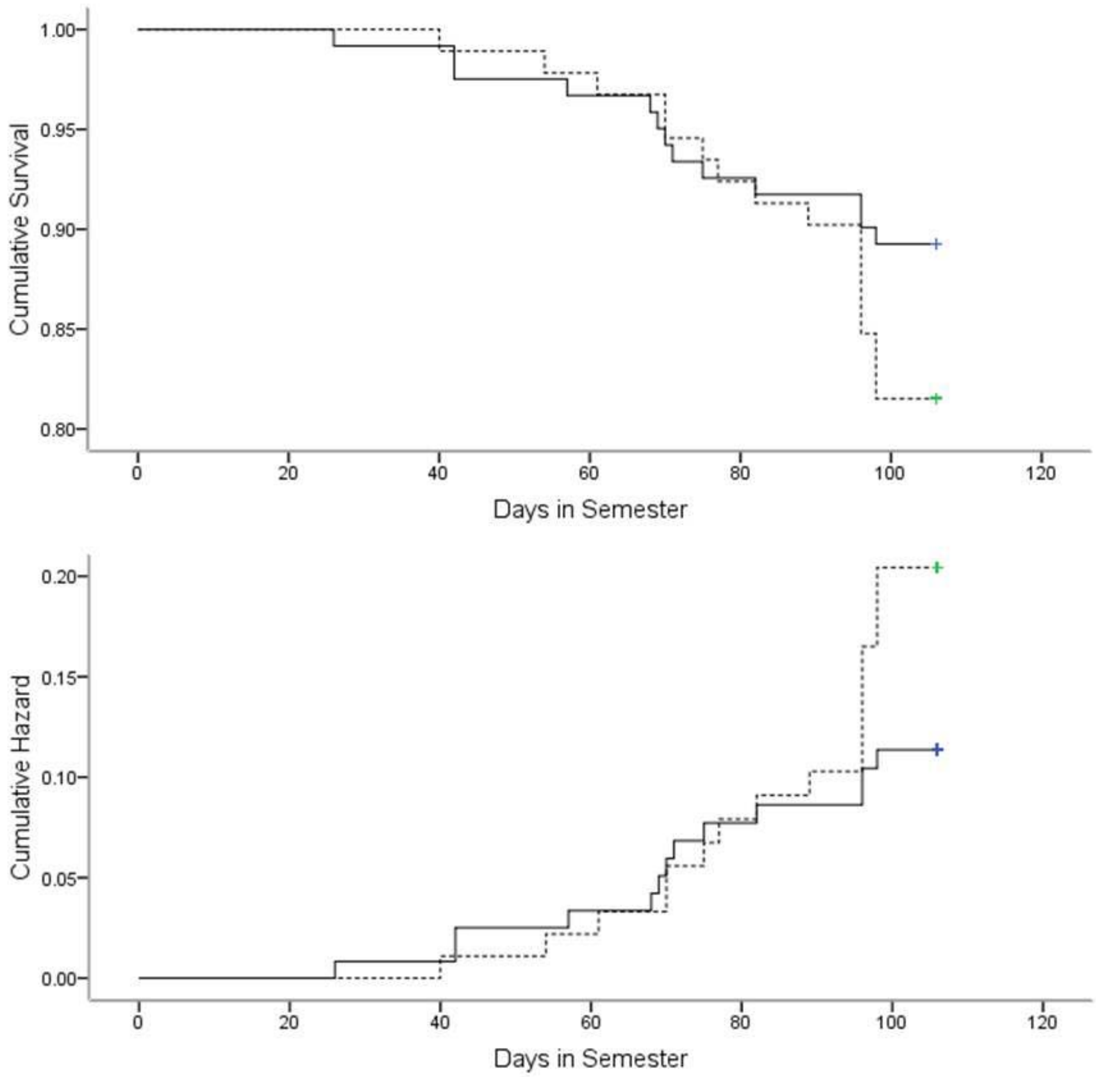


Figure 16. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low marginalization of withdrawal.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 17. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low revocable commitment.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 18. Bivariate scatterplots for revocable commitment with absenteeism and lateness variables.

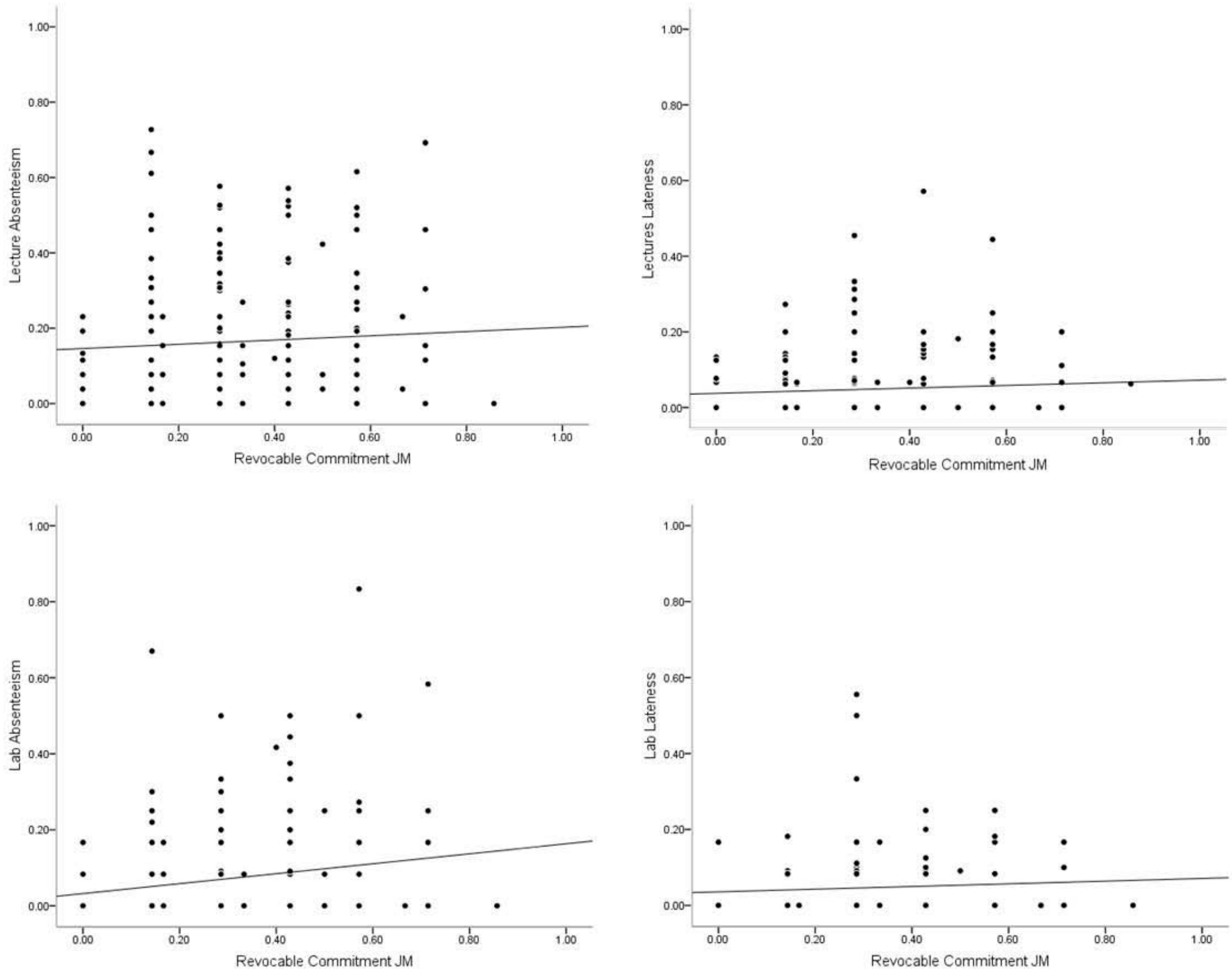


Figure 19. Residuals plots for revocable commitment with absenteeism and lateness variables.

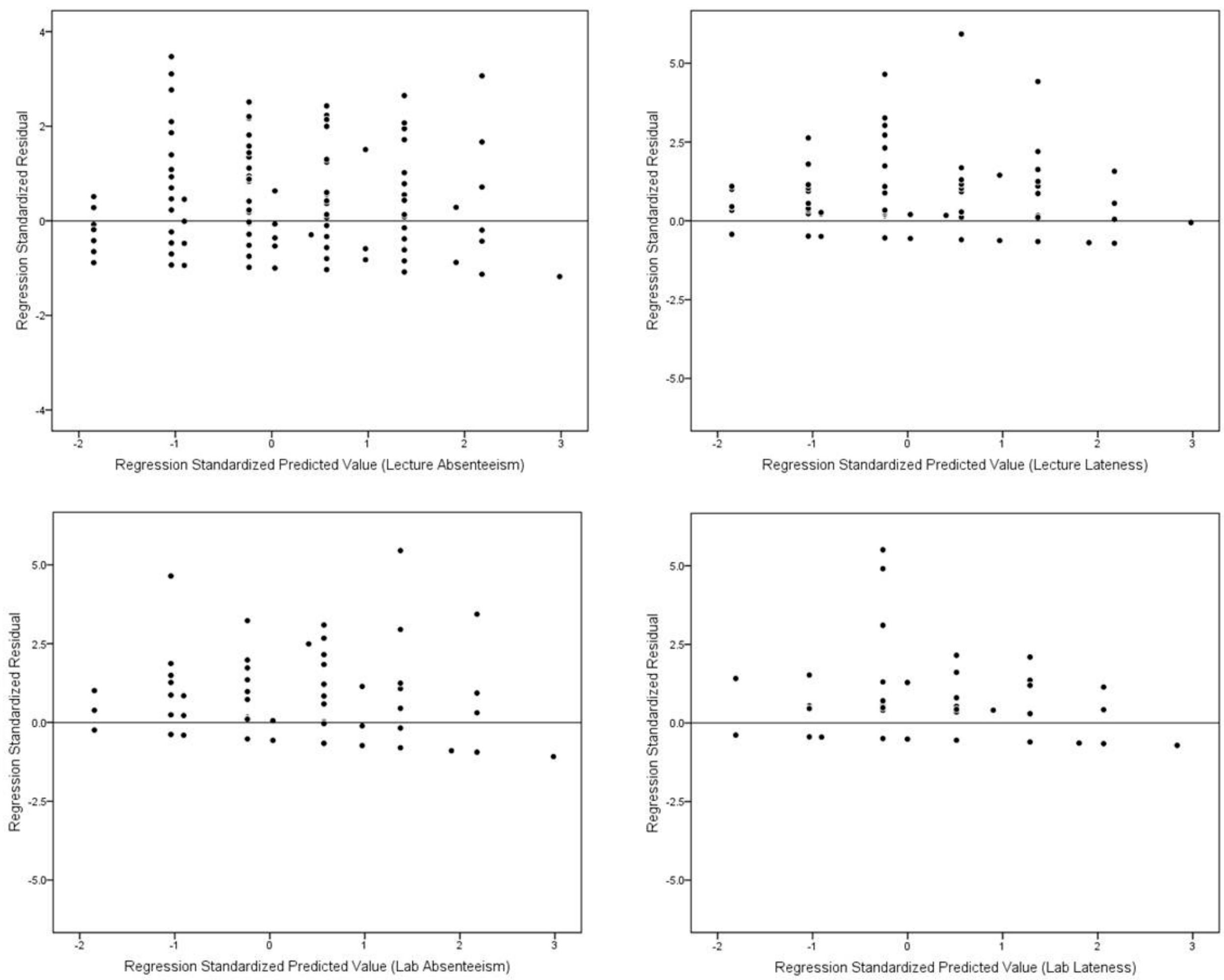


Figure 20. Bivariate scatterplots for social injustice with absenteeism and lateness variables.

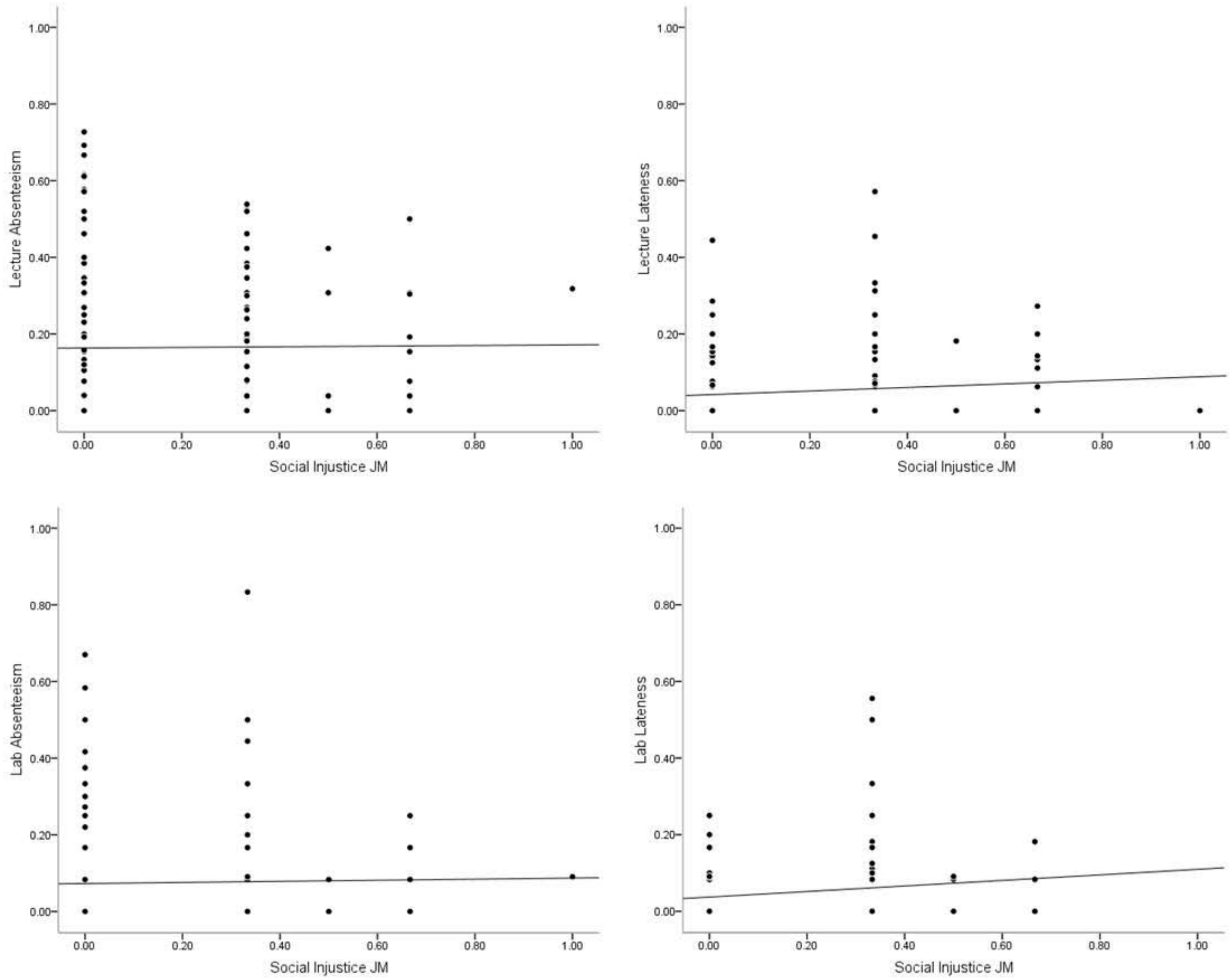


Figure 21. Residuals plots for social injustice with absenteeism and lateness variables.

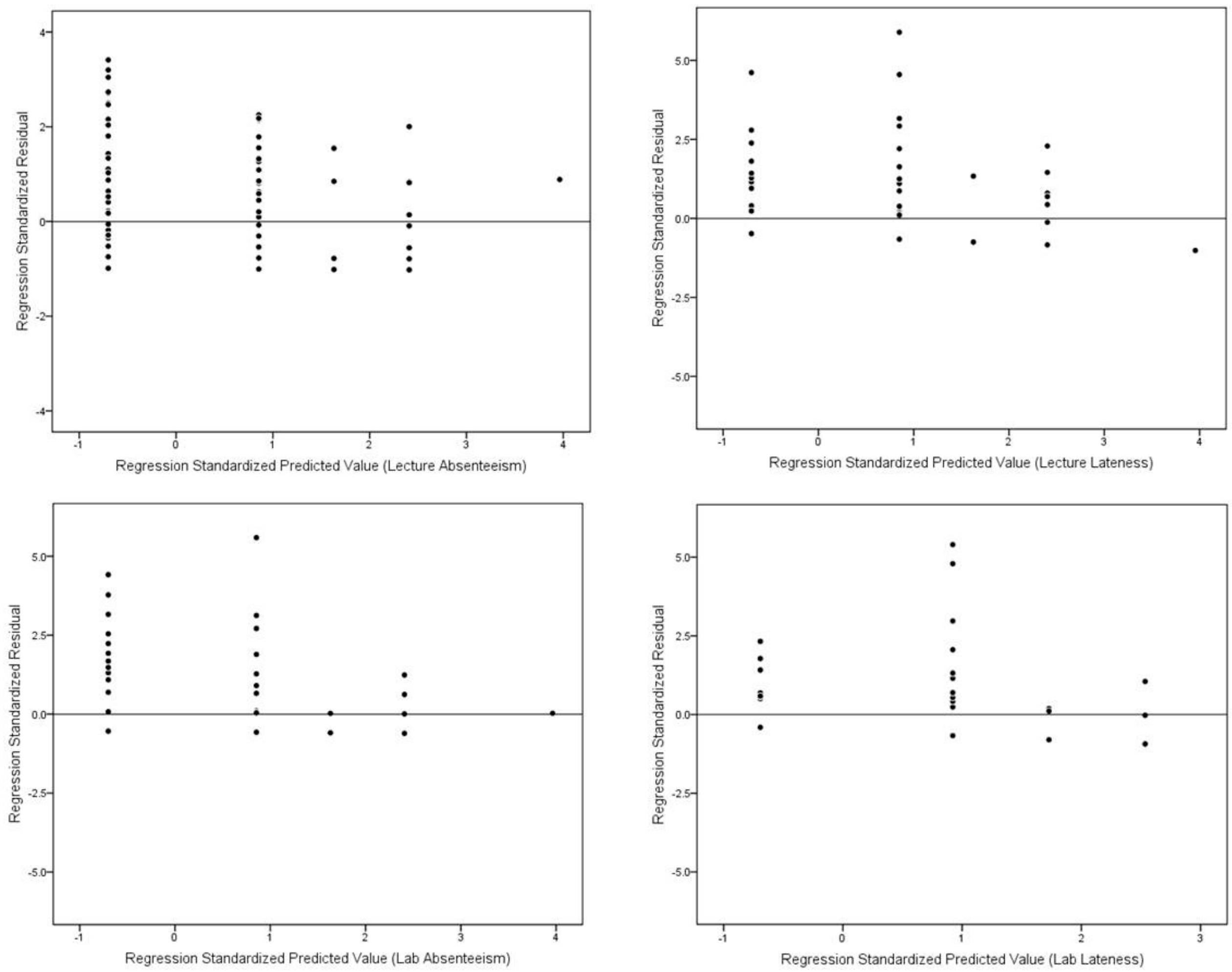
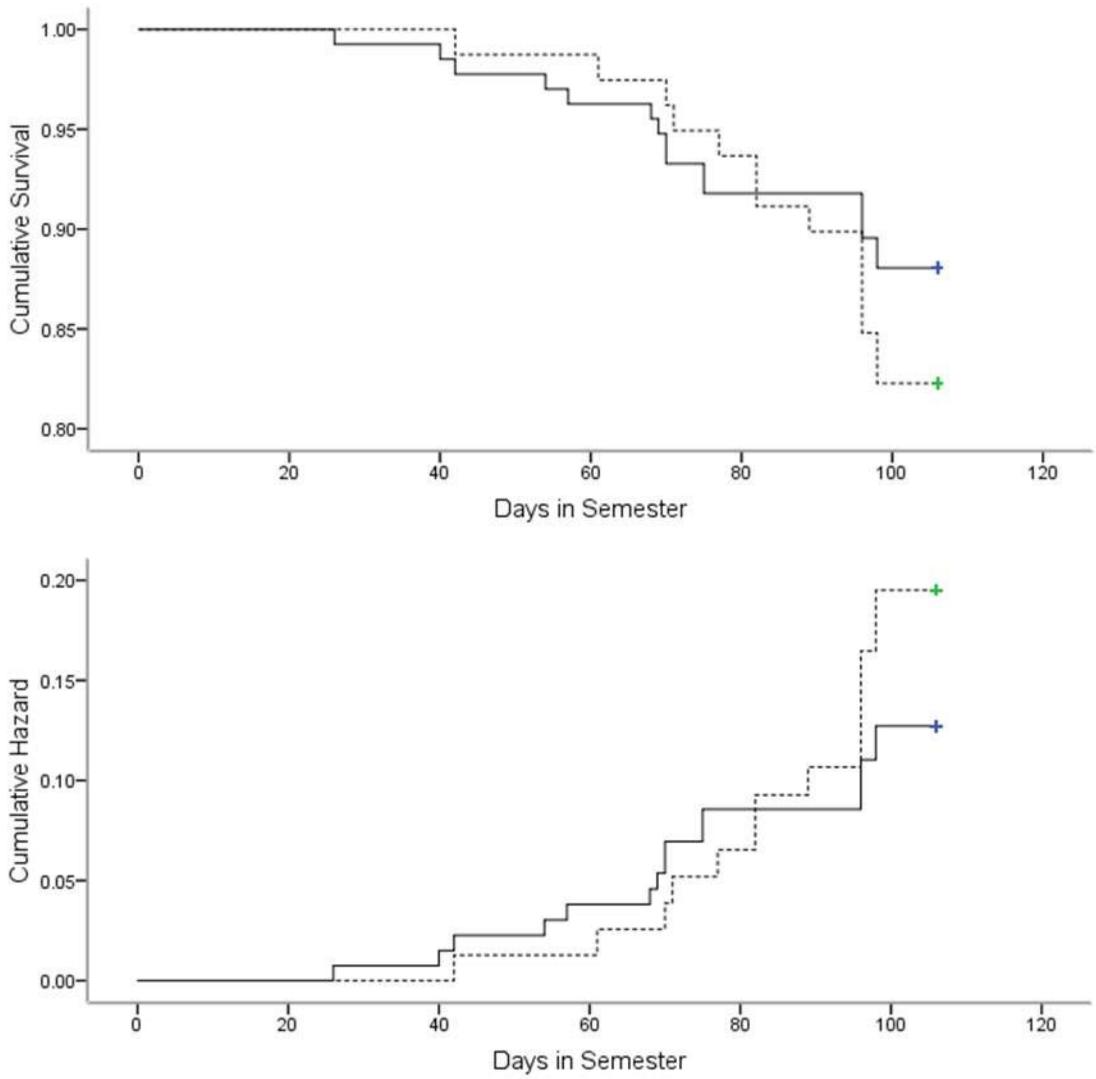


Figure 22. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low social injustice.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 23. Bivariate scatterplots for overall CRT-W scores with absenteeism and lateness variables.

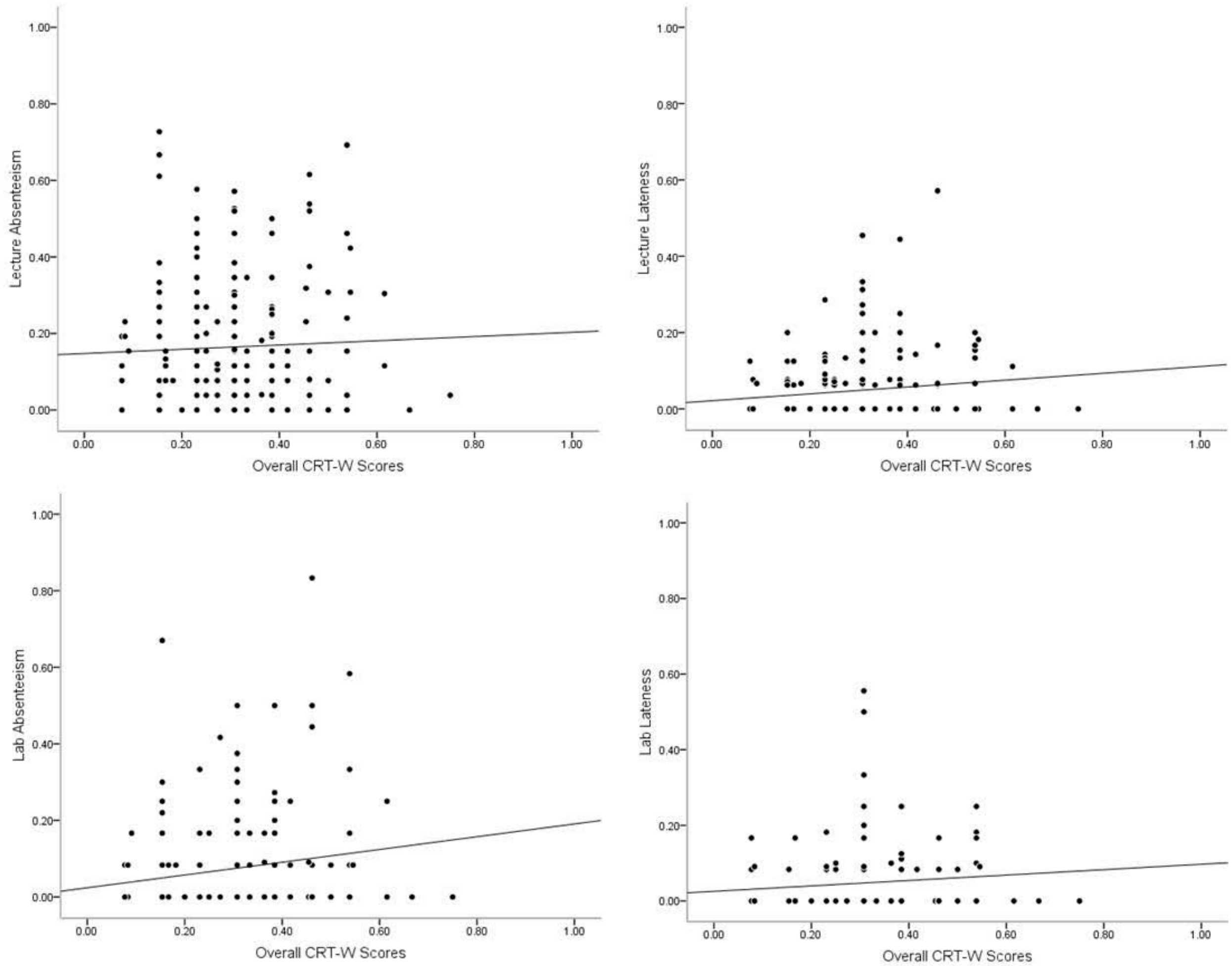


Figure 24. Residuals plots for overall CRT-W scores with absenteeism and lateness variables.

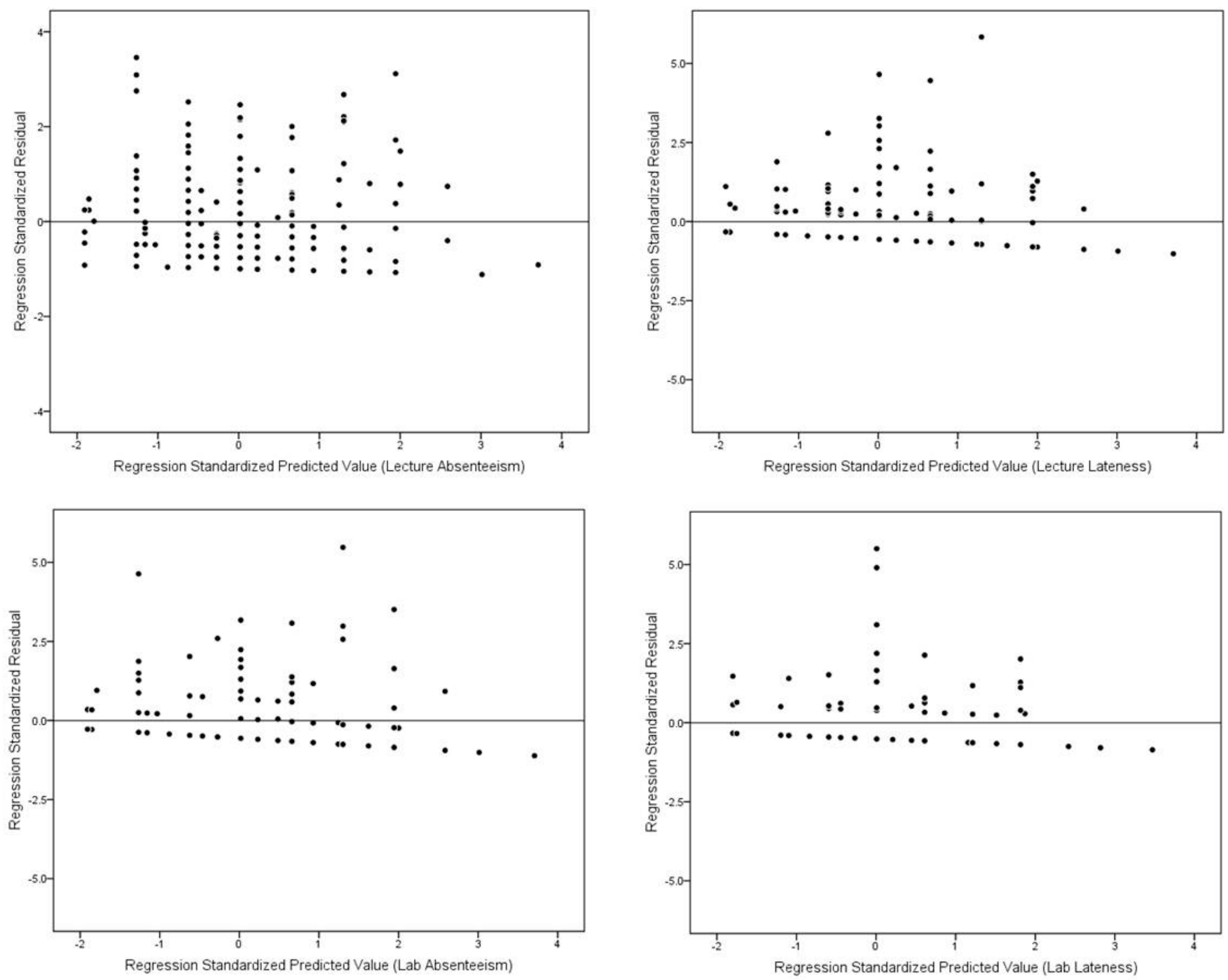
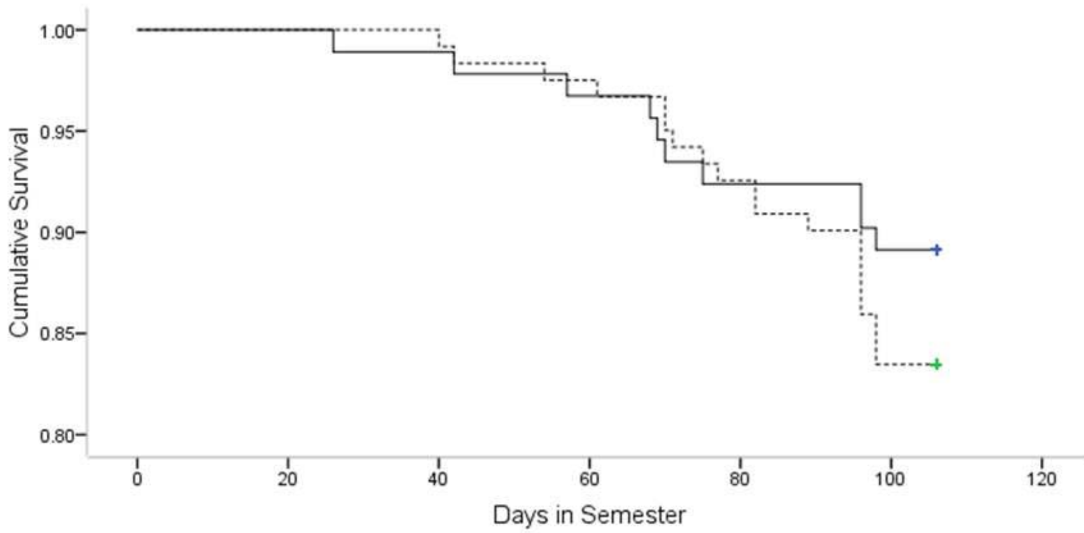
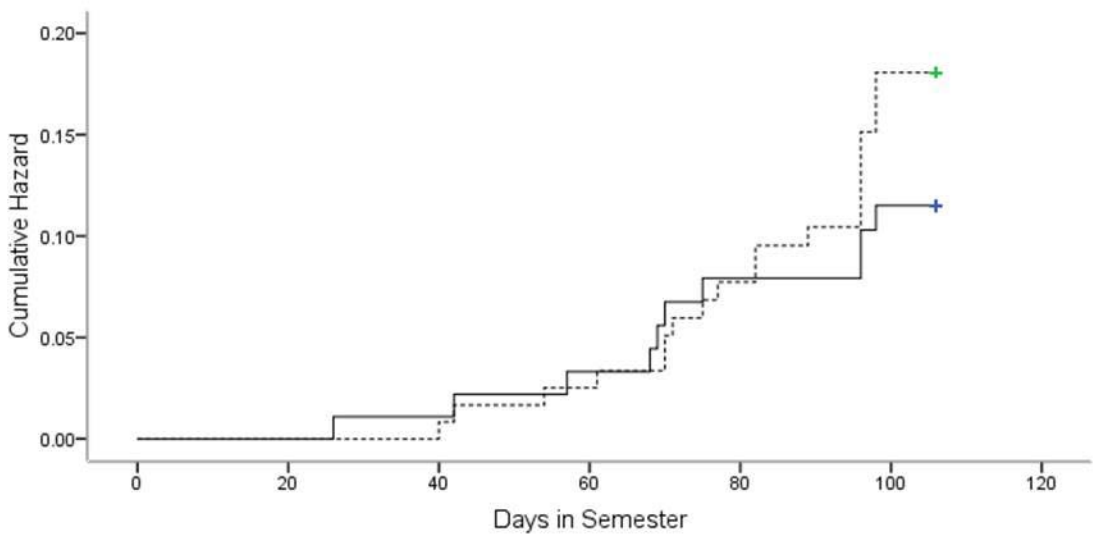


Figure 25. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low overall scores on the CRT-W.





Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 26. Bivariate scatterplots for conditional reasoning variables with early departures.

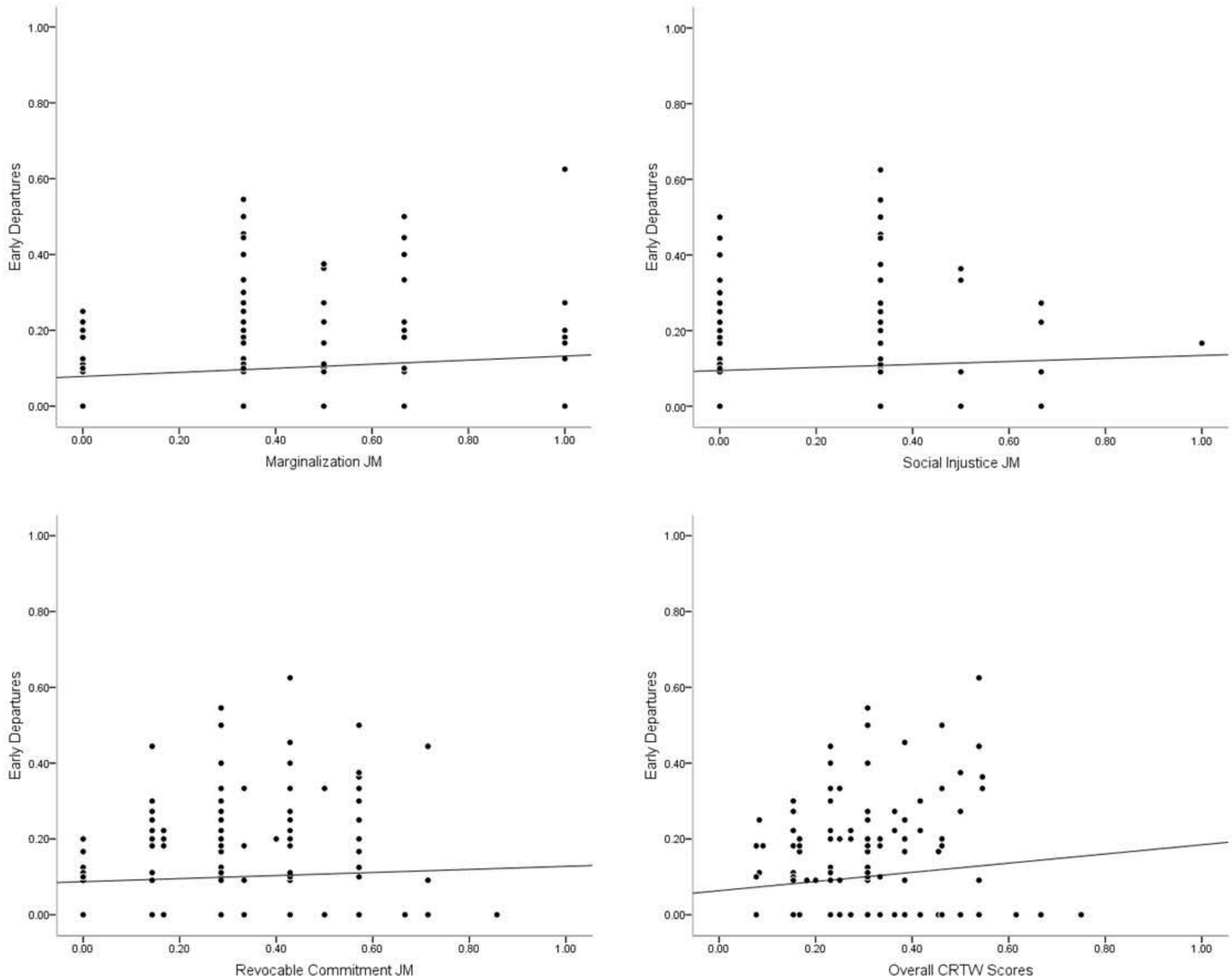


Figure 27. Residuals plots for conditional reasoning variables with early departures.

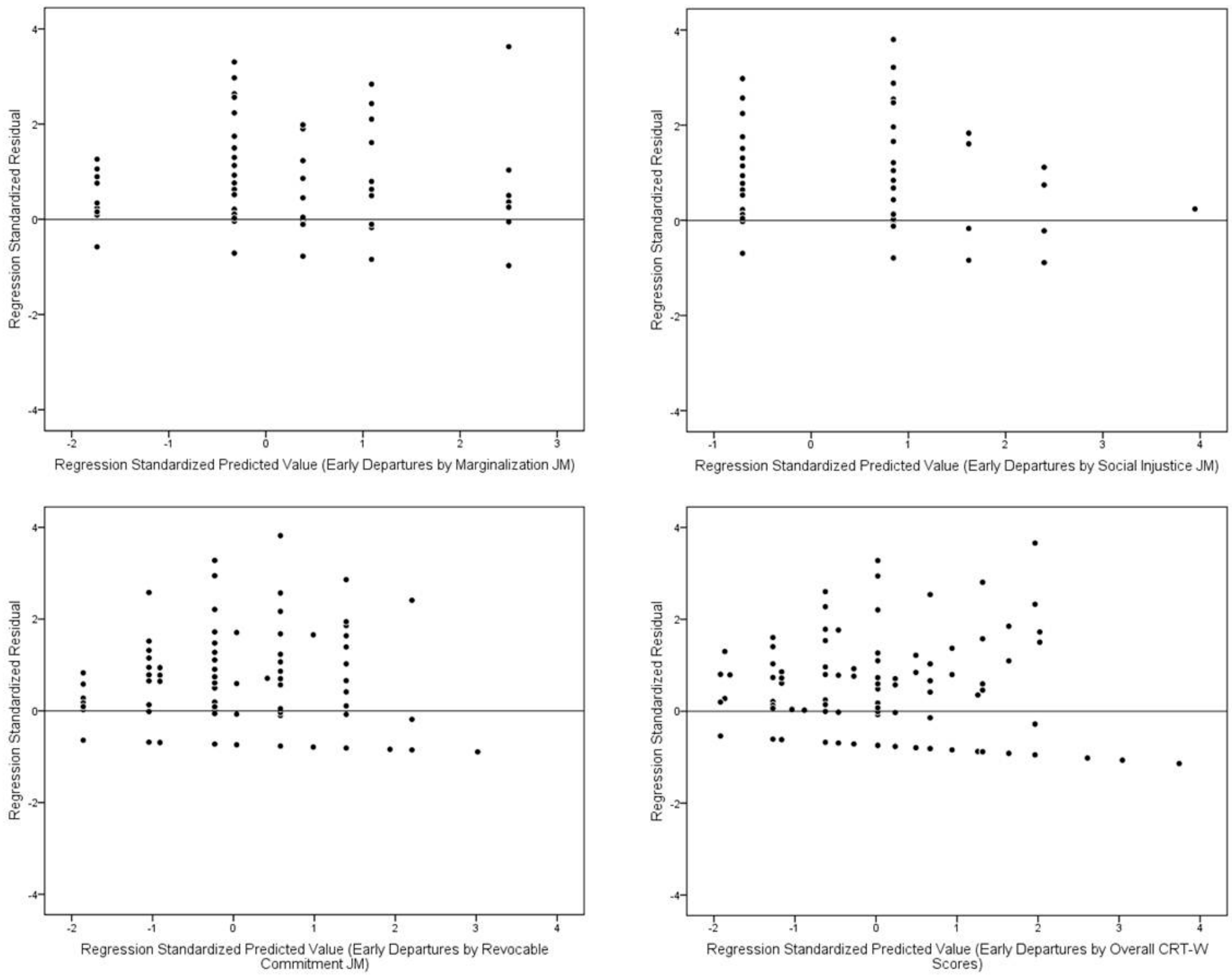


Figure 28. Bivariate scatterplots of conscientiousness with absenteeism and lateness criteria.

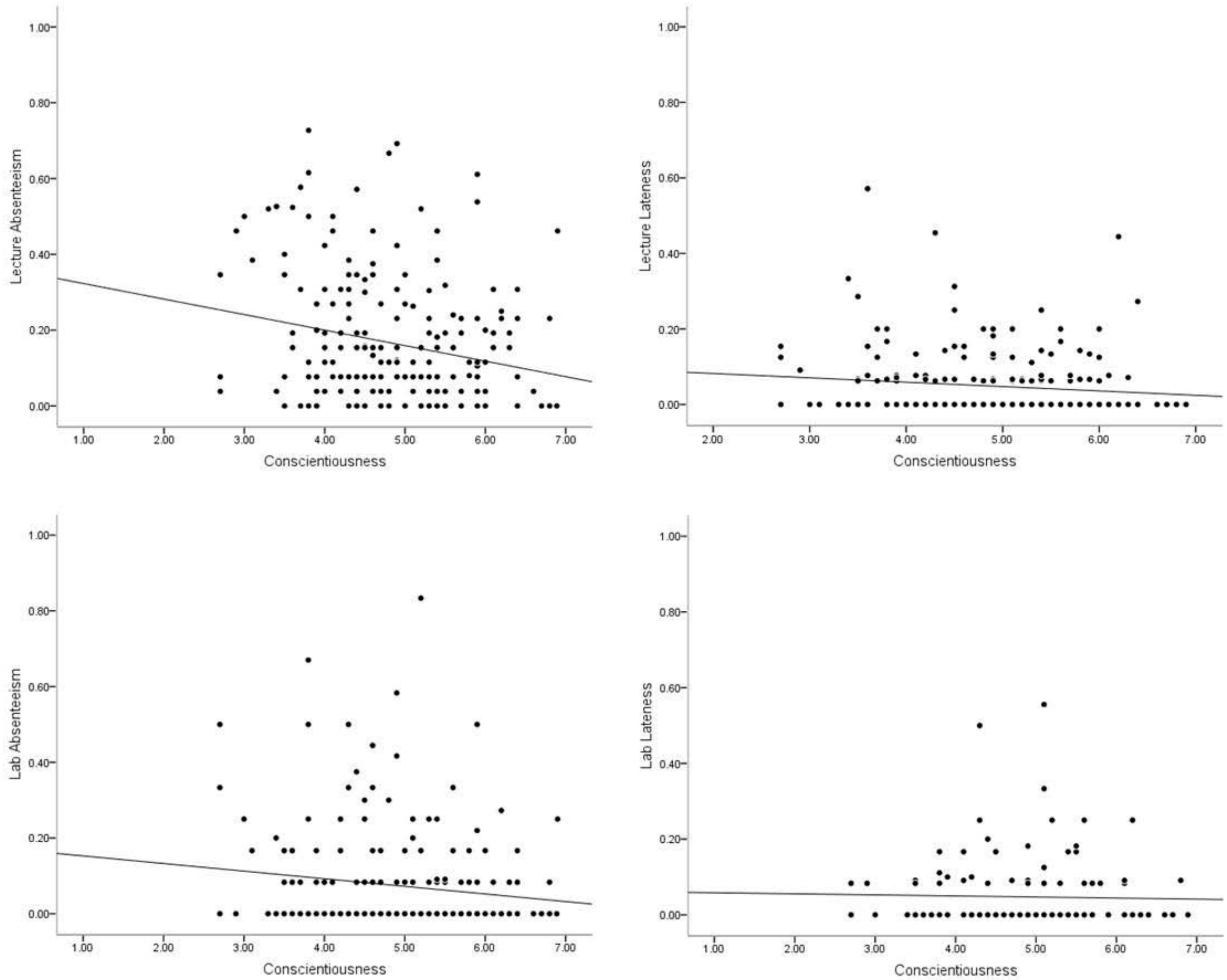


Figure 29. Residuals plots for conscientiousness with absenteeism and lateness variables.

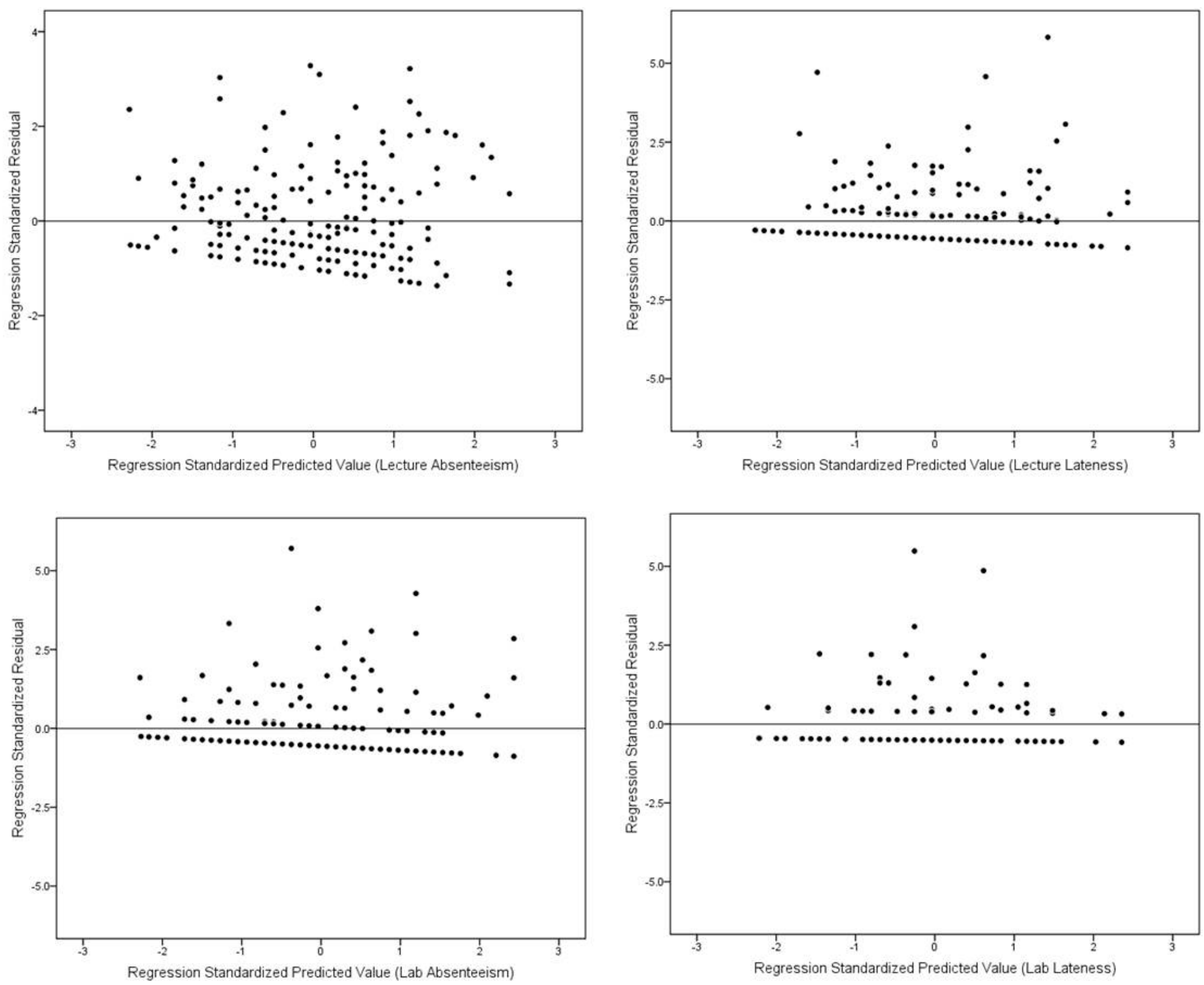
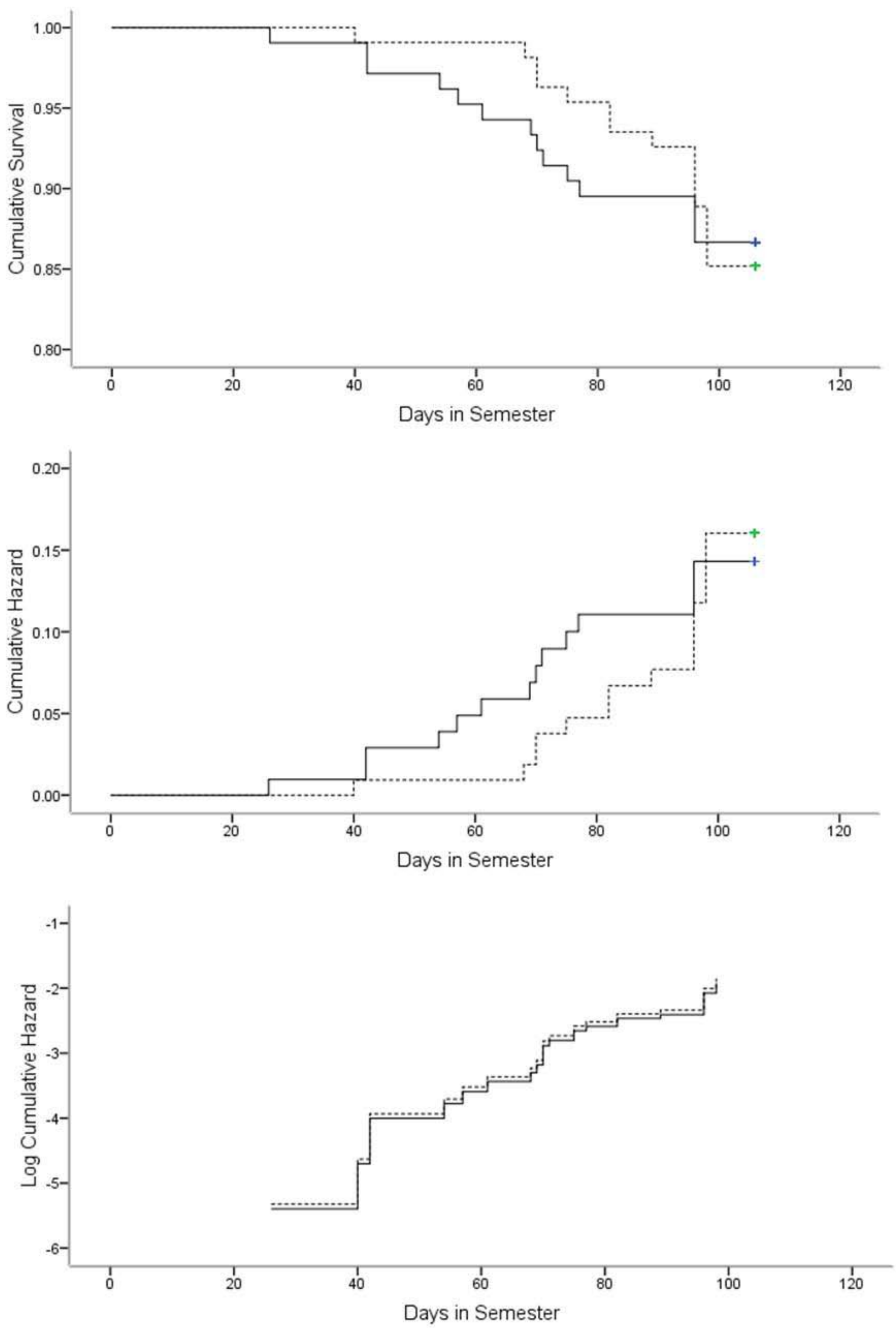


Figure 30. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low conscientiousness.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 31. Bivariate scatterplots of emotional stability with absenteeism and lateness criteria.

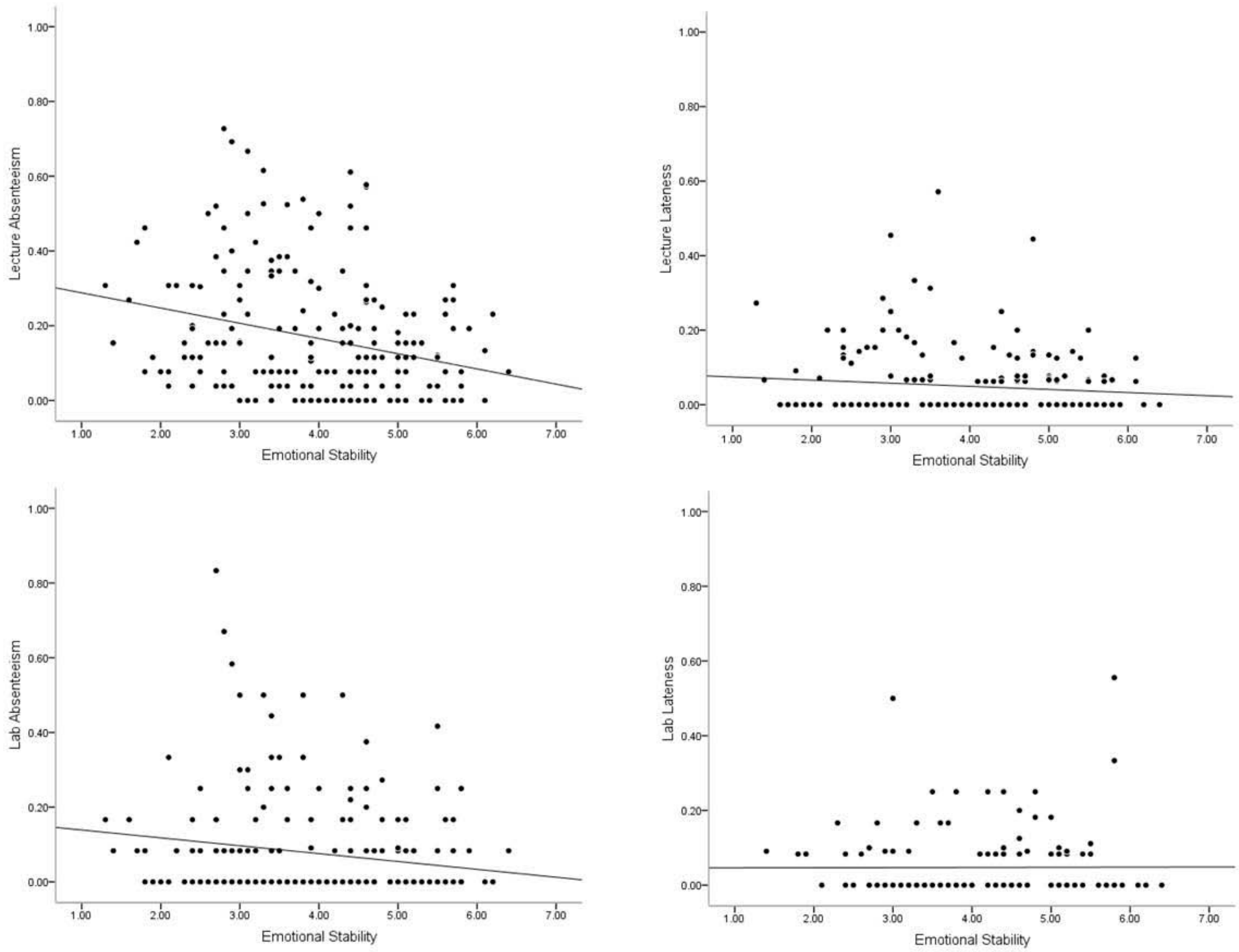


Figure 32. Residuals plots for emotional stability with absenteeism and lateness variables.

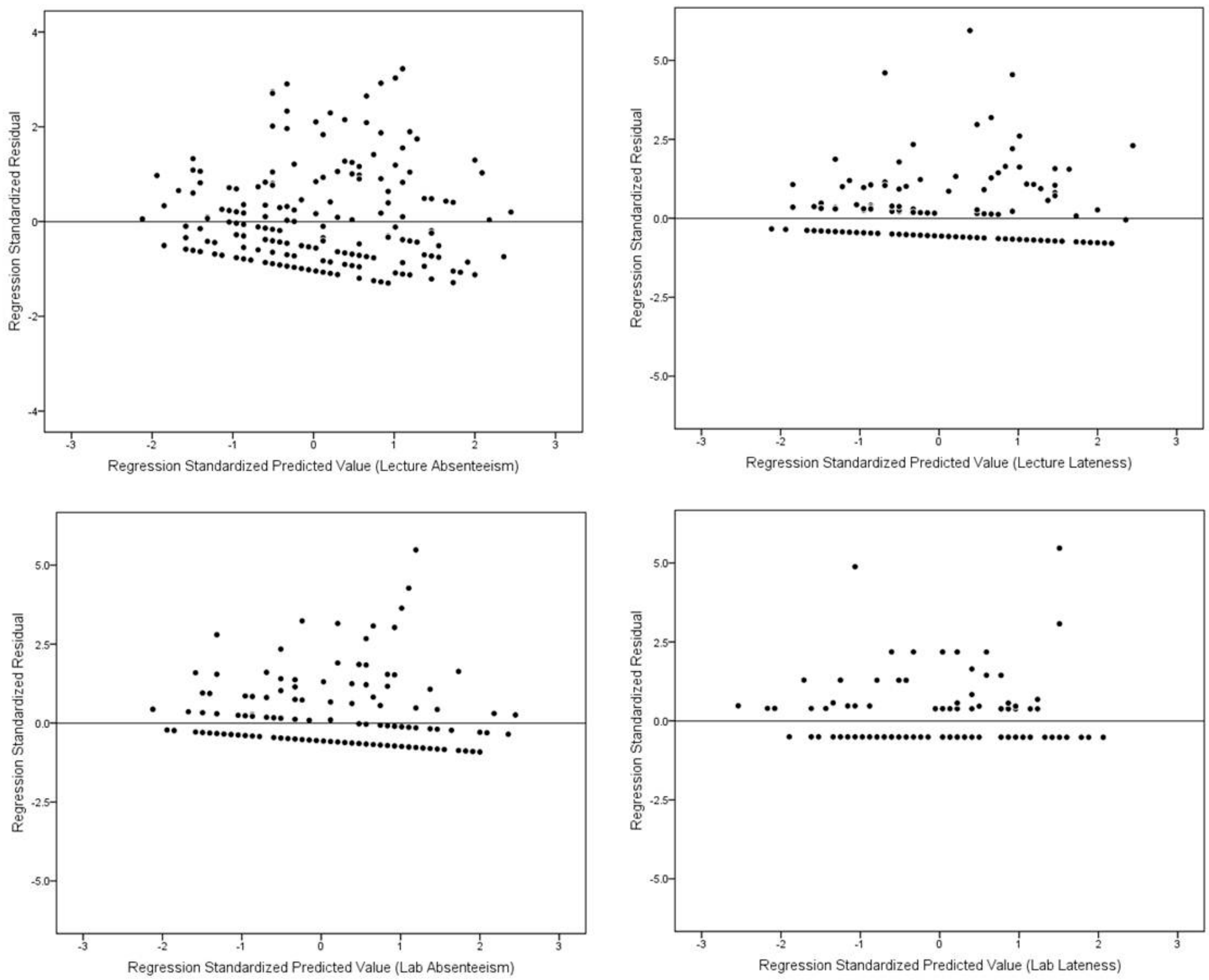
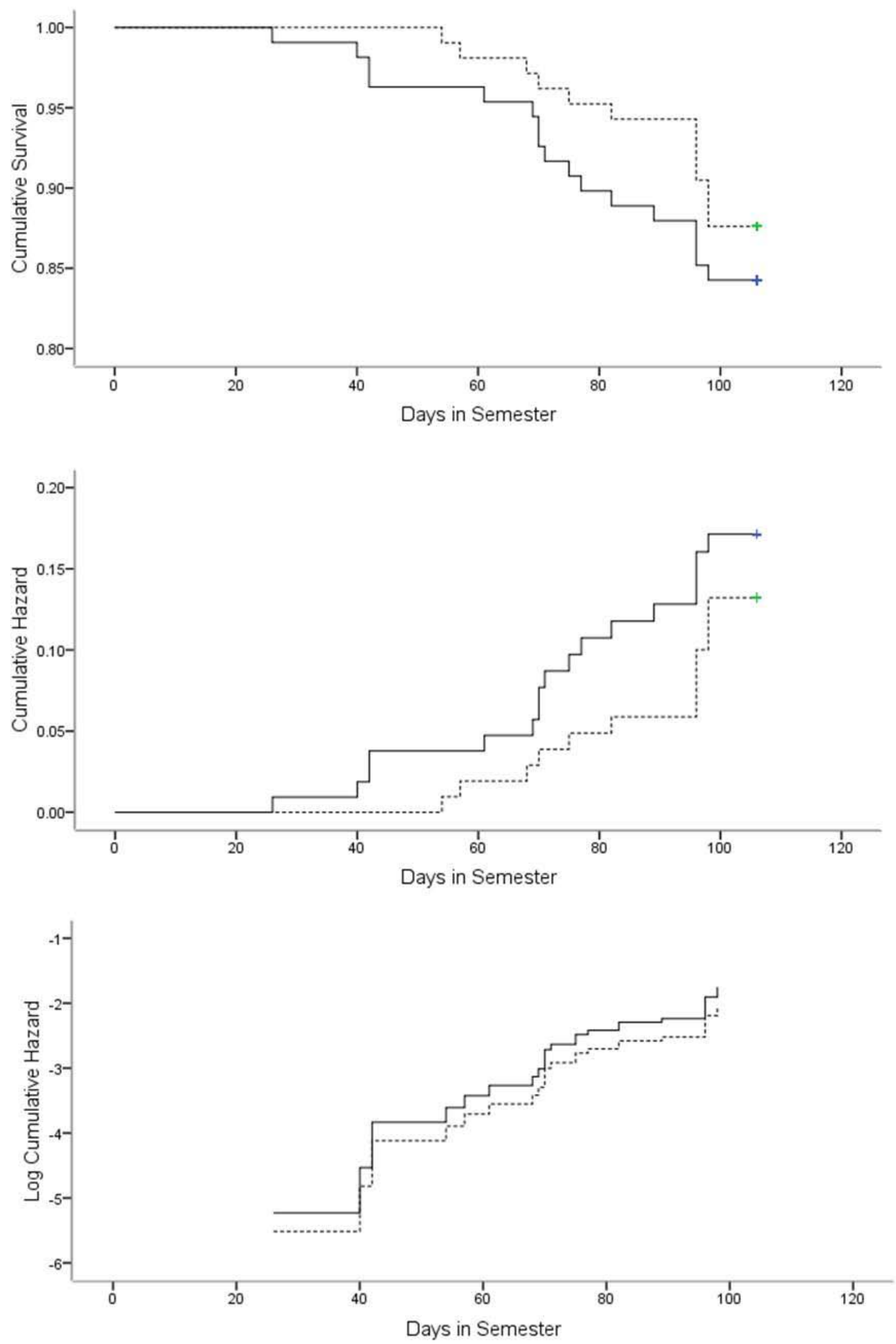


Figure 33. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low emotional stability.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 34. Bivariate scatterplots of conscientiousness and emotional stability with lecture early departures.

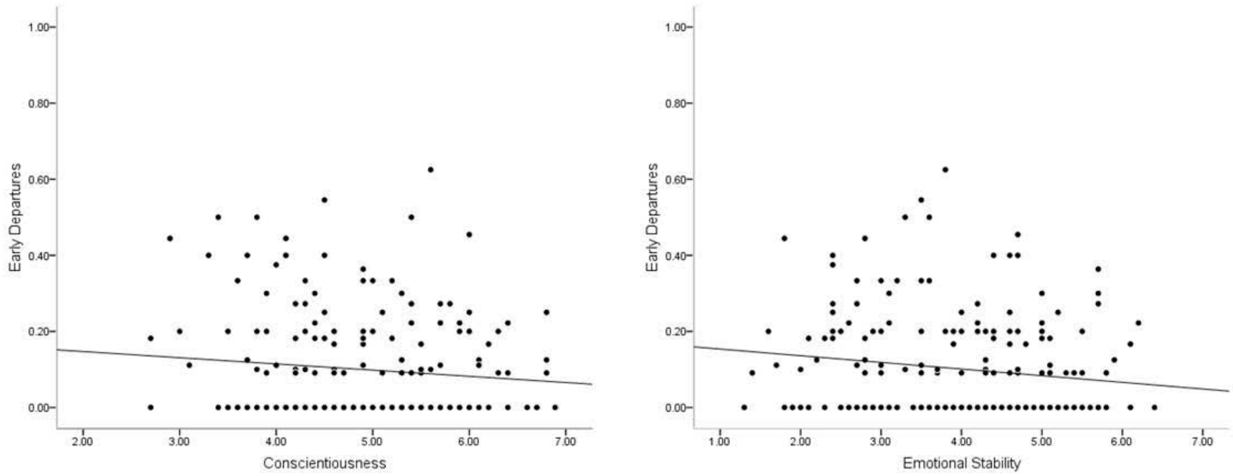
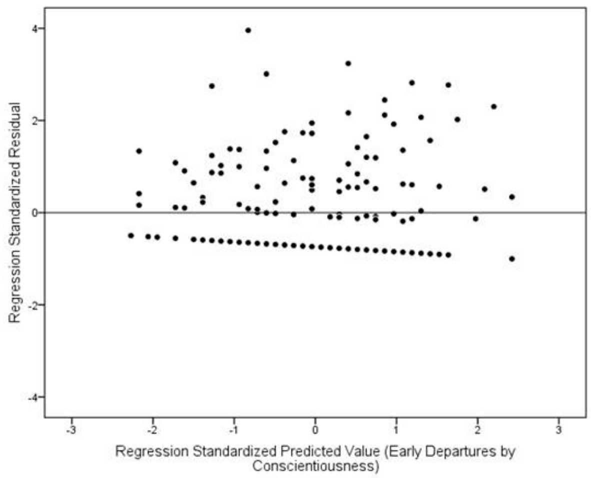


Figure 35. Residuals plots for emotional stability and conscientiousness with emotional stability.



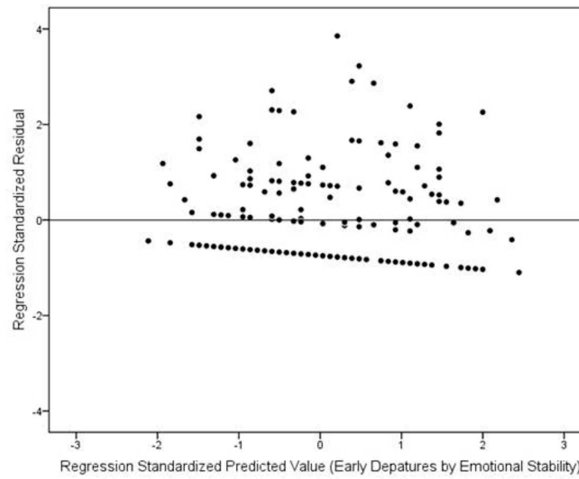
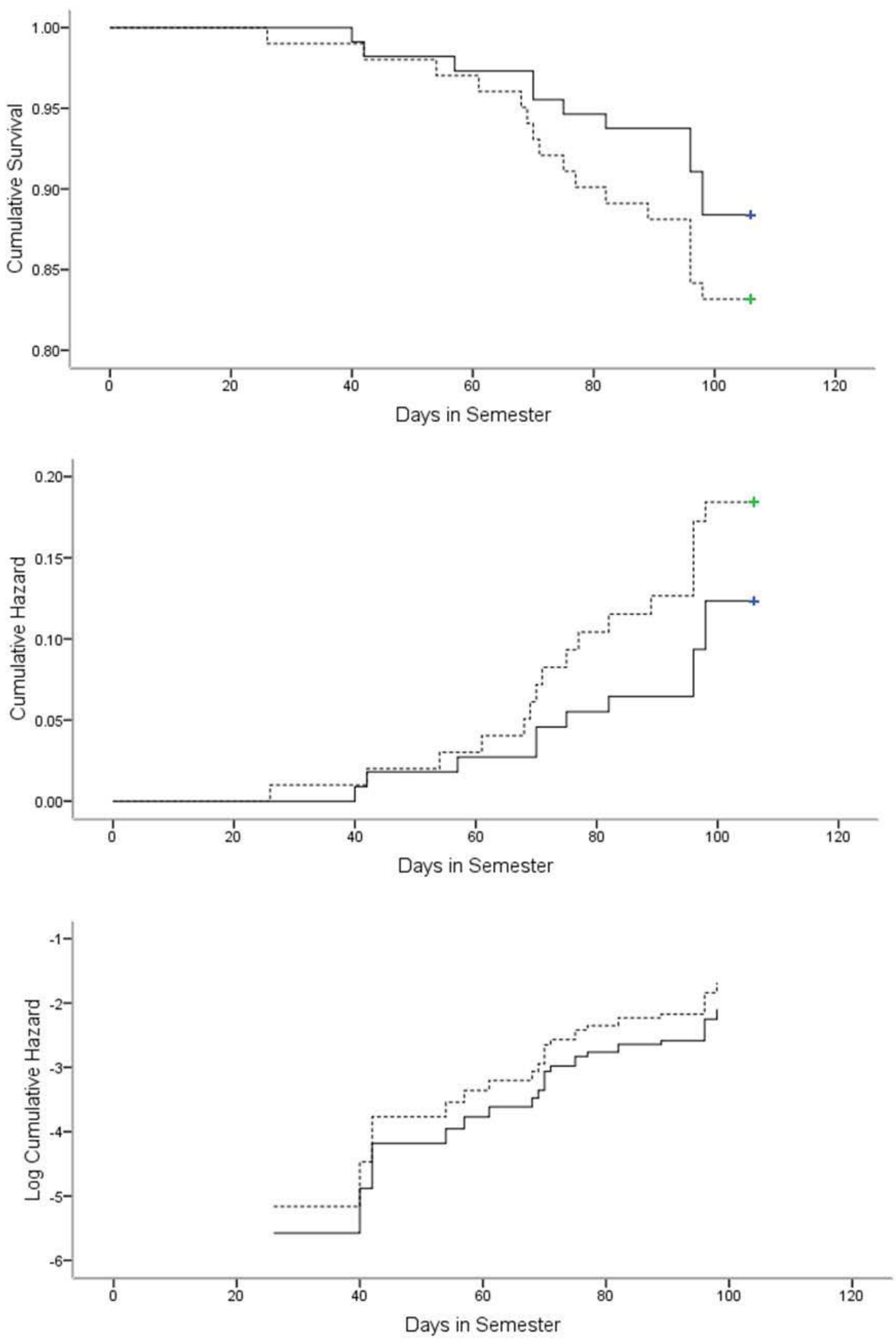


Figure 36. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low prior attendance problems.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 37. Bivariate scatterplots of social embeddedness with absenteeism and lateness variables.

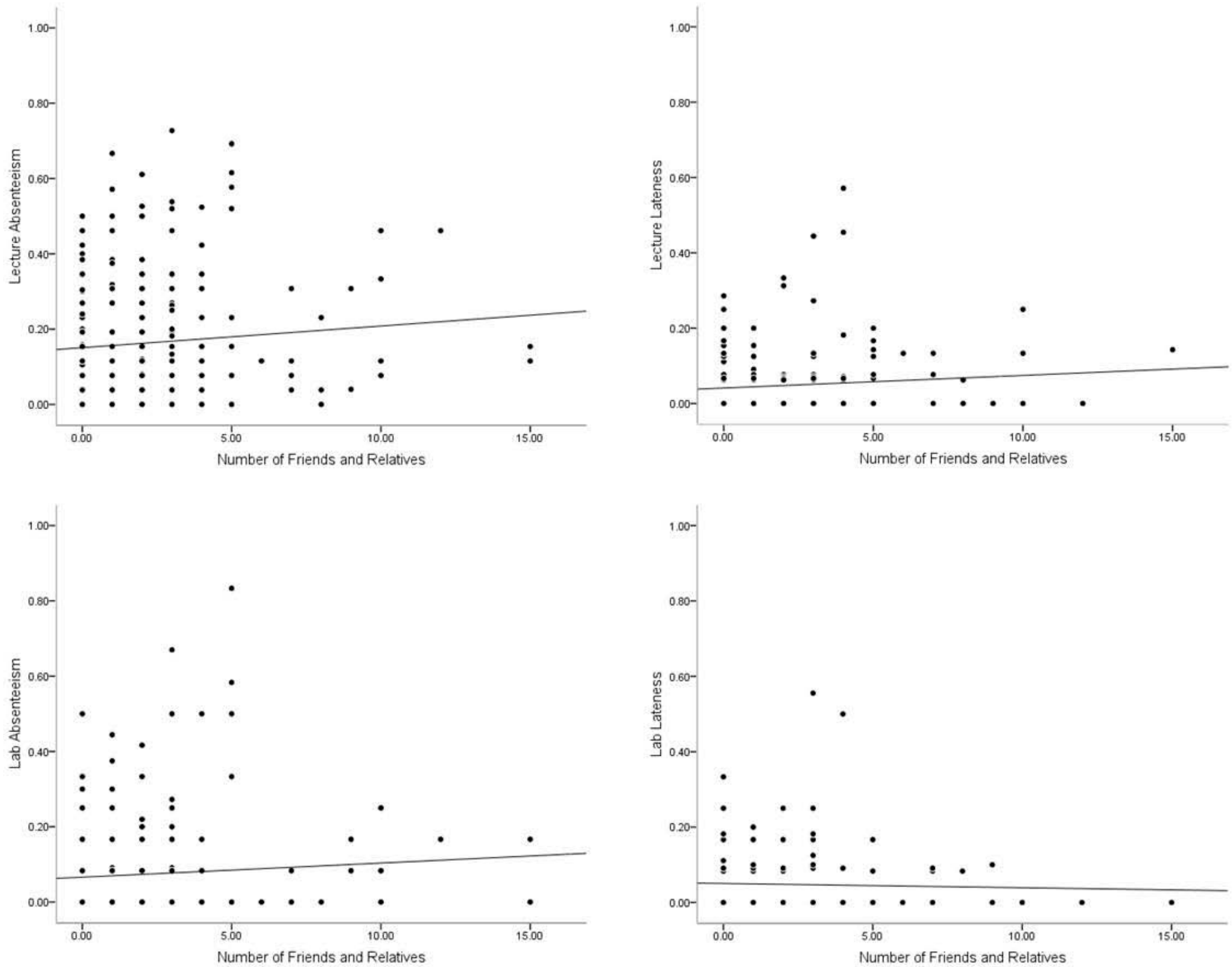


Figure 38. Residuals plots for social embeddedness with absenteeism and lateness variables.

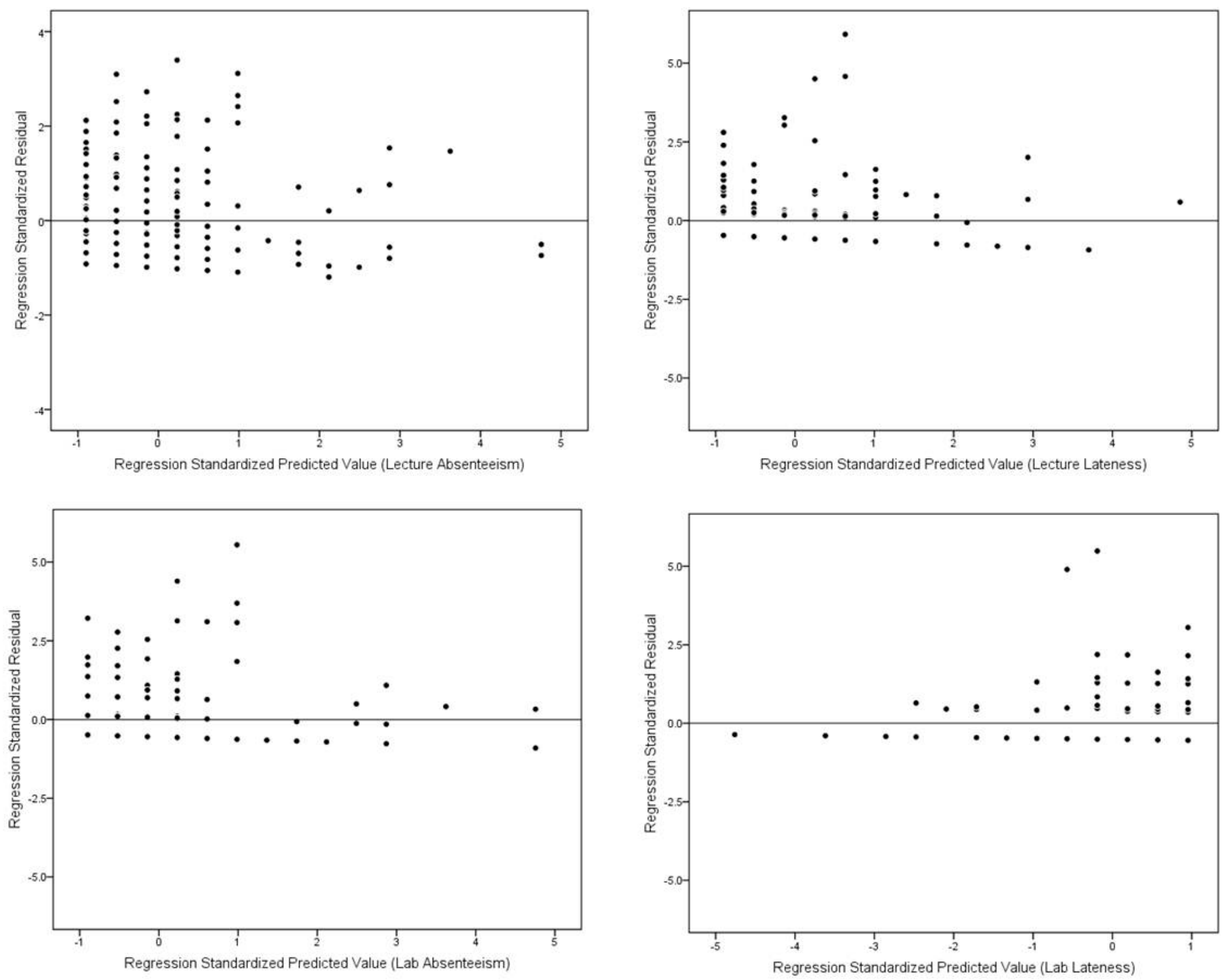
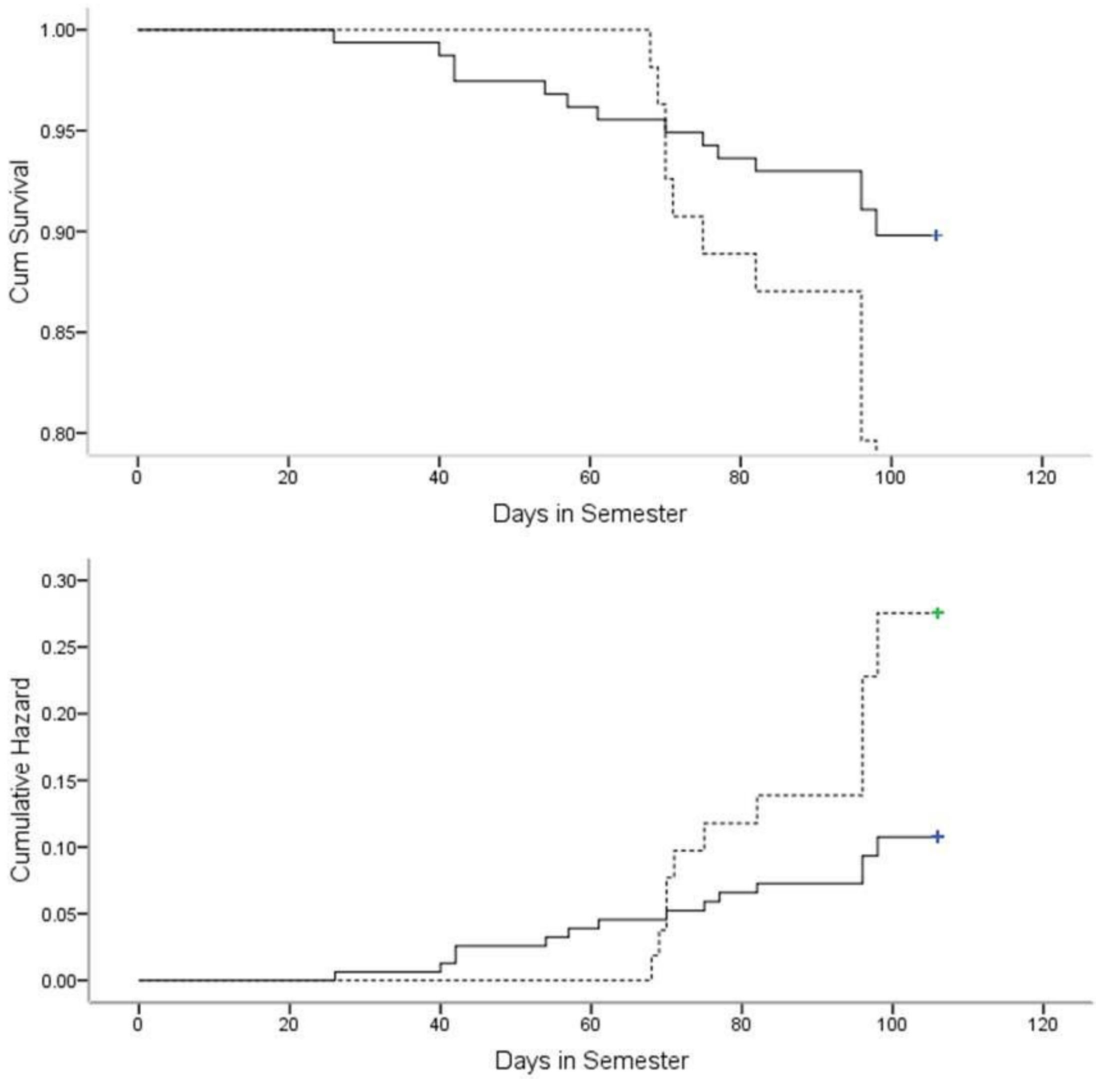


Figure 39. Kaplan-Meier cumulative survival (top) and hazard functions (middle), and Cox estimated log cumulative hazard (bottom), based on high versus low intentions to drop.



Note. Dotted line represents above-median predictor values; solid line represents below-median predictor values.

Figure 40. Bivariate scatter plots for step 3 regression models with absenteeism and lateness variables.

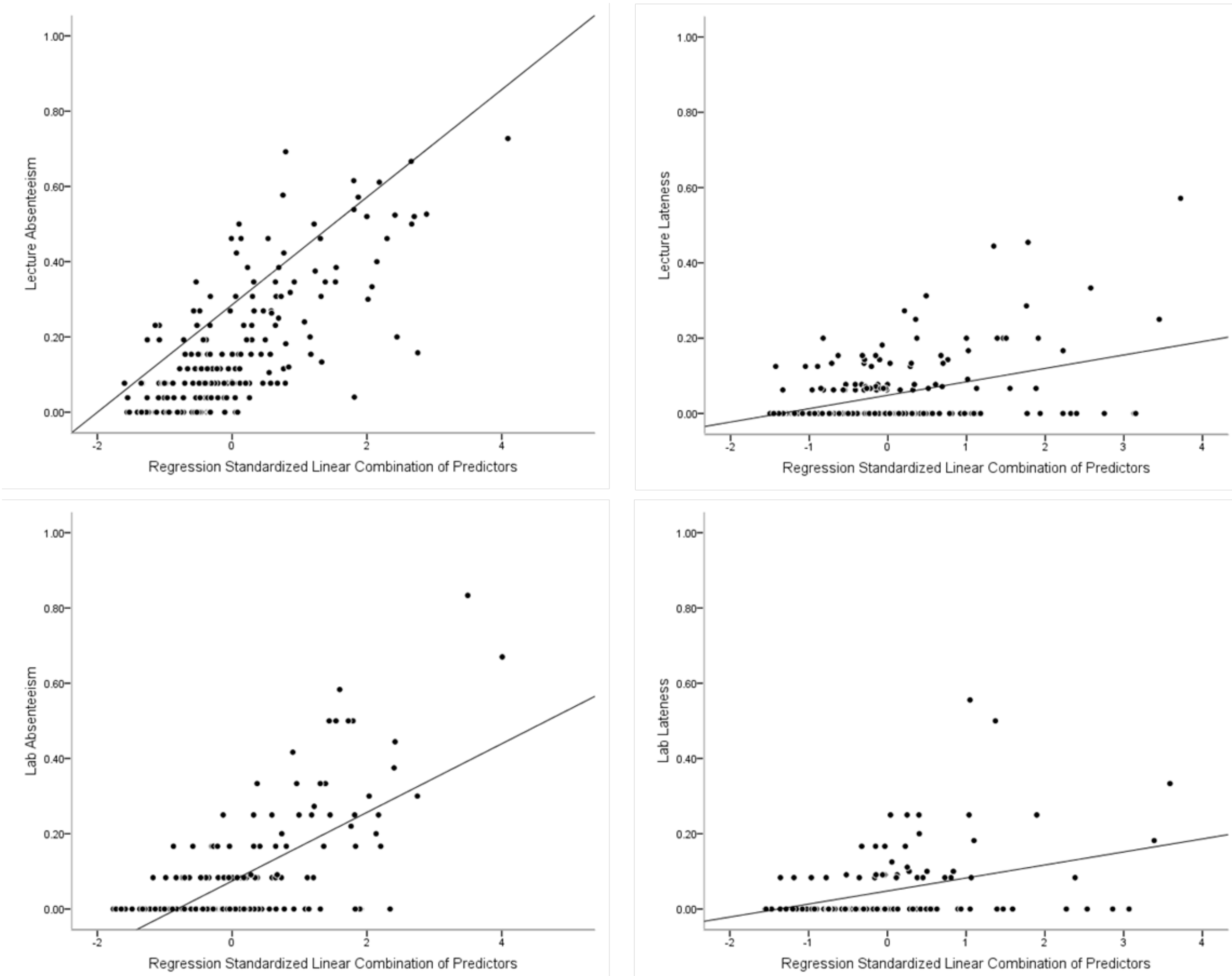
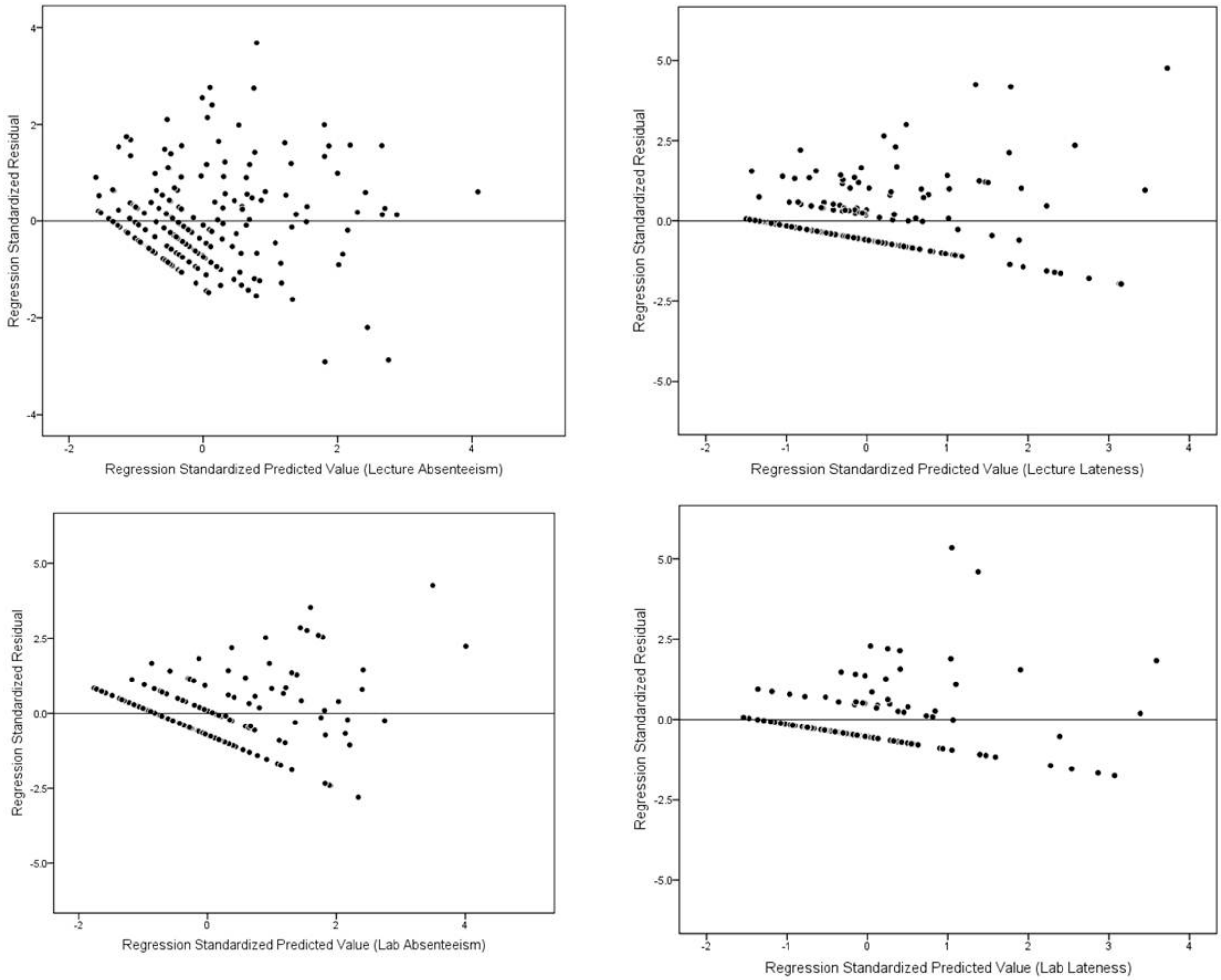


Figure 41. Residuals plots for step 3 regression models with absenteeism and lateness variables.



APPENDIX A

CRT-W Test Booklet

Reasoning Test Booklet

It is important that you write your name and ID numbers clearly below, because I will use this to assign your research credit to YOU.

Name (First and Last): ________________________________

Your WSU AccessID (for example - ab9999): _____________________

9-digit WSU Student ID: ______________________________

Instructions:

The questions on the following pages are a test of your reasoning ability. Please read each question carefully, and respond by circling the one answer that is the most logical. In other words, you can only choose one answer. Mark your response by circling the entire answer, so that it is clear which answer you have chosen. Please take a moment now to read the example item that is shown below, where the correct answer has been circled for you.

EXAMPLE ITEM:

The old saying “opposites attract” has NOT been supported by research. Couples that are more similar in terms of their interests, attitudes, and personalities are generally happier and stay together longer than couples who differ on these things. The same pattern has been found among working adults. Based on this research, which of the following can be concluded?

a. Work teams should attempt to pair up different types of people.

b. Students are too busy for long-term relationships.

c. Companies should seek out new employees who are similar to existing employees.

d. Older adults are more reliable workers.

CORRECT SOLUTION: In the example item shown above, **Option C** is the best or most logical answer. In order to respond, you would <u>circle</u> Option C as shown above.

Please Wait for Instructions.

DO NOT FLIP TO THE NEXT PAGE UNTIL YOU ARE TOLD YOU MAY BEGIN.

The test begins with Question 1 below. Make sure to respond to each question on this test. Circle the <u>one</u> answer that you feel is the most logical.

1. People living in Middle Eastern countries consume less red meat (such as beef) in their diets than people living in the U.S. The U.S. has much higher rates of stomach cancer than Middle Eastern countries, and Middle Eastern people who live in the U.S. and eat typical U.S. diets have an increased risk of getting stomach cancer. According to this trend, which of the following is the most likely to lead to a worldwide reduction of stomach cancer?

    a. Make everyone's diet more like the typical U.S. diet.

    b. Encourage more people to go to college.

    c. Consume less red meat.

    d. Use less food coloring.

2. Location is an important factor in whether or not a business will be successful. A recent survey found that business owners would be willing to pay up to three times their current rental fees in order to re-locate their business to a main street in a downtown area or a busy intersection. Similarly, companies pay more to have their products displayed near the front of department stores. Which of the following is most directly supported by these examples?

    a. College campuses should be circular in shape.

    b. Convenience affects purchasing behavior.

    c. Shoppers always go the extra mile for a product they like.

    d. Bigger signs yield bigger sales.

3. Online college classes are not recommended for all students. Only those students with good time management skills and who are comfortable working on their own should enroll. Many universities have designed short courses that students can take prior to enrolling in order to get a realistic preview of what online courses will be like. Despite such “precautions,” the number of students that drop online courses is usually very high. The most logical conclusion for educators to draw is that:

a. students use online courses as a flexible way to try out different interests.

b. students’ ability to succeed in an online learning environment should be assessed prior to enrolling.

c. writing assignments are poor ways to assess learning.

d. the effectiveness of online teaching varies by geographical region.

4. Many universities are now using a delayed schedule in which classes begin 10 minutes after the hour rather than on the hour. For example, 8am classes start at 8:10am, 9am classes start at 9:10am, and so on. According to several universities, this has led to a reduction in attendance problems, and students report liking this schedule more than the traditional one. Which of the following is the most appropriate advice for a business considering a delayed schedule?

a. Be careful not to compromise important aspects of a productive workplace.

b. Provide reasons for going green at work.

c. Use delayed schedules because they are more closely aligned with people’s natural tendencies.

d. Consider how time zones may affect travel.

5. The old saying that, “there are a lot of fish in the sea …,” suggests that one should consider many possibilities before making a choice, and that the process of exploring alternatives leads to better final decisions. For example, one should try many different jobs before deciding on a career, and come up with a number of ideas before deciding what to write about for a senior thesis. Which of the following is the biggest problem with the fish-in-the-sea saying?

a. Other cultures have different sayings.

b. It still implies that one eventually has to make a “final” decision.

c. It overlooks the possibility that the best option will be discovered early on.

d. It is rare that online dating leads to a meaningful relationship.

6. High school policies regarding attendance are usually very strict. Tardiness and unexcused absences can result in detention or other disciplinary actions. In college, professors vary widely in their views on classroom attendance. For example, some penalize absence by assigning a lower grade, and others have no attendance policy. Which of the following can be concluded based on this trend?

a. Adult students are given the opportunity to demonstrate respect.

b. High schools are moving toward shorter class sessions.

c. Attendance policies don’t work in athletic settings.

d. Adults are treated more reasonably than children and adolescents.

7. More and more first-time marriages are occurring later on in adulthood - for example, between partners that are in their 30’s and older. Marriages have higher success rates when the partners entering into the marriage are older. This trend suggests that:

a. young newlyweds have not had the opportunity to date as many potential partners.

b. Americans are similar to Europeans.

c. many people do not fully realize the great responsibility that comes with marriage.

d. some states are experiencing radical climate change.

8. Some people have made a comparison between the human circulatory system and public transportation in a city. The human circulatory system supplies all parts of the body with a continuous flow of oxygen and nutrients by pumping blood. If blood flow to a part of the body (such as an arm or leg) stops, that part of the body can die. Which of the following best shows how public transportation is similar to the human circulatory system?

a. Cities need more art and music.

b. Biking to work is very difficult for people near parks.

c. It’s important for people to regularly travel through all parts of a city.

d. There is no way to predict bus times.

9. It is common for airlines to overbook their flights by selling more tickets than there are available seats. By overbooking, airlines ensure that all seats on the flight are actually filled in case some passengers need to cancel or fail to make it on time. Which of the following most weakens overbooking as a strategy for airlines to make more money?

a. Overbooking doesn't work in those rare cases when everyone is on time for their flight.

b. Most airlines have done away with in-flight meals.

c. Overbooking is disrespectful to reliable passengers.

d. The best pilots are used for international flights.

10. Scheduling shifts for doctors on holidays and weekends is a continuous problem for hospitals. Medical residents (or doctors in training) usually get stuck working the least desirable shifts. Usually there are too few residents to allow for better flexibility. What is the best solution?

a. Marketing research should examine the needs of nursing staff.

b. More senior doctors should pick up some weekend and holiday shifts.

c. Hospitals should focus on making more money.

d. Hospitals should employ more residents.

11. New Year's resolutions (or personal goals at the beginning of the year) are rarely effective. Stopping smoking and exercising more often are two commonly failed resolutions. New gym memberships peak around the beginning of January and the resulting spike in gym attendance usually drops off by mid-February. Some research indicates a similar timeline for smokers who failed to stop. Which of the following is the most logical solution to this problem?

a. Refer a friend to a heath food specialist.

b. Use fewer sick days at work.

c. "Lower the bar" for what is expected of a resolution.

d. Take resolutions more seriously.

12. A current issue being debated in many universities is that new professors are being offered larger starting salaries than professors who have been with the university for many years. Some have argued that the incentive to work hard as a senior professor is greatly decreased as a result. Universities generally do not have the budgets to simply increase senior professor wages to a level that is more in line with what new professors are offered. What is the most logical solution?

a. Use a salary cap that limits how much new professors make at first.

b. Re-locate rural university campuses in more urban settings.

c. Discipline un-motivated senior professors as needed.

d. Pay university librarians more.

13. In general, the harshness of punishment for a particular offense is associated with its frequency – that is, how often it occurs. Parking violations are punished less harshly than is running a red light, and running a red light is punished less severely than are hit-and-run incidents. If the same general pattern was true of negative behaviors in the workplace:

a. managers would seldom file harsh disciplinary reports.

b. employees would stop talking around the water cooler.

c. additional training would be required of managers and employees.

d. managers would be unlikely to do or say much in response to employees being late for work.

14. Shaking hands is an important form of communication in business. For example, a handshake is a common signal between parties that a deal or agreement has been reached. However, research indicates that whether or not someone has a “good handshake” is not related to whether he or she will be a reliable business partner. Which of the following is the most appropriate conclusion?

a. There are other better indicators of personal character than a handshake.

b. Only some business schools are appropriately certified.

c. It is difficult to do job interviews over the phone.

d. Handshakes are not lasting agreements when circumstances change.

15. A recent study found that hand washing without the use of soap was just as effective at killing germs as hand washing with soap. A more important factor than soap versus no-soap was how long the subject scrubbed under running water. Germs were effectively killed as long as the subject scrubbed their hands for more than 10 seconds. Based on this study, which of the following is the best safety recommendation for hand washing?

a. Wash hands for a longer time and use soap as an added precaution.

b. Wear gloves at all times.

c. Wash hands for a shorter time, but doing so harder.

d. Don't use hand towels.

16. There is a popular story about the owner of a company and an employee who made a mistake that ended up costing the company over 1-million dollars. Afterwards, the employee was shocked to learn that he would NOT be fired for his mistake. In meeting with the employee, the owner explained, "why would I fire you now, after I just invested 1-million dollars in your education?" Which of the following most directly shows the lesson in this story?

a. A company moves its location to a different city.

b. An employee is given a second chance after lying about a fake sick day.

c. A CEO retires at an early age.

d. A new worker agrees to some performance improvement goals.

17. Some have argued that labor unions are no longer necessary. Labor unions were formed to argue on behalf of the common worker at a time when managers and companies were willing to exploit workers (or treat them poorly) for increased profits. However, today, companies are regulated much more tightly to ensure that human rights are respected in the workplace. Which of the following most weakens the idea that labor unions are no longer necessary?

a. There are too many independent farms.

b. It is unlikely that the auto industry will experience a full recovery.

c. Workers have endured centuries of mistreatment by their employers.

d. It would be chaos for each worker in a factory to bargain or negotiate with the company on his own behalf.

18. Company policies are always changing. Recently, many companies have adopted “family friendly” policies that allow employees greater flexibility, such as allowing them to work from home once in a while. Fifty years ago, employees were expected to keep their work and family life completely separated. Which of the following is most likely based on this trend?

a. Things like punctuality (or being on time) will be less important in the future.

b. It will be difficult to do online shopping.

c. Employees who insist on being present will be even more valuable.

d. Coastal cities, such as Los Angeles, will make a full economic rebound.

19. A prenuptial agreement (or a “pre-nup”) is an arrangement that specifies how a couple’s wealth will be divided if they get a divorce. For example, a pre-nup could indicate who gets any properties that are owned or investments. If a couple doesn’t have a pre-nup, the most likely outcome of divorce is a 50-50 split. The most logical conclusion about couples that choose NOT to get a pre-nup is that:

a. they are more certain about their relationship than couples that get a pre-nup.

b. they met while traveling.

c. neither partner has accumulated substantial wealth.

d. they attended the same college.

20. Replacing a good employee that quits can be very costly. Traditionally, companies have viewed employee loss as a negative thing that should be prevented. More recently, companies have come to recognize that not all employee losses are harmful. Some level of regular turnover is actually healthy. This change in perspective most logically indicates that:

a. climate change is real.

b. it’s increasingly easier to replace employees that quit.

c. companies are better off without the types of employees that quit.

d. the number of newspapers published is increasing.

21. Although it is well known that many high schools struggle to keep students from “skipping” school, it is somewhat surprising that the same problem often exists with teachers. A fairly common problem involves teachers who use their sick-leave days to get extra vacation or abuse their use of personal days. If research found that this problem was unique to the profession of teaching, which of the following would be the best solution?

a. Prevent students from taking multiple classes with the same teacher.

b. Find specific ways to make teachers more accountable for their behavior.

c. Encourage fewer extra-curricular activities for students.

d. Increase teacher salaries to be more in line with other professions.

22. Research suggests that losing a job can be one of life’s most stressful events. Similar studies find that unemployment rates are a strong predictor of suicide attempts over time. As unemployment rates go up, suicide attempts spike shortly thereafter. Based on this pattern, which of the following is most likely?

a. Suicide is rare in places with severe temperatures.

b. The number of service jobs has steadily increased over time.

c. How upset a person gets about losing a job is related to the number of other opportunities available.

d. Job loss is especially difficult for people who are intensely invested and committed to their work.

23. The old saying, “If you can’t beat em, then join em,” suggests that a superhero who can’t find a way to defeat his enemies, should join forces with them. For example, if the top athletes in a sport use performance-enhancing drugs, then the only way for others to compete at the highest level is to use the same substances. Which saying shown below is most closely connected to the idea behind the “If you can’t beat em, then join em” saying?

a. What’s up is down.

b. Monkey see, monkey do.

c. Honor thy neighbor.

d. People don’t make the rules, they simply play by the ones that exist.

24. Over the past 10 years, the number of older Americans going back to school has steadily increased. Some “non-traditional” students go back to college several years after having completed their first degree or after having worked in a job for several years. Many non-traditional students make the choice to go back to college after getting laid off or having a hard time finding a job. Based on this trend, which of the following is most likely?

a. The price of coffee is stable.

b. Unemployment rates are linked to university enrollment rates.

c. Disease rates are lowest in the northern states.

d. There will be very few majors for college students to choose from in the future.

25. Being labeled a “flip-flopper” during an election can have serious consequences for a political candidate. A recent study found that candidates who changed their stance on one or more issues during a campaign were less likely to be elected. Interestingly, the same study indicated that, if elected, “flip-floppers” were more likely to be re-elected for a second term. The most likely reason for this is:

a. federal and state politics differ quite a lot.

b. politicians who are flexible in their thinking are better able to adapt.

c. electronic voting machines are only available in certain states.

d. current office-holders win most elections simply because voters know them.

26. Historically, the daily commute to work has ranked among the most stressful events in the lives of working adults. In the past few years, smart phones and GPS devices have incorporated features such as instant traffic and road construction updates. In order to reduce commuter stress, the main objective of this new technology should be:

a. avoiding the frustration of sitting in traffic for extended periods of time.

b. increasing the quality of public radio.

c. eliminating new road construction.

d. making it easier for commuters to plan ahead for a reliable route to work.

27. Successful law firms are known for their planning. The “lawyer pipeline” is a specific plan for how the firm would adjust if one lawyer were to leave – for example, it might specify the person who is next in line to step into his or her position. Using a pipeline plan is particularly important in law because other firms use very aggressive recruiting strategies. Which of the following is the best reason why an *engineering firm* would need a pipeline plan for *engineers*?

a. Talented engineers are always on the lookout for better opportunities elsewhere.

b. Many lawyers are cross-trained in other disciplines.

c. A bachelor’s degree is required for engineers.

d. As with law firms, engineering firms cannot afford any dip in productivity.

28. A famous movie director was once quoted as having said that, “95% of life is showing up.” In other words, being present is an extremely important part of being successful in life. For example, one cannot hit a home run without first standing in the batter’s box. Which of the following most weakens the movie director’s quote?

a. It works best in cities.

b. It implies that showing up regularly and on time is easy to do.

c. It overstates the importance of attendance relative to other factors.

d. It discourages actors.

29. A study of parenting styles found an important difference between the previous and current generation of parents. Today, parents are more likely to encourage their children to participate in a wide variety of hobbies and activities, whereas the previous generation of parents were more likely to encourage their children to concentrate on doing one thing really well, such as playing a musical instrument or playing a sport. Which of the following is the most logical conclusion based on this trend?

a. Previous generation parents undervalued the importance of being well rounded.

b. Fewer kids attend public school than was once true.

c. “Stick-to-it-iveness” is a rare quality in parents and children today.

d. Parents today have more children on average than the previous generation.

30. An old story is said to involve a man and the man’s wife who is dying and desperately needs medication. However, the medication is expensive, and the man has no way of paying for it or taking a loan. Therefore, the story presents the question of whether or not the man should steal the medication in order to save his wife. Which of the following is the most important question for determining whether stealing is okay in this example?

a. Has the man been rejected by society?

b. Is the man attractive?

c. What is the woman’s favorite food?

d. Can the man make amends (or make up for) for his crime?

APPENDIX B

50-item IPIP Trait Scales based on the FFM

On the following pages, there are phrases describing peoples' behaviors. Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your response will be kept in absolute confidence. Please read each statement carefully, and write your response in the space provided.

1	2	3	4	5	6	7
Very Inaccurate	**Moderately Inaccurate**	**Slightly Inaccurate**	**Neither Accurate nor Inaccurate**	**Slightly Accurate**	**Moderately Accurate**	**Very Accurate**

__________ 1. I am the life of the party.

__________ 2. I feel little concern for others.

__________ 3. I am always prepared.

__________ 4. I get stressed out easily.

__________ 5. I have a rich vocabulary.

__________ 6. I don't talk a lot.

__________ 7. I am interested in people.

__________ 8. I leave my belongings around.

__________ 9. I am relaxed most of the time.

__________ 10. I have difficulty understanding abstract ideas.

__________ 11. I feel comfortable around new people.

__________ 12. I insult people.

__________ 13. I pay attention to details.

__________ 14. I worry about things.

__________ 15. I have a vivid imagination.

continued…

1	2	3	4	5	6	7
Very Inaccurate	**Moderately Inaccurate**	**Slightly Inaccurate**	**Neither Accurate nor Inaccurate**	**Slightly Accurate**	**Moderately Accurate**	**Very Accurate**

__________ 16. I keep in the background.

__________ 17. I sympathize with others' feelings.

__________ 18. I make a mess of things.

__________ 19. I seldom feel blue.

__________ 20. I am not interested in abstract ideas.

__________ 21. I start conversations.

__________ 22. I am not interested in other people's problems.

__________ 23. I get chores done right away.

__________ 24. I am easily disturbed.

__________ 25. I have excellent ideas.

__________ 26. I have little to say.

__________ 27. I have a soft heart.

__________ 28. I often forget to put things back in their proper place.

__________ 29. I get upset easily.

__________ 30. I do not have a good imagination.

__________ 31. I talk to a lot of different people at parties.

__________ 32. I am not really interested in others.

__________ 33. I like order.

__________ 34. I change my mood a lot.

__________ 35. I am quick to understand things.

continued…

1	2	3	4	5	6	7
Very Inaccurate	**Moderately Inaccurate**	**Slightly Inaccurate**	**Neither Accurate nor Inaccurate**	**Slightly Accurate**	**Moderately Accurate**	**Very Accurate**

__________ 36. I don’t like to draw attention to myself.

__________ 37. I take time out for others.

__________ 38. I shirk my duties.

__________ 39. I have frequent mood swings.

__________ 40. I use difficult words.

__________ 41. I don’t mind being the center of attention.

__________ 42. I feel others’ emotions.

__________ 43. I follow a schedule.

__________ 44. I get irritated easily.

__________ 45. I spend time reflecting on things.

__________ 46. I am quiet around strangers.

__________ 47. I make people feel at ease.

__________ 48. I am exacting in my work.

__________ 49. I often feel blue.

__________ 50. I am full of ideas.

APPENDIX C

Biodata Measure

Instructions – READ CAREFULLY

Please answer the following questions by writing in the appropriate response on the blank provided or by circling the appropriate answer when necessary.

1. For how long have you worked in your current job? (Write in the number of years and months – for example, 1 year and 3 months etc.)

______ year(s) and ______ month(s)

2. For how long did you work in your most recent job prior to the one you have now? (Write in the number of years and months – for example, 1 year and 3 months etc.)

______ year(s) and ______ month(s)

3. How many jobs have you held in the past five years? (Write in the number on the blank)

____

4. How many times have you changed majors since you started college? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: _________________

5. How many academic institutions (including colleges, community colleges, technical colleges, and universities) have you attended on a **full-time basis** since you graduated high school including Wayne State University? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: _________________

6. How many academic institutions (including colleges, community colleges, technical colleges, and universities) have you attended on a **part-time basis** (for example, taken a summer course at) since you graduated high school including Wayne State University? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: _________________

7. How many times have you dropped a course **during high school**? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: _________________

8. How many times have you dropped a course **during college**? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

9. How many friends do you have in the introductory psychology course that you are taking this semester? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

10. How many relatives (for example, a cousin, brother, etc.) do you have in the introductory psychology course that you are taking this semester? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

11. About how often have you changed your mind about future career plans since the time you entered **high school**? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

12. About how many times have you been scolded or disciplined in previous jobs for showing up late or missing a scheduled work shift? (If more than 5, please write in the number on the blank)

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

13. Since the time that you started high school, about how many times have you been scolded or disciplined for showing up late or missing a class? (If more than 5, please write in the number on the blank).

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5
g. More than 5. Please write in the number: ________________

14. Approximately how far from WSU's campus do you live currently? (Please write in the **number of miles** or if you live on campus, write "campus")

____

15. How do you usually get to campus?

a. I drive myself in a car
b. I ride in a friend's or family member's car
c. I walk
d. I ride my bike
e. I use public transportation

16. How would you describe your commute to Wayne State University to a friend who was considering living where you currently live?

a. It is very easy.
b. It is moderately easy.
c. It is somewhat difficult.
d. It is very difficult.

APPENDIX D

Background Questionnaire

Instructions – READ CAREFULLY

Please use the blanks below to write your age in years and months.

Age: _____ Years _____ Months

Instructions – READ CAREFULLY

Please indicate your gender and ethnicity below by placing a check mark in the appropriate space.

Gender: _____ Female _____ Male

Ethnicity:

_____ Asian, Asian American, or Oriental

_____ Arab American

_____ Black or African American

_____ Hispanic or Latino

_____ American Indian

_____ White, Caucasian, European, not Hispanic

_____ Mixed; parents are from two different groups

_____ Other (write in on blank): ___________________

Instructions – READ CAREFULLY

Please answer the following questions by circling the appropriate answer and when necessary writing in the appropriate response on the blank provided.

1. My major at Wayne State University is:

    a. Psychology
    b. Undecided
    c. Other. Please write in your Major on the blank:
    ____

2. According to Wayne State University, I am currently considered a:

    a. Freshman – 1st year
    b. Sophomore – 2nd year
    c. Junior – 3rd year
    d. Senior – 4th year
    e. Senior – 5th year

3. I am currently enrolled in ______________ credit hours. Please write in the number on the blank provided.

4. Not including courses I am taking this semester, I have completed _____________ credit hours to date. Please write in the number on the blank provided.

5. My current status as a student at Wayne State University is:

    a. part-time – taking fewer than 12 credits this semester
    b. full-time – taking at least 12 credits this semester

6. What was your level of interest in Psychology prior to enrolling in this introductory psychology course?

a. very strong
b. strong
c. moderately strong
d. not strong – I had no interest in psychology
e. not strong – I was unsure what my interest in psychology would be

7. How frequently do you expect to **miss** introductory psychology class this semester (to miss a lecture or lab meeting)?

a. very frequently
b. somewhat frequently
c. rarely
d. never
e. not sure

8. How frequently do you expect to be **late** (to arrive after the class start time) to introductory psychology class this semester?

a. very frequently
b. somewhat frequently
c. rarely
d. never
e. not sure

9. How likely are you to **drop** this introductory psychology course this semester?

a. very likely
b. somewhat likely
c. not likely
d. extremely unlikely
e. not sure

10. What is the final grade that you expect to receive in this introductory psychology course?

a. A
b. B
c. C
d. D or lower
e. not sure

11. On the days that introductory psychology **lectures** are scheduled, do you have prior classes during the day?

a. Yes
b. No – Introductory Psychology is my first class of the day

12. On the days that your introductory psychology **laboratory** session is scheduled, do you have prior classes during the day?

a. Yes
b. No – Introductory Psychology Laboratory is my first class of the day

13. What is your most recent score on the ACT or SAT?

My ACT score was (possible scores from 1 to 36): _________________

My SAT score was (possible scores from 600 to 2400): _________________

14. How did you pay for the credit hours to take Introductory Psychology this semester? Or if you haven't paid yet, how do you plan to pay for these credit hours?

a. On my own.
b. Family support.
c. A student or personal loan.
d. A scholarship or university grant.
e. Other (Please explain): ______________________________

APPENDIX E

Instruction Set for Testing Sessions

1. Ask students to be seated with at least one open desk between them and the next closest student.

2. Once students are seated, read the following instructions:

   Before we begin the study, I'm going to ask you to take care of two things for me. First, if you have a cell phone, please turn it off at this time and leave it off for the duration of the study. Second, you'll need a pen or a pencil. If you do not have one with you, please raise your hand and I'll bring one to you. Next, I am going to pass out the informed consent form for this study. The informed consent is going to provide you with some basic information about the study, as well as information about any risks and benefits to you as a participant in the study. Please take a few minutes to read through the information that is provided, and when you are comfortable to proceed, please sign your name on the third page. If you have any questions, please raise your hand, and I will come over to assist you.

3. Collect informed consent and distribute CRT-W. Then read the following instructions:

   Please write your full name, access ID (for example: ab9999) and your 9-digit WSU student ID number on the blanks provided on the first page of the test booklet. It is very important that you write these in clearly so that I know who to assign the research credit to. When you are finished, please look up. [Pause and wait for students to finish].

   Ok, go ahead and flip over to page 2 at this time. I'm going to read through the instructions on this with you. The questions on the following pages are a test of your reasoning ability. Please read each question carefully, and respond by identifying the one answer that is the most logical. In other words, you must choose only one answer that is the best solution to each problem. Mark your response by circling the entire answer, so that it is clear which answer you have chosen. Please take a moment now to read the example item that is shown below, where the correct answer has been circled for you.

   Are there any questions before we proceed? [Pause for student questions]. I have just a few more instructions, and then we will begin. First, I want to make you aware of an opportunity to win a $20.00 gift card. Fifty - that's 5-0 - gift cards will be given to the students who achieve the highest score on the reasoning test. So, you will want to make sure to do your best. Incorrect guesses will not count against you; therefore, you should mark a response for every item on the test. If you are unsure which answer is correct, choose the one that you think is the best solution. The research assistants present will be wandering around the room to make sure that everyone is completely the test fairly.

   This test is timed. There are 30 items total and you have exactly 30 minutes to work on the test. Many of you will be done well before the 30 minutes is up. When you finish, please raise your hand, and an assistant will come by and give you the materials that you will need to complete for the next part of the study. Those materials will have instructions for how to proceed. Ok, you may begin the reasoning test now. I'm going to press start on the timer now. I will announce when you have 15-min, 10-min, and 5-min remaining.

4. Collect the CRT-W when participants indicate they have finished. Briefly check to make sure that all items are complete, and then write the current time on the top page of the test booklet. After collecting each test booklet, distribute the second packet of materials containing the remaining measures for the study (i.e., personality scales, biodata measure, demographics and background questionnaires).

APPENDIX F

HIC Approval Letter

WAYNE STATE UNIVERSITY

HUMAN INVESTIGATION COMMITTEE
87 East Canfield, Second Floor
Detroit, Michigan 48201
Phone: (313) 577-1628
FAX: (313) 993-7122
http://hic.wayne.edu

NOTICE OF EXPEDITED APPROVAL

To: Levi Nieminen
Psychology
5057 Woodward 7th Floor

From: Ellen Barton, Ph.D. E. Barton, PhD / ea
Chairperson, Behavioral Institutional Review Board (B3)

Date: August 31, 2010

RE: HIC #: 087310B3E

Protocol Title: A Study of Reasoning Ability

Funding Source: Unit: Psychology

Protocol #: 1008008715

Expiration Date: August 30, 2011

Risk Level / Category: Research not involving greater than minimal risk

---

The above-referenced protocol and items listed below (if applicable) were **APPROVED** following *Expedited Review* (Category 7*) by the Chairperson/designee *for* the Wayne State University Behavioral Institutional Review Board (B3) for the period of 08/31/2010 through 08/30/2011. This approval does not replace any departmental or other approvals that may be required.

- Recruitment Script
- Debriefing Email
- SONA Systems Advertisement
- Consent Form (dated 8/31/10)

- Federal regulations require that all research be reviewed at least annually. You *may* receive a *"Continuation Renewal Reminder"* approximately two months prior to the expiration date; however, it is the Principal Investigator's responsibility to obtain review and continued approval ***before*** the expiration date. Data collected during a period of lapsed approval is unapproved research and can ***never*** be reported or published as research data.
- All changes or amendments to the above-referenced protocol require review and approval by the HIC **BEFORE** implementation.
- Adverse Reactions/Unexpected Events (AR/UE) must be submitted on the appropriate form within the timeframe specified in the HIC Policy (http://www.hic.wayne.edu/hicpol.html).

NOTE:

1. Upon notification of an impending regulatory site visit, hold notification, and/or external audit the HIC office must be contacted immediately.
2. Forms should be downloaded from the HIC website at ***each*** use.

*Based on the Expedited Review List, revised November 1998

REFERENCES

Abelson M. A. (1993) Turnover cultures. *Research in Personnel and Human Resources Management, 11*, 339-376.

Adler, S., & Golan, J. (1981). Lateness as a withdrawal behavior. *Journal of Applied Psychology,* 66, 544-554.

Barrick, M. R., Mount, M. K., & Strauss, J. P. (1994). Antecedents of involuntary turnover due to a reduction in force. *Personnel Psychology, 47,* 515-535.

Barrick, M. R., & Zimmerman, R. D. (2005). Reducing voluntary, avoidable turnover through selection. *Journal of Applied Psychology, 90*, 159–166.

Barrick, M. R., & Zimmerman, R. D. (2009). Hiring for retention and performance. *Human Resource Management, 48*, 183-206.

Bernardin, J. H. (1977). The relationship of personality variables to organizational withdrawal. Personnel Psychology, 30, 17-27.

Berry, C. M., Sackett, P. R., Tobares, V. (2010). A meta-analysis of conditional reasoning tests of aggression. *Personnel Psychology, 63*, 361-384.

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

Bersoff, D. M. (1999).Why good people sometimes do bad things: Motivated reasoning and unethical behavior. *Society for Personality and Social Psychology*, *25*, 28-39.

Bing, M. N., LeBreton, J. M., Davison, H. K., Migetz, D. Z., & James, L. R. (2007). Integrating Implicit and Explicit Social Cognitions for Enhanced Personality Assessment: A General Framework for Choosing Measurement and Statistical Methods. *Organizational Research Methods, 10*, 136-179.

Blau, G. (1994). Developing and testing a taxonomy of lateness behavior. *Journal of Applied Psychology, 79*, 959-970.

Blau, G. (1998). On the aggregation of individual withdrawal behaviors into larger multi-item constructs. *Journal of Organizational Behavior, 19*, 437-451.

Blau, G. (2002). New conceptualizations of lateness since Blau, 1994. In Koslowsky, M. & Krausz, M. (Eds.), *Voluntary Employee Withdrawal and Inattendance*. New York: Kluwer Academic/Plenum Publishers.

Boudreau, J. W, & Berger, C. J. (1985). Decision-theoretical utility analysis applied to employee separations and acquisitions. *Journal of Applied Psychology, 70,* 581-612.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics, 30*, 88-99.

Campbell, D. T. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47*, 15-38.

Campion, M. A. (1991). Meaning and measurement of turnover: Comparison of alternative measures and recommendations for research. *Journal of Applied Psychology, 76*, 199-212.

Cascio, W. (1987). *Costing human resources: The financial impact of behavior in organizations* (2nd ed.). Boston: Kent.

Caza, B. B., & Cortina, L. M. (2007). From insult to injury: Explaining the impact of incivility. *Basic and Applied Social Psychology, 29*, 335-350.

Chadwick-Jones, J. K., Brown, C. A., Nicholson, N., & Sheppard, C. (1971). Absence measures: Their reliability and stability in an industrial setting. *Personnel Psychology, 24*, 463-470.

Clegg, C. W. (1983). Psychology of employee lateness, absence, and turnover: A methodological critique and an empirical study. *Journal of Applied Psychology, 68*, 88-101.

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

Cooper, C., & Dewe, P. (2008). Well-being: Absenteeism, presenteeism, costs, and challenges. *Occupational Medicine, 58*, 522–524.

Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D. (2001). Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology, 6*, 64-80.

Dalton, D. R., Krackhardt, D. M., & Porter, L. W. (1981). Functional turnover: An empirical assessment. *Journal of Applied Psychology,* 66.716-721.

Dalton, D. R., & Mesch, D. (1991). On the extent and reduction of avoidable absenteeism: An assessment of absence policy provisions. *Journal of Applied Psychology, 76*, 810-817.

Dalton, D. R., & Todor, W D. (1979). Turnover turned over: An expanded and positive perspective. *Academy of Management Review, 4,* 225-235.

Dalton, D. R., & Todor, W D. (1982). Turnover: A lucrative hard dollar phenomenon. *Academy of Management Review, 7,* 212-218.

Dalton, D. R., & Todor, W. D. (1993). Turnover, transfer, absenteeism: An interdependent perspective. *Journal of Management,* 19:193-219.

Dalton, D. R, Todor, W D, & Krackhardt, D. M. (1982). Turnover overstated: The functional taxonomy. *Academy of Management Review,* 117-123.

Deery, M. A., & Shaw, R. N. (1997). An exploratory analysis of turnover culture in the hotel industry in Australia. *International Journal of Hospitality Management, 16*, 375-392.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association, 83*, 414-425.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology, 58*(2), 203-210.

Fisher, C. (1980). On the dubious wisdom of expecting job satisfaction to correlate with performance*. Academy of Management Review, 5*, 607-612.

Fried, Y., & Ferris, G. R. (1987). The validity of the job characteristics model: A review and meta-analysis. *Personnel Psychology, 40*, 287-322.

Froggatt, P. (1970). Short-term absence from industry. III. The inference of 'proneness' and a search for causes. *British Journal of Industrial Medicine, 27*, 297-312.

Frost, B. C., Ko, C. H. E., & James, L. R. (2007). Implicit and explicit personality: A test of a channeling hypothesis for aggressive behavior. *Journal of Applied Psychology, 92*, 1299-1319.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C., Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Goodman, P. S., & Atkin, R. S. (1984). Effects of absenteeism on individuals and organizations. In P. S. Goodman & R. S. Atkin (Eds.), *Absenteeism: New approaches to understanding, measuring, and managing employee absence* (pp. 276-321). San Francisco: Jossey-Bass.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*, 449-510.

Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine, *The handbook of research synthesis and meta-analysis* (2nd ed.), pp. 417-433. New York: Russell Sage Foundation.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Greenwald, A. G., Bellezza, F. S., & Banaji, M. R. (1988). Is self-esteem a central ingredient of the self-concept? *Personality and Social Psychology Bulletin*, *14*, 34-45.

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management, 26*, 463-488.

Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley.

Hackett, R. D., & Bycio, P. (1996). An evaluation of employee absenteeism as a coping mechanism among hospital nurses. *Journal of Occupational & Organizational Psychology,* 69, 327-338.

Hammer, T. H., Landau, J. C., & Stern, R. N. (1981). Absenteeism when workers have a voice: The case of employee ownership. *Journal of Applied Psychology, 66*, 561-573.

Hanisch, K. A. (1995). Behavioral families and multiple causes: Matching the complexity of responses to the complexity of antecedents. *Current Directions in Psychological Science, 4*, 156-162.

Hanisch, K. A., & Hulin, C. L. (1990). Job attitudes and organizational withdrawal: An examination of retirement and other voluntary withdrawal behaviors. *Journal of Vocational Behavior, 37*, 60-78.

Hanisch, K. A., & Hulin, C. L. (1991). General attitudes and organizational withdrawal: An evaluation of a causal model. *Journal of Vocational Behavior, 39*, 110-128.

Hanisch, K. A., Hulin, C. L., & Roznowski, M. (1998). The importance of individuals' repertoires of behaviors: The scientific appropriateness of studying multiple behaviors and general attitudes. *Journal of Organizational Behavior, 19*, 463-480.

Harrison, D. A. (2002). Meaning and measurement of work role withdrawal: Current controversies and future fallout from changing information technology. In Koslowsky, M. & Krausz, M. (Eds.), *Voluntary Employee Withdrawal and Inattendance*. New York: Kluwer Academic/Plenum Publishers.

Harrison, D. A., & Martocchio, J. J. (1998). Time for absenteeism: A 20-year review of origins, offshoots, and outcomes. *Journal of Management, 24*, 305-350.

Harrison, D. A., & Price, K. H. (2003). Context and consistency in absenteeism: studying social and dispositional influences across multiple settings. *Human Resource Management Review, 13*, 203-225.

Harrison, D. A., & Shaffer, M. A. (1994). Comparative examinations of self-reports and perceived absenteeism norms: Wading through Lake Wobegon. *Journal of Applied Psychology, 79*, 240-251.

Hausknecht, J. P., Trevor, C. O., Howard, M. J. (2009). Unit-level voluntary turnover rates and customer service quality: Implications of group cohesiveness, newcomer concentration, and size. *Journal of Applied Psychology, 94*, 1068-1075.

Hayes, L. J., O'Brien-Pallas, L., Duffield, C., Shamian, J., Buchan, J., Hughes, F., et al. (2006). Nurse turnover: A literature review. *International Journal of Nursing Studies, 43*, 237-263.

Herzberg, F., Maunser, B., Peterson, R., & Capwell, D. (1957). *Job attitudes: Review of research and opinion*. Pittsburgh: Psychological Services.

Hill, J. M., & Trist, E. L. (1955). Changes in accidents and other absences with length of service: A further study of their incidence and relation to each other in an iron and steel works. *Human Relations, 8,* 121-152.

Hoffman, B. J., & Woehr, D. J. (2006). A quantitative review of the relationship between person-organization fit and behavioral outcomes. *Journal of Vocational Behavior, 68*, 389-399.

Hom, P. (1992). *Turnover costs among mental health professionals*. College of Business, Arizona State University, Tempe, Ariz.

Hulin, C. L., Roznowski, M., & Hachiya, D. (1985). Alternative opportunities and withdrawal decisions: Empirical and theoretical discrepancies and an integration. *Psychological Bulletin, 97*, 233-250.

Huseman, R. C., Hatfield, J. D., & Miles, E. W. (1985). Test for individual perceptions of job equity: Some preliminary findings. *Perceptual and Motor Skills, 61*, 1055-1064.

Huseman, R. C., Hatfield, J. D., & Miles, E. W. (1987). A new perspective on equity theory: The equity sensitivity construct. *Academy of Management Review, 12*, 222-234.

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences. Retrieved on April 20, 2010 from: http://ipip.ori.org/

Iverson, R. D. (1996). Employee acceptance of organizational change: The role of organizational commitment. *International Journal of Human Resource Management,* 7, 121-148.

Iverson, R. D. & Deery, M. (1997). A causal model of employee intent to leave: The role of turnover culture in the hospitality industry. *Human Resource Management Journal,* 7, 71-82.

Iverson, R. D., & Deery, S. J. (2001). Understanding the “personological” basis of employee withdrawal: The influence of affective disposition on employee tardiness, early departure, and absenteeism. *Journal of Applied Psychology, 86*, 856-866.

James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods, 1*, 131-163.

James, L. R., & LeBreton, J. M. (2011). *Assessing the Implicit Personality Through Conditional Reasoning*. Washington, D.C.: American Psychological Association.

James, L. R., & Mazzerolle, M. D. (2002). *Personality in Work Organizations*. Thousand Oaks, CA: Sage.

James, L. R., McIntyre, M. D., Glisson, C. A., Bowler, J. L., & Mitchell, T. R. (2004). The Conditional Reasoning Measurement System for aggression: An overview. *Human Performance, 17*, 271-295.

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., et al. (2005). A Conditional Reasoning Measure for Aggression. *Organizational Research Methods, 8*, 69-99.

Johns, G. (1994a). Absenteeism estimates by employees and managers: Divergent perspectives and self-serving perceptions. *Journal of Applied Psychology, 79*, 229-239.

Johns, G. (1994b). How often are you absent? A review of the use of self-reported absence data. *Journal of Applied Psychology, 79*, 574-591.

Johns, G. (1998). Aggregation or aggravation? The relative merits of a broad withdrawal construct. *Journal of Organizational Behavior, 19*, 453-462.

Judge, T. A., Martocchio, J. J., & Thoresen, C. J. (1997). Five-factor model of personality and employee absence. Journal of Applied Psychology, 82, 745–755.

Kacmar, K. M., Andrews, M. C., Van Rooy, D. L., Steilberg, R. C., & Cerrone, S. (2006). Sure, everyone can be replaced . . . but at what cost? Turnover as a predictor of unit-level performance. *Academy of Management Journal, 49,* 133–144.

Koslowsky, M., & Dishon-Berkovits, M. (2001). Self-report measures of employee lateness: Conceptual and methodological issues. *European Journal of Work and Organizational Psychology, 10*, 145-159.

Koslowsky, M., Sagie, A., Krausz, M., & Singer, A. D. (1997). Correlates of employee lateness: Some theoretical considerations. *Journal of Applied Psychology, 82*, 79-88.

Krackhardt, D., & Porter, L. W. (1985). When friends leave: A structural analysis of the relationship between turnover and stayers' attitudes. *Administrative Science Quarterly, 30*(2), 242-261.

Krackhardt, D., & Porter, L. W. (1986). The snowball effect: Turnover embedded in communication networks. *Journal of Applied Psychology, 71*(1), 50-55.

Krausz, M. (2002). The many faces of voluntary employee turnover: A multifacet and multilevel perspective. In Koslowsky, M. & Krausz, M. (Eds.), *Voluntary Employee Withdrawal and Inattendance*. New York: Kluwer Academic/Plenum Publishers.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

LeBreton, J. M., Barksdale, C. D., Robin, J., & James, L. R. (2007). Measurement issues associated with conditional reasoning tests: Indirect measurement and test faking. *Journal of Applied Psychology, 92*, 1-16.

Lee, T. L., & Mitchell, T. R. (1994). An alternative approach: The unfolding model of voluntary employee turnover. *Academy of Management Review, 19*, 51-89.

Lenth, R. V. (2006-9). Java Applets for Power and Sample Size [Computer software]. Retrieved July 28, 2010 from http://www.stat.uiowa.edu/~rlenth/Power.

Maertz, C. P., Jr., & Campion, M. A. (2004). Profiles in Quitting: Integrating Process and Content Turnover Theory. *Academy of Management Journal, 47*, 566-582.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.

Martocchio, J. J. and Harrison, D. A. (1993). To be there or not to be there? Questions, theories, and methods in absenteeism research. *Research in Personnel and Human Resources Management, 11*, 259-328.

Meyer, J. P., Bobocel, D. R., & Allen, N. J. (1991). Development of organizational commitment during the first year of employment: A longitudinal study of pre- and post-entry influences. *Journal of Management, 17*, 717-733.

Miles, E. W., Hatfield, J. D., & Huseman, R. C. (1989). The Equity Sensitivity construct: Potential implications for worker performance. *Journal of Management, 15*, 581-588.

Miles, E. W., Hatfield, J. D., & Huseman, R. C. (1994). Equity sensitivity and outcome importance. *Journal of Organizational Behavior, 15*, 585-596.

Mirvis, P, M., & Lawler, E. E. Measuring the financial impact of employee attitudes. *Journal of Applied Psychology,* 1977, *62,* 1-8.

Mitra, A., Jenkins, G. D., & Gupta, N. (1992). A meta-analytic review of the relationship between absence and turnover. *Journal of Applied Psychology, 77*(6), 879-889.

Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology, 62*, 237-240.

Mobley, W H. (1982). Some unanswered questions in turnover and withdrawal research. A*cademy of Management Review, 7,*111-116.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683-729.

Morita, J. G., Lee, T. W., & Mowday, R. T. (1989). Introducing survival analysis to organizational researchers: A selected application to turnover research. *Journal of Applied Psychology, 74*, 280-292.

Morse, B. J., & Popovich, P. M. (2009). Realistic recruitment practices in organizations: The potential benefits of generalized expectancy calibration. *Human Resource Management Review, 19*, 1-8.

Muchinsky, P. M. (1977). Employee absenteeism: A review of the literature. *Journal of Vocational Behavior, 10*, 316-340.

Muchinsky, P. M., & Morrow, P. C. (1980). A multidisciplinary model of voluntary employee turnover. *Journal of Vocational Behavior, 17*, 263-290.

Mueller, C. W., Price, J. L. (1989). Some consequences of turnover: A work unit analysis. *Human Relations, 42*, 389-402.

Munson, H. L., & Rubenstein, B. J. (1992). School IS work: Work task learning in the classroom. *Journal of Career Development, 18*, 289-297.

Murphy, K. R. (1983). Fooling yourself with cross-validation: Single sample designs. *Personnel Psychology, 36*, 111-118.

Neuman, J. H., & Baron, R. A. (1998). Workplace violence and workplace aggression: Evidence concerning specific forms, potential causes, and preferred targets. *Journal of Management*, *24*, 391–419.

van der Net, J. B., Janssens, A. C., Eijkemans, M. J., Kastelein, J. J., Sijbrands, E. J., & Steyerberg, E. W. (2008). Cox proportional hazards models have more statistical power than logistic regression models in cross-sectional genetic association studies. *European Journal of Human Genetics, 16*, 1111-1116.

Nicholson, N. (1977). Absence behavior and attendance motivation. A conceptual synthesis. *Journal of Management Studies, 14:231-252.*

Nicholson, N., & Johns, G. (1985). The absence culture and the psychological contract: Who's in control of absence? *Academy of Management Review, 10*, 397-407.

Payne, R., & Nicholson, N. (1987). Absence from work: Explanations and attributions. *Applied Psychology: An International Review, 36*, 121-132.

Pelz, C. J., & Klein, J. P. (n.d.). Analysis of survival data: A comparison of three major statistical packages (SAS, SPSS, BMDP). Unpublished manuscript retrieved from: http://www.mcw.edu/FileLibrary/Groups/Biostatistics/TechReports/tr017.pdf

Porter, L. W., & Steers, R. M. (1973). Organizational work and personal factors in employee turnover and absenteeism. *Psychological Bulletin, 80*, 151-176.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660-679.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2003). Personality and absenteeism: A meta-analysis of integrity tests. *European Journal of Personality, 17*, 19-38.

O'Reilly, C. A., & Caldwell, D. F. (1981). The commitment and job tenure of new employees: Some evidence of postdecisional justification. *Administrative Science Quarterly, 26*, 597-616.

Rosse, J. G. (1988). Relations among lateness, absence, and turnover: Is there a progression of withdrawal? *Human Relations,* 41(7): 517-531.

Rotundo, M., & Spector, P. E. (2010). Counterproductive work behavior and withdrawal. In Farr, J. L., & Tippins, N. T. (Eds). *Handbook of Employee Selection*. Taylor & Francis Group: New York.

Sackett, P. R., & DeVore, C. J. (2002). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.). *Handbook of industrial, work and organizational psychology, Volume 1: Personnel Psychology* (pp. 145-164). Thousand Oaks, CA: Sage.

Saige, A., Birati, A., & Tziner, A. (2002). Assessing the costs of behavioral and psychological withdrawal: A new model and an empirical illustration. *Applied Psychology: An International Review, 51*, 67-89.

Sagie, A., Koslowsky, M., & Hamburger, Y. A. (2002). Antecedents of employee lateness: A multiple-level model. In Koslowsky, M. & Krausz, M. (Eds.), *Voluntary Employee Withdrawal and Inattendance*. New York: Kluwer Academic/Plenum Publishers.

Salgado, J. F. (2002). The Big-Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment, 10*, 117-125.

Schmitt, N., Cortina, J. M., Ingerick, M. J., Wiechmann, D. (2003). Personnel selection and employee performance. In Borman, W. C., Ilgen, D. R., & Klimoski, R. J.

(Eds), *Handbook of Psychology, Vol. 12 Industrial and Organizational Psychology*. John Wiley & Sons, Inc. Hoboken: New Jersey.

Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437-453.

Schneider, B., Goldstein, H. W., & Smith, D. B. (1995). The ASA framework: An update. *Personnel Psychology*, *48*, 747-773.

Schotzhauer, D. L., & Rosse, J. G., (1985). A five-year study of a positive incentive absence control program. *Personnel Psychology, 38*, 575-585.

Silvera, D. H., & Seger, C. R. (2004). Feeling Good about Ourselves: Unrealistic Self-Evaluations and Their Relation to Self-Esteem in the United States and Norway. *Journal of Cross-Cultural Psychology, 35*, 571-585.

Singer, J. D., & Willett, J. B. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110*, 268-290.

Somers, M. J. (1996). Modelling employee withdrawal behaviour over time: A study of turnover using survival analysis. *Journal of Occupational and Organizational Psychology, 69*, 315-326.

Somers, M. J., & Birnbaum, D. (1999). Survival versus traditional methodologies for studying employee turnover: Differences, divergences and directions for future research. *Journal of Organizational Behavior, 20*, 273-284.

Spector, P. E., Suzy Fox, S., Penney, L. M., Bruursema, K., Goh, A., Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior, 68*, 446–460.

Staw, B. M. (1980). The consequences of turnover. *Journal of Occupational Behavior, 1*, 253-273.

Staw, B. M., & Oldham, G. R. (1978). Reconsidering our dependent variables: A critique and empirical study. *Academy of Management Journal, 21:* 539-555.

Steers, R. M. and Mowday, R. T. (1980). Employee Turnover and post-decision accomodation processes. In: Cummings, L. L. and Staw, B. M. (Eds), *Research in Organizational Behavior*. JAI Press: Greenwich, Connecticut.

Steers, R. M., & Rhodes, S. R. (1978). Major influences on employee attendance: A process model. *Journal of Applied Psychology, 63*, 391-407.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*, 5th ed. Boston, MA: Allyn and Bacon.

Taylor, P. J. (1968). Personal factors associated with sickness absence: A study of 194 men with contrasting sickness absence experience in a refinery population. *British Journal of Industrial Medicine, 25*, 106-118.

Totterdell, P., Spelten, E., Smith, L., Barton, J., & Folkard, S. (1995). Recovery from work shifts: How long does it take? *Journal of Applied Psychology,* 80, 43-57.

Wanous, J. P. (1973). Effects of a realistic job preview on job acceptance, job attitudes, and job survival. *Journal of Applied Psychology, 58*, 327−332.

Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin, 124*, 333-371.

ABSTRACT

THE DEVELOPMENT AND VALIDATION OF A CONDITIONAL REASONING TEST OF WITHDRAWAL

by

LEVI R. G. NIEMINEN

May 2012

Advisor: Dr. Sebastiano Fisicaro

Major: Psychology (Industrial / Organizational)

Degree: Doctor of Philosophy

This study developed and evaluated a measure of implicit dispositional tendencies associated with lateness, absenteeism, and permanent withdrawal behaviors. The conditional reasoning framework developed by Lawrence James and colleagues was adopted. Novel cognitive biases or *justification mechanisms* associated with withdrawal were proposed, drawing on research and theory from the attribution (*marginalization of withdrawal*), commitment (*revocable commitment*), and fairness/equity (*social injustice bias*) domains. As part of the empirical validation design, college students enrolled in an Introductory Psychology course completed the conditional reasoning measure, and corresponding behavioral withdrawal criteria were collected unobtrusively throughout the 16-week course. Results of scale development analyses pointed to a subset of items (13 of 25) with positive evidence of predictive validity and indicated that the items assess largely heterogeneous content, possessing low internal consistency. Results of hypothesis testing revealed positive and statistically significant predictive relationships for revocable commitment and social injustice (i.e., higher scores on the conditional reasoning items associated with higher frequencies of

the withdrawal behaviors), but not for marginalization of withdrawal. Taken together, these results provide initial evidence for the role of implicit dispositional tendencies in the withdrawal process and underscore potential avenues for further development of a conditional reasoning test of withdrawal.

AUTOBIOGRAPHICAL STATEMENT

Education

2012: Ph.D. in Psychology (Industrial / Organizational), Wayne State University, Detroit, MI
2009: M.A. in Psychology (Industrial / Organizational), Wayne State University, Detroit, MI
2004: B.S. in Psychology, Lake Superior State University, Sault Ste. Marie, MI

Professional Experience

2010-present: *Denison Consulting, LLC* – Research consultant.
2010: *Polaris Assessment Systems* – Intern.
2006-2009: *Applied Psychology Organizational Research Group* – Junior consultant.

Publications

Nieminen, L. R. G. & Denison, D. (In Press). Organizational culture and effectiveness. In E. Kessler (Ed.), *Encyclopedia of Management Theory*.

Nieminen, L. R. G., Rudolph, C. W., Casper, C. M., Wynne, K. T., Kirby, L. C., & Baltes, B. B. (In Press). The combined effect of ratee bodyweight and past performance information on performance judgments. *Journal of Applied Social Psychology*.

Dickson, M. W., Nieminen, L. R. G., & Biermeier-Hanson, B. (2011). Nepotism and organizational homogeneity: How the ASA process is accelerated by non-merit-based decision-making. In Jones, R. G. (Ed.), *Nepotism in Organizations*, pp. 93-128. New York: Routledge Academic.

Schmitt, N. W., Arnold, J. D., Nieminen, L. R. G. (2010). Validation strategies for primary studies. In Farr, J.L., Tippins, N.T. (Eds.), *Handbook of Employee Selection*, pp. 51-71. New York: Psychology Press.

Nieminen, L. R. G., Nicklin, J. M., McClure, T. K., & Chakrabarti, M. (2010). Meta-analytic decisions and reliability: A serendipitous case of three independent telecommuting meta-analyses. *Journal of Business and Psychology, 26,* 105-121.