

CORRELATES OF A PAST BEHAVIOR INTERVIEW FOR THE BUSINESS

UNIT LEADER: EXPERIENCE, MOTIVATION, PERSONALITY,

AND COGNITIVE ABILITY

Lane A. Conner, M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2008

APPROVED:

Charles A. Guarnaccia, Major Professor

C. Edward Watkins, Committee Member

Paul Lambert, Committee Member

Linda Marshall, Chair of the Department of
Psychology

Sandra L. Terrell, Dean of the Robert B. Toulouse
School of Graduate Studies

UMI Number: 3385777

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3385777

Copyright 2009 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Conner, Lane A. *Correlates of a Past Behavior Interview for the Business Unit Leader: Experience, Motivation, Personality and Cognitive Ability*. Doctor of Philosophy (Industrial/Organizational Psychology), August 2008, 103 pp., 14 tables, references, 56 titles.

This research evaluates the relationship between various individual differences constructs and performance on a past behavior interview (PBI)—one of the most popular forms of personnel selection interviews used today—within a sample of business unit leader level incumbents and applicants from organizations across the United States. Correlation analysis is conducted on the relationship between overall performance on a PBI and four work-related constructs: Experience, Motivation, Personality, and Cognitive Ability. The existing literature on PBIs and the four independent variables is critically reviewed. As limited research has been conducted on the influence of Experience and Motivation on PBI performance, this study makes unique contributions to the literature regarding impact of these two constructs. The major hypotheses stated that Experience and Motivation would yield significant, positive correlations with PBI performance while Personality and Cognitive Ability would not be significantly correlated with PBIs. Results partially supported the hypotheses—Experience, Motivation, and Personality were significantly related to overall PBI score, while Cognitive Ability was not. Implications for the findings as well as suggestions for future research are discussed.

Copyright 2008

by

Lane A. Conner

ACKNOWLEDGEMENTS

This work is the culmination of a lifetime of formal education. There more people that have been involved in the pursuit of this goal than could ever be acknowledged; that being said, those who have had the most significant and profound impact must be acknowledged. First and foremost, I must acknowledge my mother. She has given more support, time, money, love, and sanity than I will ever begin to understand. To my parents, this outcome is the result of a carefully planned existence, beginning literally before I was born. Strategic decisions were made in every aspect of my life, from being born in the fall to a baseball being placed in my left hand and a pencil in my right, to the hand-selection of every teacher I've ever had. For this, I am eternally in your debt and only hope the pride you feel in this achievement can begin to repay you for the infinite blessings you have bestowed upon me. I must also acknowledge The Buddies, one of my best friends and editor Zach Sherman, and my classmate Kevin Impelman for his friendship and infinite statistical wisdom and guidance. Do you like this gig? That wasn't the question—do you like your job? Finally I must acknowledge the consistent support of Michael Beyerlein, Charles Guarnaccia, Linda Marshall, Amy Mayfield, and the rest of the psychology department at the University of North Texas for their support of an extinct program and its few remaining specimens.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	v
Chapters	
1. INTRODUCTION	1
History and Structure	
Dependent Variable—Past Behavior Interview	
Experience and the Interview	
Motivation and the Interview	
Personality and the Interview	
Cognitive Ability and the Interview	
Current Study	
2. METHOD	51
Participants	
Measures	
Procedure	
Analysis	
3. RESULTS	67
Experience and Past Behavior Interview Performance	
Motivation and Past Behavior Interview Performance	
Personality and Past Behavior Interview Performance	
Cognitive Ability and Past Behavior Interview Performance	
4. DISCUSSION	75
Interpretations	
Limitations	
Future Research	
APPENDIX: SAMPLE QUESTIONS.....	97
REFERENCES	100

LIST OF TABLES

	Page
1. Overview of Campion, Palmer, & Campion's (1997) Fifteen Elements of Interview Structure.....	83
2. Assessment Methods' Ability to Measure Elements of Performance by Determinants Model and Type of Performance Demonstrated.....	85
3. Super-factors and Subscales of the Leadership Experience Inventory (LEI).....	85
4. Super-factors and Subscales of the Global Personality Inventory (GPI).....	86
5. Global Personality Inventory Subscales Restructured along the Big Five Personality Factors, with Original GPI Super-factors	87
6. 12 Competencies Measured in the Past Behavior Interview and Their Definitions	88
7. Descriptive Statistics for the Past Behavior Interview Competencies and Overall Rating	89
8. Descriptive Statistics for the Super-factors and Subscales of the Leadership Experience Inventory (LEI).....	90
9. LEI Super-factors and Overall LEI Score Correlations with PBI Overall Average	92
10. Descriptive Statistics for the Super-factors and Subscales of the Global Personality Inventory (GPI).....	93
11. Overall Motivation and Motivation Subscales Correlation with PBI Overall Average	95
12. Hierarchical Regression of Motivation Subscales and 7 Remaining Subscales under the Big Five Factor Extraversion	95
13. GPI Super-factors Correlation with PBI Overall Average	96
14. Big Five Factor Correlations with PBI Performance from Current Study, Conway and Peneno (1999) and Huffcutt et al. (2001)	96

CHAPTER 1

INTRODUCTION

History and Structure

The selection interview has been one of, if not the, most prominent method of personnel selection over the past century and it continues to play a substantial role in a majority of hiring decisions made today. Research on the topic began to appear early in the 20th century; the first scholarly journal article on the interview as it relates to the selection of salesmen was published by Scott in 1915. Even then the selection interview was not just a research consideration, but was being used to make personnel decisions—and being used often. A survey of 236 companies in 1930 found that fully 93% of those firms reported conducting an interview on applicants before making a hiring decision (Spriegel & James 1958). These researchers replicated the study again some 30 years later, and by 1957 99% of 852 firms surveyed reported interviewing applicants before hiring (Spriegel & James, 1958).

The trend of increased use of the interview for personnel selection continued through the second half of the 20th century. In a 1965 review of the existing literature on selection interviews Ulrich and Trumbo (1965) declared, “apparently the interview is used almost universally as one of the sources of information on which personnel decisions of hiring, placement, and, in all likelihood, transfer and promotion are made,” (page 235). Subsequent research continued pointing to the popularity of the selection interview; Guion’s (1976) research on the extent of use of various personnel selection methods led him to conclude that the interview was the most widely used employment technique in America. Similarly, research found the interview to be the most popular hiring method in the United Kingdom as well (Robertson & Makin, 1986). Despite years of considerable difficulty in scientifically justifying its consistent use (which will be

discussed later), today the employment interview is as popular as ever and an integral part of selection practices around the globe.

As the popularity of the interview for use in personnel selection began to grow in the early part of the 20th century, so too did the amount of scientific research published on this topic. As an example of the growth in research published on the selection interview during this time, from 1915 to 1965 (50 years) 134 studies were published on the selection interview; from 1965 to 1989 (24 years) approximately 260 studies were published. The problem was that although the sheer amount of literature being published on the topic was increasing, the reliability and validity estimates of the interview were not following suit.

From the earliest published articles on the selection interview, reliability and validity were well below acceptable levels. For example, the personnel selection interview developed for the Scott (1915) article possessed obvious problems. I found that of 36 interviews for sales positions conducted by a panel of 6 interviewers, the panel disagreed on something as simple as whether the applicant should be placed in the top- or bottom-half of the candidate pool in 28 of the 36 (nearly 80%) of the interviews. Another early example of problems with the interview as a selection device comes from Hollingworth (1922). In this study, twelve different sales managers “experienced with personnel selection” interviewed 57 different candidates for a sales position and were asked to rank-order the candidates. Overall, interviewer ranking of applicants consistently varied wildly. In fact, one candidate was ranked 1st by one interviewer and 57th—absolutely last—by another.

The fact is, despite the increase in volume, research on the selection interview continued to yield unimpressive results until the 1980’s. Although a detailed, comprehensive discussion of the voluminous research finding poor validity and reliability data on the selection interview since

these early examples is well beyond the scope of this discussion, the findings of the major literature reviews published since then makes it apparent the poor statistical results continued to appear over the following decades. For example, the first narrative review of the existing literature on the selection interview was conducted by Wagner (1949). In his review of 104 research articles published on the selection interview, he identified 27 studies reporting validity coefficients ranging from .09 to .94 and 43 studies reporting reliability data with reliability coefficients ranging from -.20 to .97. Based on widely varied and inconsistent findings as these, Wagner concluded that the interview should be limited in its use to those factors which cannot be more accurately measured by other means. He quickly followed this statement by noting that many of those dimensions which were once only measured by the interview were being increasingly—and more accurately—evaluated instead by other methods such as standardized tests.

Mayfield (1964) and Ulrich and Trumbo (1965) published the next substantial reviews of the selection interview literature in subsequent years. While these articles varied to some degree in the studies reviewed, the general conclusions reached, and recommendations for future research, one issue on which they wholeheartedly agreed was the lack of reliability and validity found for the selection interview. Mayfield (1964) concluded that “knowledge of the selection interview is only a little more advanced” than it was at the time Wagner conducted his study and that the interview is of “dubious” value (page 245). One of Ulrich and Trumbo’s (1965) major positions was supporting Wagner’s (1949) position that due the inability of the selection interview to accurately assess individual attributes, ancillary data (that is, additional information about the interviewee such as work samples and standardized tests) should be taken into consideration and “weighed” when making decisions about interview performance as often as

possible in order to help raise the validity of the selection interview. Similarly, in Schmitt's (1976) review of the literature published since the last review in 1969, the author concludes "there is not much in the research of the last half-dozen years to bolster the confidence of a personnel interviewer concerned with the reliability and validity of his decisions" (p. 97).

Over the course of nearly a century, multiple factors influenced the evolution of the interview into a selection instrument which yields validity and reliability data comparable to that of long accepted assessment and selection tools such as cognitive ability measures and work sample tests. That being said, the change in the selection interview that played the biggest role in the evolution to its current state was researchers' gradual identification of the importance of conducting interviews that were structured. Although evidence of the importance of interview structure was available quite early in the research, as will be revealed, the understanding of the essential role of interview structure in creating valid, reliable interviews was a gradual process that emerged over time and became apparent as research began to identify the numerous confounding variables that inhibited interview validity and reliability. Research also began to shed light upon the variables that yielded the lack of statistical support for the use of the personnel interview for selection purposes.

Although the influence of structure on the interview did not lead to meaningful statistical improvements of the validity and reliability of the selection interview in the published research until the 1980s, the importance and usefulness of structure in the interview was documented by researchers as early as the 1930's. In 1937, Wonderlic emphatically stated that the only way to obtain useful information from interviews was by standardizing the entire process. Wonderlic's assertions were not just speculation; they were based on scientific data. For example, in Hovland and Wonderlic (1939) published an article on the reliability and validity and reliability of a

standardized interview they developed with an accompanying rating guide (one of the earliest behaviorally anchored rating scales, or BARS; discussed below) called the Diagnostic Interviewer's Guide. Although no predictive validity was reported, the authors found a corrected even-odd reliability of .82. Another study which found statistical support for the application of structure to interviews that also failed to attract much attention was McMurray's (1947) publication of what he termed the "patterned" interview. A precursor of the modern structured interviews used today, McMurray's patterned interview consistently asked the same questions of candidates and their answers were objectively scored according to a scoring guide he had developed. This interview format yielded a predictive validity of .68 between interview score and job performance ratings in an automobile factory. Despite occasional publications and findings such as these, a vast majority of the interviews conducted during this time in both research and applied settings were unstructured interviews.

Little had changed by the time Wagner published the first narrative review of the existing selection interview literature in 1949 mentioned earlier. Wagner noted that as of 1949, "no standardized procedure or pattern for conducting an interview has generally been adopted" (page 33). It is my opinion this inaction (i.e., lack of adopting a standardized procedure for conducting interviews) was a function of two factors, the first of which being a lack of belief that an interaction as complex as the selection interview could be meaningfully controlled. For example, in summarizing the research conducted to that point, Wagner concluded that "the validity and reliability of the interview [was] highly specific to both the situation and to the interviewer" (page 43). From my perspective, this statement exemplifies Wagner's belief that too many extraneous variables were involved in the interview to maintain any meaningful control on the interview—situational variables such as: type of position being interviewed for, amount of

additional information available about the candidate, and number/type of questions asked. Other extraneous variables Wagner believed could not be controlled originated from the individual interviewer, including: ability to establish rapport, ability to ask job-related questions, and ability to decipher honest from dishonest answers—it was Wagner’s view that nothing could be done to systematically keep the numerous situational and the interviewer variables in check so as to yield meaningful data.

The second factor that prevented a standardized procedure for conducting interviews from being established was simply a general lack of “know-how”. Even if an individual had the desire, generally speaking there was not sufficient knowledge of scientific rigor to control an interaction as dynamic as an employment interview (those few examples of the successful structured interview discussed above—which were not widely distributed or replicated—were the rare exception).

The next significant development in understanding the importance of structure in the interview came from two key publications on the selection interview were released in 1964. Mayfield’s (1964) review mentioned earlier proposed two new directions of research which he believed would be beneficial. First he recommended that interview research should change its focus to studying the decision making processes inherent in the interview and work to uncover the factors influencing interviewer judgments. Additionally, Mayfield (1964) proposed a shift to a “microanalytic approach” to interview research. Specifically, this would involve dividing the interview into smaller subunits and conducting controlled studies one or two units at a time. Thus, Mayfield’s approach can be seen as a potential solution to Wagner’s (1949) stated concern about the too large number of factors to potentially be aware of and control in such dynamic human interactions as selection interviews.

Ironically, that same year, Webster (1964) published a book on nine years of research conducted by him and his colleagues. Just as Mayfield (1964) had requested, Webster's (1964) book reported a series of microanalytic studies focusing on factors that influence the decision-making of the interviewer. The actual findings of Webster's (1964) work made significant contributions to the literature on the selection interview; but as Wright (1969) notes, the greater significance of Webster's work came in terms of the subsequent research it fostered. Wright (1969) surmises, "It would be difficult to over-estimate the importance of the work done by Webster and his colleagues," (page 394) based on the amount of replication research as well as new research avenues opened by Webster.

On the surface, neither Mayfield nor Webster's work may seem to have an impact on the development of structure in the interview. However, closer examination of their results makes it apparent that this literature was strongly pointing to the need for increased control in the employment interview. For example, a few of Webster's (1964) key findings were: interviewers developed a stereotype of a good candidate and sought to match interviewees with the stereotype; interviewers establish biases early on in interviews and those biases tend to reflect in interview decisions; and an interviewers' decision is different when information is fed piece-by-piece as opposed to simultaneously. What these microanalytic studies of the interviewer decision-making process—and many others like it—were bringing to light was the fact there were numerous confounding variables from a variety of sources that directly influenced the interviewer's ratings and thus they were a major cause of the poor quality data consistently appearing in the research. As the identification of variables confounding the utility of the interview steadily increased, it gradually became apparent that these variables were in large part responsible for the low observed validity and reliability statistics, and that structure needed to be

introduced in the interview to control the influence of these variables.

Not only did Webster's (1964) work directly provide examples of multiple variables that interfered with interviewers' accurate ratings of performance, but as Schmitt (1976) notes, each of Webster's "significant findings has lead to a series of subsequent research" (page 82). With this, it is evident that Webster's (1964) research was especially important to the acknowledgment of the necessity of structure in the interview not only for the numerous variables he identified that influenced interviewer ratings but also for the considerable research this inspired addressing the same topic.

Campion, Palmer, and Campion (1997) note that every major review published since Wagner (1949) has encouraged the use of the structure in the interview. However, it was not until after the aforementioned realization that there were numerous confounding variables influencing interviewer decision-making which had to be minimized via the micro-analytic studies of the 1960s and 1970s and the publication of formalized "how-to" articles on two types of standardized interviews—patterned behavior description interviews (PBDIs—a precursor to the Past Behavior Interview; Janz, 1982) and situational interviews (SIs; Latham, Saari, Pursell, and Campion, 1980)—that researchers began to consistently apply structure to the interviews being developed and make progress in terms of establishing credible reliability and validity figures. These two types of interviews which, as will be discussed, have become the most popular methods of interviewing used today, were similar in a number of ways regarding the elements of structure emphasized and built into their development process (discussed below). The main difference between the interviews centered on the time frame of the interview question; specifically, one type of interview question (the SI) focused on future behavior and the other (the PBDI) focused on past behavior. The situational interview was founded on Locke's

(1968) goal-setting theory which states that a good indicator of what people are going to do is what they say they are going to do; therefore, situational-type interview questions ask about intended future behavior in a hypothetical situation similar to one the interviewee might actually face on the job. For example, a question might begin, “What would you do if you encountered...” in an effort to predict how the interviewee might actually perform in the future. On the other hand, the patterned behavior description interview (PBDI) is based upon the behavior consistency principle, which states that “the best predictor of future behavior is past behavior,” (Janz, Hellervik, and Gilmore, 1986; page 32). To be clear, the terms “patterned behavior description interview” (PBDI) and “past behavior interview” (PBI) are to be considered synonymous for reasons that will be described momentarily. In an attempt to infer actual future job behaviors/performance, patterned behavior description (or past behavior) questions ask participants about specific situations encountered in the past which required them to demonstrate behaviors relevant to successful performance on the job being applied for. A patterned behavior interview question might start out “Tell me about a time when you...”

These two seminal articles brought together various methods of standardizing the interview which were not altogether new or original, but discussed and organized them in a straightforward way with clear examples that future researchers could use as guides. For example, both Janz (1982) and Latham et al. (1980) made clear the essential need for developing the interview based upon a critical incident job analysis (described momentarily) to ensure the interview questions are based on, and specifically measure, overt employee behavior. Another example of an element of structure in the interview specifically emphasized and demonstrated by Janz (1982) was the necessity of ensuring the interviewer was properly trained to focus only on the actual behavioral responses of the interviewee when assigning ratings, not attending to

hunches or preliminary judgments. Additionally, Latham et al. (1980) re-emphasized the notion (originally suggested by Smith and Kendall, 1963) of basing the scoring guide, referred to as a behaviorally anchored rating scales (BARS) on the job analysis to ensure a valid, objective scoring key. BARS are quite similar to a 5-point Likert-type rating scale except that sample behavioral responses are developed from the critical incident job analysis for the lowest, middle, and highest ratings and included on the guide to serve as benchmarks, or “anchors” for identifying the score a particular response should receive.

While the publication of these two articles proved to be predominant forms of structured interviews used today, in my opinion, the gradual progress made over the years in understanding of the importance of structure in the interview culminated with the research published by Campion et al. in 1997. Campion et al. (1997) were finally able to organize the elements of interview structure in a comprehensive manner. In this study, the authors reviewed the existing literature on the selection interview and compiled a comprehensive a list of 15 elements of structure; they reported on how each element is operationalized in the literature and discuss each component’s potential impact on validity and reliability. The authors broadly defined structure as, “any enhancement of the interview that is intended to increase psychometric properties by increasing standardization or otherwise assisting the interviewer in determining what questions to ask or how to evaluate responses,” (page 656). The 15 elements of interview structure were classified into two general categories: structural elements that impact the content of the interview and those that impact the evaluation process. Table 1 lists the 15 elements of structure described by Campion et al. (1997) and makes a general comment about each.

Dependent Variable—Past Behavior Interviews

As mentioned, the Past Behavior Interview (PBI), initially termed the “patterned behavior description interview,” (PBDI) was developed by Tom Janz in 1982. Though not the first introduction of structure to the interview, the PBDI, along with Latham et al.’s (1980) situational interview (SI), seemed to be introduced just at right time, when a multitude of research was being published identifying the vast number of confounding variables influencing interviewer ratings and detracting from interview validity and reliability. The essential role structure played in the development and administration of valid and reliable interviews and, in turn, the ability of interviews to meaningfully predict future job performance had never been more apparent—and Janz’ PBDI provided a clear framework to instill that structure.

Before exploring the process involved in developing this type of interview, it is first beneficial to understand how the patterned behavior description interview (PBDI) has evolved since its introduction. Janz’ (1982) PBDI was designed with certain unique characteristics—some of which have been considered integral for attaining the specific, behaviorally focused information essential for accurately predicting future work performance. These elements, such as interview questions exclusively focusing on past experiences to make evaluations, and the development of competencies and interview questions according to a critical incident job analysis, have remained staples of the process over the years. On the other hand, certain components of Janz’s patterned behavior description interview—including its name—have gradually been phased out by researchers and replaced by techniques more capable of increasing the reliability and validity of the instrument.

Two main alterations to Janz’s (1982) original version of the PBDI have been made over the years, the most important being the introduction of behaviorally anchored rating scales

(BARS) for scoring the interviews, which has had a considerable impact on improving the validity of past-oriented interviews. As previously discussed, Latham et al.'s (1980) situational interview utilized the BARS for providing accurate, focused interview ratings; however, Janz' PBDI offered little assistance in terms of direction or structure regarding how to best score the PBDI; therefore, nearly all past behavior interviews today employ BARS.

The second noteworthy alteration to the original version of the PBDI, which, unlike the integration of behaviorally anchored rating scales has not been universally accepted, was first introduced by Motowidlo, Carter, Dunnette, Tippins, Werner, Burnett, and Vaughn (1992) and addressed the lack of standardization of questions in the PBDI. Janz termed his interviews "patterned" because the intent was to allow the interviewer to selectively choose from "patterns" of questions established for each dimension being assessed, depending on the particular interviewee and the direction of the interview. Then the interviewer was encouraged to probe and ask follow-up questions to gather the additional information necessary to rate the behavior dimension. Researchers such as Motowidlo et al. (1992) felt the high degree of interviewer discretion allotted by the PBDI could lead to inconsistent questioning/probing and therefore unreliable results. Motowidlo et al. (1992) developed a variation of the PBDI in which interviewees were still asked questions about past experiences and behaviors, but instead of interviewers being free to choose from a variety of possible questions, a standardized list of questions (and probes) were provided for the interviewers to ask each applicant. While some researchers today follow Motowidlo et al.'s (1992) example and ask standardized past-behavior focused questions, others (such as Day and Carroll, 2003) feel standardization of questions severely limits the amount of quality, job-relevant data that can be elicited.

The functional elements of Janz's (1982) patterned behavior description interview are not

the only aspects that have changed with time. For example, when Motowidlo et al. (1992) introduced their variation of Janz's PBDI, they termed the interview the "Structured Behavioral Interview." Numerous additional iterations of the PBDI with a variety of names have appeared since 1992. For example, Pulakos and Schmitt (1995) developed an interview asking past-behavior type questions for their study and labeled it the "experience based interview."

The publication of numerous additional versions of Janz's PBDI interview with subtle variations in the structural elements combined with the introduction of various terms describing these types of structured interviews such as "behavior description interviews" (BDIs) or "past behavior interviews" (PBIs) has led to the somewhat confusing current situation in which all of these terms are basically used interchangeably. However, working from two key thoughts on the topic, the present discussion proposes a solution to this confusion. First, Krajewski, Goffin, McCarthy, Rothstein, and Johnston's (2006) statement regarding current terminology of what once was known as the PBDI is insightful: "Janz labeled his interview format the 'patterned behavior description interview' (PBDI); however, as the norm has been in past research, we will use the label *PBI* to refer, more generally, to structured interviews using the past-behavior format," (page 413; italics added). In addition, it is my observation that the term "past behavior interview" seems to be the most straight-forward and comprehensive characterization of these types of structured interviews; therefore this paper calls for the universal adoption and exclusive use of the term "past behavior interview" or "PBI," to refer to any structured interview that asks questions oriented toward eliciting participants' descriptions of past behavior in situations.

As is apparent from the multiple examples of structured interviews discussed above, structured interviews can vary in any number of ways. That being said, the key issue in differentiating one structured interview from the next—especially for researchers attempting to

generalize interview research findings by meta-analytically evaluating data—seems not to be the subtle differences in interview structure such as level of question standardization described above. Instead, the key issue is the more general factor of whether the interview asks future-focused, situational-type questions or past-oriented, past behavior-type questions. A data-driven (in terms of differential reliability, predictive validity, and even the constructs measured) as well as intuitive “fork in the road” has emerged, with the two clear paths of structured interviews being past behavior interviews and situational interviews. As will be discussed, while most of the modern structured interview research today focuses on this key element of structure to differentiate structured interviews, much of the early research on structured interviews failed to emphasize this differentiating factor.

Developing a past behavior interview (PBI; as well as a situational interview for that matter), to be clear, is a time-consuming, detailed process. To provide a general understanding of the steps involved in developing a PBI, an overview of the process as detailed by the Janz, Hellervik, and Gilmore (1986) comprehensive text on the past behavior interview will be provided. The first step in developing a PBI, or any structured interview, begins with a thorough critical incident job analysis of the position in order to identify critical job activities. In general, a job analysis is a systematic process that begins with identification of most, if not all, components involved in the job and eventually narrows the components down to those most essential for job success. According to Janz et al.’s (1986), critical incidents—specific descriptions of real behaviors (both effective and ineffective) that have actually occurred on the job as reported from a variety of job experts—are the key to the job analysis process.

According to Janz et al. (1986), the key to this step is recording the critical incidents in a specific fashion. Well-written critical incidents help to ensure a smooth question development

process. When recording the critical incidents for the job at hand, incidents should be written in such a way that they accurately describe, rather than evaluate behavior; additionally each incident should include the outcome or result of actions taken. Once a sufficient number of critical incidents have been gathered—the specific number depends on the amount of previous accurate information available about the job, such as prior job analyses—the next step is the formation of performance dimensions. This process involves the systematic reduction of the critical incidents into 5 to 10 roughly organized groups involving generally similar behaviors.

When this process is complete, actual item writing can begin. The first step Janz et al. (1986) describe is deciding whether or not an interview question should be written for a specific performance dimension at all. This decision is made by considering which of two different types of job performance the dimension best demonstrates. The authors explain that there are two types of job performance – maximum performance and typical performance. Maximum performance refers to instances where the individual must display the highest degree of knowledge, skills or abilities he or she possesses, such as producing as many widgets as possible in a set amount of time. On the other hand, typical performance describes the average output one demonstrates under normal circumstances on a day-in, day-out basis. For example, assessing an individual's ability to collaborate with co-workers is a measure of typical performance because it is based on how the individual displays that characteristic on a daily basis. As it relates to the PBI, if the dimension under consideration is more commonly going to be displayed at its maximum level, then Janz et al. (1986) advise that typically a standardized assessment such as a cognitive ability test or job knowledge test, which by nature requires the participant to demonstrate his or her maximum performance, will be a better measure for that particular dimension. On the other hand, if the dimension is in most cases going to be displayed at its typical performance level,

then the past behavior interview is appropriate, and the process of developing a question for that dimension can begin. (Familiarity with the dynamic of maximum performance and typical performance will serve the reader well, as this concept will receive further discussion later.)

Actually writing the question begins with composing question stems based on the critical incidents within the particular performance dimension. The stem is the part of the interview question that leads the applicant to consider certain past situations and focus on the one that matches the stem. The next step is the development of the probes. Probes are the questions about the situation and the applicant's specific behavioral response to that situation. The goal of a probe is to elicit precise descriptions of behavior and the results of that behavior. Although the above description is an overly simplistic portrayal of the process, most past behavior interviews, including the one used as the dependent variable in this study and others reviewed throughout this discussion, have been developed in a manner very similar to this.

In addition to evaluating the method in which an assessment technique such as the past behavior interview is developed, reviewing the psychometric properties of the instrument is another important factor in gaining insight into the instrument. While the focus of this study is, in fact, just the past behavior interview, often a discussion of the validity and reliability data on the PBI cannot be separated from a discussion of the validity and reliability evidence of the situational interview (SI) as well. As the two emerged around the same time in the early 1980s, the body of research examining the SI and the PBI progressed along a similar route as well. First, early research simply attempted to establish credible validity and reliability figures for each (Orpen, 1985; Janz, 1989; Latham & Saari, 1984; Weekley & Gier, 1987). Gradually though, as these structured interviews consistently demonstrated respectable psychometric properties, the research emphasis shifted from establishing validity to comparing validity of the PBI and the SI.

The first study to compare the predictive validities of the two types of structured interviews was conducted by Campion, Campion and Hudson in 1994. In a study of 70 pulp mill employees, the PBI correlated .51 with job performance while the SI only correlated .39, although the difference was not statistically significant. In the years following Campion et al.'s (1994) study, enough research comparing the PBI and SI (e.g., Conway & Peneno, 1999; Huffcutt, Weekley, Wiesner, DeGroot, and Jones, 2001; Latham & Skarlicki, 1995, Pulakos & Schmitt, 1995) was published that in 2002 Taylor and Small were able to aggregate the existing data and conduct meta-analyses on the criterion-related validities and inter-rater reliabilities of both types of interviews in order to paint a more complete picture of the psychometric properties of the PBI and the SI.

Results of the analysis showed that PBIs, when scored via behaviorally anchored rating scales (as mentioned, some of the early PBIs were not scored using BARS; because their use increases predictive validity considerably, this factor was controlled for) yielded significantly higher criterion-related validities than SIs scored with BARS. The estimated population mean for PBIs scored with BARS was .63, while SIs scored using BARS was .47. Regarding inter-rater reliabilities, SIs and PBIs yielded similar sample-weighted mean inter-rater reliabilities when use of BARS for interview scoring was being controlled for (.79 and .77 respectively; Taylor and Small, 2002). Two additional studies, Day and Carroll (2003) and Krajewski et al. (2006) have been published since the Taylor and Small (2002) meta-analysis. Day and Carroll (2003) found the SI and PBI predicted job performance at nearly identical levels in a sample of undergraduates, as they correlated .37 and .36 respectively ($p < .01$). More recently, Krajewski et al. (2006) found results considerably different than Day and Carroll (2003) when comparing the criterion-related validities of PBIs and SIs in a sample of applicants to managerial positions.

Similar to Taylor and Small's (2002) findings, in Krajewski et al.'s (2006) study, PBIs significantly predicted job performance ratings ($r = .32, p < .01$), while SIs did not ($r = .09, ns$).

Numerous constructs have been proposed as being predictors and correlations of Past Behavior Interview success, from well-established individual difference variables such as cognitive ability (e.g., Huffcutt, Roth, & McDaniel, 1996) and personality (Huffcutt et al. 2001), to more obscure and difficult-to-define variables such as tacit knowledge (Conway & Peneno, 1999) and person-organization fit (Harris, 1989). Specifically, this current study will examine the relationship between PBIs and an individual's work experience, motivation, cognitive ability and personality. Each of these constructs and its relationship with the interview will be discussed in turn.

Experience and the Interview

The relationship between interviews and the amount of work experience a person possesses has not received much attention until fairly recently. McDaniel, Whetzel, Schmidt & Maurer (1994) conducted a meta-analysis of the 10 studies available at the time that evaluated the correlation between interview performance and tenure (a common measure of job experience). The corrected estimated population mean of those studies for this relationship was .20. The problem with this meta-analysis—which will become a recurring theme amongst much of the initial structured interview research and was touched on in the discussion of how question type (past- or future-focused) is the key differentiator between the two types of structured interviews—was that it failed to clearly differentiate between interviews which asked situational/future-focused questions and interviews which posed past behavior-focused questions in the meta-analysis. The authors lumped the findings of research that used SIs together with the

findings of research that used PBIs in their analysis. Therefore, the findings of this meta-analysis and many studies like it that explore the impact of individual differences constructs on the interview but only differentiate between “structured” and “unstructured” interviews are limited in their ability to provide insight about the two separate, highly structured interviews most common today—the SI and the PBI—between which meaningful differences have been established.

The first study to examine the impact of an applicant’s amount of work experience on the interview and also differentiate between situational interviews and past behavior interviews was Pulakos and Schmitt (1995). This research, designed to compare the general psychometric properties of the PBI and the SI, reported on two separate studies utilizing employees in a large federal organization; Study I consisted of 216 participants and Study II involved 464 participants. A between-subjects design was used to compare the characteristics of the SI and the PBI; that is, in each study the sample was split into two independent experimental groups, one of which was administered a SI, while the other group took a PBI (each interview was designed to measure the exact same dimensions).

It should be noted that the between-subjects design structure utilized by Pulakos and Schmitt (1995) is arguably superior to the within-subjects design used by other studies comparing the SI and PBI where each participant is asked both situational- and past behavior-type questions in the same interview. This is the case because the within-subjects design eliminates potential influence of order effects—artifacts that could confound study results by causing participants to respond differentially to one type of question specifically because it consistently appears before or after the other type of question. Additionally, results from between-subjects design studies are potentially more meaningful in “real world,” applied settings

where interviews typically do not combine question types. Finally, assessments developed for use with between-subjects design are more appropriate for meta-analyses in that they are a true measure of the instrument and will provide data more representative of the measure.

Regarding the between-subjects design of the Pulakos and Schmitt (1995) study, half of the participants – 108 in Study I and 232 in Study II – were randomly assigned to take the PBI, the other half were administered the SI (only the results of the research related to the PBI will be discussed here). The 16-question PBI was developed and administered in accordance with the procedures generally used for maximizing the predictive validity of a structured interview. First, the questions were developed based on a thorough critical incident job analysis. Furthermore, the PBI was administered in a standardized fashion by a panel of 3 interviewers each separately scoring the interaction according to a 7-point behaviorally anchored rating scale (BARS) with behavioral samples for each possible level of effectiveness (it should be noted that the same anchored rating scale was used for both the SI and the PBI). Interviewers discussed the interview afterward to reach consensus within one point. The interviewers came from a pool of 72 supervisors who received a one-day training on administering and scoring the interviews. Scores on a cognitive ability measure were also collected in Study II (discussed later). In both studies, job experience was measured by years of experience (the participants varied from 1 to 6 years); no significant relationship between experience and score on the PBI was found in either instance (Study I, $r = .11$ *ns*; Study II $r = -.01$, *ns*).

A potential limitation of this study's exploration of the relationship between experience and Past Behavior Interview performance exists in that the sample only included participants ranging from one to six years of experience. If a wider range of experience was included in the

sample—from new employees to those with 10 or 15 years of experience, it seems more likely that effects of experience on PBI performance would have been observed.

Another limitation exists in that the same BARS was used to score both the situational interview and the past behavior interview. While the authors did this to ensure that any differences in validities between the two interview types were due to interview format and not rating scale differences, this methodology could potentially impact the ability of the rating scale to fully capture differences in responses between more experienced and less experienced interviewees. Specifically, it is plausible that due to the fact that example responses (the anchors) had to be appropriate for both situational- and past behavior-type questions, the response anchors might have been more generic in wording and might not have been to accurately differentiate between subtle, yet meaningful differences in responses of more experienced participants—that had a real impact on outcomes—from similar responses of less experienced participants without the subtle actions. Although Pulakos and Schmitt's (1995) study was rigorously developed and utilized a sound research design, these potentially significant factors influence the analysis of the experience/PBI relationship; therefore these results must be interpreted with caution.

Like Pulakos & Schmitt (1995), Huffcutt et al.'s (2001) comparison of the psychometric properties of situational interviews and past behavior interviews also pointed to a lack of a meaningful relationship between work experience and PBI scores. Unlike Pulakos and Schmitt (1995), however, Huffcutt et al. (2001) chose a within-subjects design to compare the two interview types. Thus, of 93 district managers for a national chain of stores, each participated in a phone interview that contained both situational and past behavior questions. However, in order to minimize potential order effects resulting from using the within-subjects methodology, PBI questions were asked first exactly half of the time, while SI questions were asked first during the

other half of interviews. Each question was developed according to a critical incident job analysis. The interviewers were trained to administer and score both types of interview using a 3-point scale to rate responses for that particular question. Job experience was measured in terms of tenure—number of years in the current position. Participants' age also was noted and correlated with PBI score. The authors reasoned that age was a viable measure of “general life experience” that, like job experience, could potentially impact PBI performance. Additionally, each participant in this study was administered a cognitive ability measure and a Big Five personality scale (discussed later). Results indicated no significant correlation between experience and PBI score ($r = .02$). Interestingly, a significant negative correlation was found between age and PBI performance ($r = -.26, p < .05$). The authors hypothesized this result was either due to an age-related bias inherent in the test content or that older participants were in general poorer performers who had not been promoted to higher positions.

Although Huffcutt et al.'s (2001) results, like Pulakos and Schmitt (1995), imply that no significant relationship exists between experience and PBIs, limitations in the research design lead one to interpret the findings on the experience/PBI relationship with caution in this study as well. One potential limitation of this study was that the behaviorally anchored rating scale only differentiated between three different levels of question response quality (i.e., only “low”, “medium”, and “high” rating options). A more thorough anchored rating scale differentiates between at least five different levels of responses and provides behavioral examples for each. Just as Pulakos and Schmitt's (1995) use of a generic BARS to score both the SI and the PBI could fail to differentiate between subtle differences in behavior responses of those with more experience from those with less experience, so too could a BARS with limited options for rating quality of responses. Additionally, although it is unclear how it would impact the findings

regarding the relationship between amount of work experience and PBI scores, a general limitation of the study acknowledged by the authors was the fact that the interviews were conducted over the phone.

While evaluating and/or comparing the psychometric properties of situational interviews and past behavior interviews, the research discussed above found no significant work experience/PBI correlation. However, to be sure, there is an equal amount of research similarly exploring the properties of SIs and PBIs that has found a meaningful, positive relationship between work experience and performance on PBIs. For example, in an attempt to compare the construct validity and applicant reactions resulting from three different interview types (SIs, PBIs, and another type of interview), Conway and Peneno (1999) collected data on 179 applicants for a Resident Assistant (RA) position on a university campus over a three-year time period. A standardized, two-round selection interview process was developed. A brief screening interview narrowed the sample to 137 applicants who participated in the more advanced second-round structured interview. This second-round interview consisted of both situational and past behavior interview questions (7 of each type of question; a within-subjects study design). The questions, developed from a critical incident job analysis, were administered and scored by thoroughly trained interviewers using a five-point rating scale with benchmark sample answers provided for each question. In order to explore potential sources of construct validity of the three types of interviews, in addition to past leadership experience, data on personality and cognitive ability was collected as well (results of which are discussed later). Again, as PBIs are the focus of the current study, only findings related to this type of interview will be discussed. In this study, leadership experience was operationalized as number of leadership roles in extracurricular activities both on- and off-campus. It should be noted that of the three years this interview data

was collected, applicant experience data was only available from one of these years, thus reducing the actual sample size available for analysis of this particular construct to 48. Analysis revealed a significant positive correlation ($r = .43, p < .05$) between past leadership experience and PBI performance.

Two minor qualifications of the experience/PBI relationship data in this study should be noted. First, while an N of 48 is acceptable, it is by no means ideal; typically a higher N is preferred to allow for greater confidence in results. Also, while the skills learned from off- and on-campus leadership roles could reasonably be expected to impact performance for the job of R.A., such an operationalization of experience may make it difficult to generalize these findings to applied business world settings where work leadership might involve different skill sets.

Day and Carroll (2003) also found evidence that work experience has a significant impact on PBI outcomes. Like Conway and Peneno (1999), the authors set out to further explicate differences in the construct and criterion validity of past behavior interviews and situational interviews (SIs), as well as differences in perceptions of fairness for each. This experiment used a between-subjects design to compare the two types of structured interviews; therefore half of the 120 undergraduate participants took the PBI and half took the SI. The authors developed a four-question PBI and a four-question SI to decide admittance to a hypothetical academic program. The interviews and accompanying behaviorally anchored rating scales (with behavioral examples for each of 5 levels of response quality), were developed according to job-analytic techniques. Under laboratory conditions, each interview was administered by the same 2-person team of interviewers, both trained on administering and scoring the PBI and the SI. In addition to experience, cognitive ability and prior knowledge of questions before entering the interview

were introduced as independent variables in order to observe their effect on PBI scores (discussed later).

Individual experience was operationalized in terms of academic experience and quantified by number of years of experience in the participant's particular university program. Results of the statistical analysis showed a statistically significant positive correlation between individual level of experience and score on the PBI ($r = .33, p < .01$). No obvious problems or limitations in study design could be identified that significantly impact the interpretation of these findings regarding the relationship between experience and PBIs.

On the surface, studies on the relationship between work experience and past behavior interviews performance yield mixed results. However, as discussed later, when a more detailed investigation is conducted, this data seems to point generally to the existence of a relationship between these two variables such that those who possess more work experience tend to perform better on PBIs.

Motivation and the Interview

Of the constructs evaluated in this study, individual motivation and its relationship to interview performance has received the least amount of attention from researchers. In fact, despite my extensive efforts, no prior quantitative research could be located directly testing the potential link between motivation and interview ratings. The meta-analysis of the relationship of the SI and PBI with experience by Taylor and Small (2002) was the only study that even addressed the topic and proposed a hypothesis about the strength and direction of the relationship between motivation and the PBI based on some indirect statistical evidence. This study, along with research from related areas in the selection interview literature which provides some

potential insight about the nature of the relationship between motivation and past behavior interviews will be discussed as well.

Taylor and Small's (2002) hypothesis that motivation plays a role in determining an individual's performance on the past behavior interview integrates information from two concepts concerning what is typically the main criterion variable in personnel selection—future job performance. First, Taylor and Small (2002) describe the three specific elements that determine job performance—declarative knowledge, procedural knowledge and skills, and motivation—from McCloy, Campbell, and Cudeck's (1994) job performance determinants model. Second, Taylor and Small (2002) integrate a concept introduced earlier in this paper—the two types of performance observable in an applicant or employee—maximum performance and typical performance. These two concepts, as well as how Taylor and Small (2002) employ them to link motivation to past behavior interviews will be discussed.

As mentioned, McCloy, et al. (1994) hypothesized that three (and only three) elements comprise an individual's job performance: declarative knowledge, procedural knowledge and skills, and motivation. Declarative knowledge refers to “the ability to state or *declare* the facts, rules, principles, or procedures that are a prerequisite for successful task performance,” (italics added; page 494). Procedural knowledge and skills, on the other hand, is defined as, “the capability attained when Declarative Knowledge (knowing what to do) has been successfully combined with knowing how and [actually] being able to perform a task,” (page 495). A practical example helps clarify the two—while an individual who has never been in an airplane may read all the manuals and memorize the steps involved in landing a passenger jet and may be able to recite those steps with ease (demonstrating declarative knowledge), actually sitting behind the controls and landing the plane (demonstrating procedural knowledge and skills)

would obviously be considerably different. Declarative knowledge, then, is a necessary but insufficient condition for procedural knowledge and skills (Taylor and Small, 2002). McCloy et al.'s (1994) model of the determinants of job performance is as follows:

$$\text{Job Performance} = f(\text{DK, PKS, M}),$$

where DK is Declarative Knowledge, PKS is Procedural Knowledge and Skills, and M is Motivation.

The authors step away from this concept momentarily to further discuss the two types of job performance: maximum performance and typical performance. As discussed, maximum job performance is the resulting output when an individual is motivated to perform at the highest level of his or her capability. Typical job performance, on the other hand is the average productivity level that one displays under normal conditions—one's day-to-day job performance. Much of what Taylor and Small (2002) cover regarding maximum and typical performance was covered in the review of Janz et al. (1986) process of developing a PBI; however, as the topic is integral to understanding the motivation/past behavior interview relationship, the key components will be reviewed. Taylor and Small (2002) explain how some personnel evaluation measures (assessments) are measures of maximum performance and others are measures of typical performance. Measures of maximum performance are designed to elicit a demonstration of the highest level of performance possible; for example work sample tests are measures of maximum performance in that the participant attempts to produce the highest output he or she is capable of in the amount of time allotted. Whereas measures of typical performance, such as supervisor ratings, are so named because they portray how an individual performs day-in, day-out, on a typical basis.

The key difference, however, between measures of maximum and typical performance,

lies in differences in the ability to measure the three determinants of job performance described by McCloy et al. (1994) above. While some measures of maximum performance only effectively measure Declarative Knowledge (e.g., a job knowledge test only evaluates knowledge of what to do in situations on the job); other measures tap both Declarative and Procedural Knowledge and Skills (e.g., a work sample test requires the individual not only to possess knowledge of what to do—Declarative Knowledge, but he or she must also be able apply that knowledge in a simulated work scenario—Procedural Knowledge and Skills). None of the measures of maximum performance is able to accurately assess an individual’s general level of Motivation however, in that each measure’s design artificially sets motivation to “high,” as participants aim to perform to the best of their abilities during the allotted time Taylor and Small (2002). Measures of typical performance on the other hand are not only able to demonstrate declarative knowledge and procedural knowledge and skills, but they also demonstrate differences in individuals’ motivation to apply that declarative and procedural knowledge and skills on a daily basis. Table 2 lists the various personnel assessment measures, the elements of job performance measured, and whether the measure assesses maximum or typical performance.

According to Taylor and Small (2002) past behavior interviews—the type of interview utilized in this study—not only assess Declarative Knowledge and Procedural Knowledge and Skills, but PBIs are also able to tap Motivation, in that as the interviewee relays an example of how he or she responded effectively in the past to a certain situation, the individual not only demonstrates the requisite knowledge and skills but also that he or she was “sufficiently motivated to apply the knowledge/skills in that situation,” (page 280). Thus, Taylor and Small (2002) hypothesize that as PBIs are measures of typical performance which, unlike measures of

maximum performance, are also able to tap an individual's Motivation to perform, PBIs should be meaningfully correlated to the Motivation construct.

As mentioned, Taylor and Small (2002) also present indirect statistical support for their contention that motivation and PBI performance are related. One of the main goals of their research was to meta-analytically compare the predictive validities of the situational interview (SI) and the past behavior interview (PBI). It should be noted that situational interviews are considered measures of maximum performance in that it is assumed the interviewee will give the best possible answer to each hypothetical question. Therefore, motivation is not being assessed because there is no way to determine whether the interviewee would actually handle the situation in the way he or she indicated (Taylor and Small, 2002). Analyses revealed both SIs and PBIs successfully predicted job performance, but that PBIs produced a significantly higher correlation coefficient than the SIs (.56 to .45, respectively). Taylor and Small pointed to these findings as indirect statistical support for their contention that although both interview types assess declarative knowledge and procedural knowledge and skills, the significantly higher validity of PBIs suggests that PBIs also tap an interviewee's motivation due to their ability to assess typical performance.

Additional indirect evidence for the existence of a relationship between motivation and interview performance stems from research conducted on related aspects of the selection interview. The research discussed below addresses the effects of various methods of preparing for an interview ahead of time. Assuming that individuals who possess higher levels of motivation overall are more likely to make the extra effort to participate in some form of pre-interview preparation, it follows that they would also be more likely to exhibit the improved interview scores resulting from this preparation as described in the research below. It should be

noted that although no specific research could be found explicitly linking an individual's intrinsic motivation to likelihood of pre-interview preparation, it is not a great theoretical leap to hypothesize that highly motivated individuals in general would have a stronger desire to perform well in a selection interview and therefore be more likely to take action to prepare ahead of time.

For example, Caldwell and Burger (1998) found a link between conducting research about the company or job for which the applicant was interviewing ahead of time and interview performance. The authors explored whether preparing for the interview via two different methods: "social preparation," talking to people who do the job or work for the company and "background preparation," gathering information on the company/job from newspapers, company publications, etc. had an impact on interview performance. Results indicated that both methods of interview preparation were significantly related to interview success.

As Day and Carroll (2003; details of this study are described above) point out, another method of pre-interview preparation for which "real world" human resource professionals are concerned is applicants gaining knowledge of interview questions from acquaintances who have already been through the interview and using the actual questions to prepare answers ahead of time. If an applicant were motivated to perform on the interview to the extent that he or she is willing to use such a technique to prepare, Day and Carroll (2003) found—not surprisingly—that advance knowledge of interview questions have a significant impact on interview success; individuals who were given a copy of the questions prior to the interview received consistently higher interview ratings than those with no advance knowledge of the questions.

The method of pre-interview preparation which has probably received the most attention is interviewee training. Research on the impact of various forms of interviewee training goes back more than 30 years (e.g., Stevens and Tornatzky, 1976). Although this body of literature

has not found universal support for the contention that those who are motivated to undergo interview training will universally enjoy increased interview performance ratings, there is sufficient evidence to suggest interview training typically yields beneficial results. The aforementioned Stevens and Tornatzky (1976) found that members of the treatment group who received interviewer training were more likely to receive higher paying jobs than control group individuals who received no training. Another early study by Keith, Engelkes, and Winborn (1977) found that a significantly higher percentage of individuals in a rehabilitation program who underwent interview training obtained jobs than those who did not. In more recent research, Mauer, Solamon, and Troxtel (1998) and Mauer, Solamon, Andrews and Troxtel (2001) conducted studies of police officers and fire fighters and found that those participants who volunteered for coaching were more likely to perform well in an interview. Given that a link exists between an applicant's motivation and the likelihood of pre-interview preparation, one can point to research such as the positive impact of participant training and the other interview preparation methods discussed above as additional evidence that a link exists between motivation and interview scores.

Personality and the Interview

While the topic of personality and the numerous ways it can affect organizations has long been studied by industrial/organizational psychologists and human resource professionals, the influence of an individual's personality on the selection interview did not begin receiving attention from researchers until the 1980s. Harris' 1989 narrative review of the selection interview literature only identified two studies on the topic, both of which focused on the ability of the interviewer to accurately assess personality components of the interviewee. Specifically,

the review discussed Jackson and colleagues (Jackson, Peacock and Holden, 1982; Paunonen, Jackson, & Oberman, 1987) evaluation of interviewers' ability to accurately evaluate the personality of applicants. Jackson, Peacock and Holden (1982) measured this ability in professional recruiters in a field setting and Paunonen, Jackson, & Oberman (1987) replicated the research in a controlled lab setting with undergraduate students. Results from both studies showed that interviewers may in fact be able to accurately assess applicant personality. Both studies were quick to point out that although interviewers may indeed be able to accurately assess interviewee personality, additional research would be needed to determine the validity of these personality trait assessments for accurately predicting future job performance.

I identified a few additional studies published in the 1980s that were not included in Harris' (1989) review. For example, Keenan (1982) published a study on the impact of candidate personality on selection interviews and concluded that personality was an important factor in interview behavior and outcomes. White (1986) conducted a lab study of 122 undergraduates and evaluated the impact of interviewee assertiveness (an element of an individual's personality) on interview outcomes. His analysis revealed that interviewee assertiveness did in fact have a significant impact on interview success. Similarly, Fletcher (1987) conducted a laboratory study utilizing university undergraduates ($N = 138$) and explored the relationship between various applicant personality characteristics and ratings on an interview which determined admittance to a particular psychology class. Results offered some indication that extraverts were accepted to the class at higher rates. Fletcher also found that neuroticism and self-monitoring scores did not correlate with interviewer scores.

Research on the impact of personality on the selection interview did begin to increase in the 1990s. For example, in Posthuma, Morgeson, and Campion's (2002) review of the literature

published on the selection interview, the authors noted 26 different articles published on the topic of personality as it relates to the interview. While it is beyond the scope of this paper to review all of the studies published on the personality/selection interview relationship since 1989, Posthuma et al. (2002) note that topics on the relationship ranged from the impact of obscure personality aspects such as body image or communication apprehension on interview performance (Ayres and Crosby, 1995) to how personality influences individual interview preparation (Caldwell and Burger, 1998; see Posthuma et al., 2002 for a comprehensive list and a review of this literature). It is worthwhile to observe that few studies published in the 1990s did explore the influence of Big Five factors of personality such as Extraversion, Conscientiousness and Need for Achievement (Barrick and Mount, 1991; Walters, Miller, and Ree, 1993) and found mixed results.

While the early personality/interview research discussed above continued the focus on the numerous variables that could influence ratings in the selection interview, as was the case with the initial literature on the work experience/interview relationship, the applicability of this research to the relationship between personality and the highly structured interviews common today is limited because a clear distinction has been established between situational interviews and past behavior interviews.

The first article to differentiate past behavior and situational interview questions while exploring the relationship between personality and interview performance was Conway and Penno (1999). This experiment evaluated 179 applicants for an RA position on a university campus (details of this study are fully described in the Experience and the Interview section). After the initial screening interview, the 131 remaining participants were administered a combined SI/PBI interview as well as a Big Five personality measure. None of the traits

measured by the Big Five personality scale yielded significant correlations with composite ratings on the PBI: extraversion, $r = .10$, agreeableness, $r = -.02$; neuroticism, $r = .07$; conscientiousness, $r = .00$ and intellectance (another term for openness; $r = .10$, all ratings non-significant. Although as mentioned this research did employ a within-subjects study design, no meaningful impact on the exploration of the personality/PBI relationship was apparent, and no other limitations regarding interpretation of this data could be identified.

Huffcutt et al. (2001) also compared a Big Five personality scale to PBI performance for 93 district managers (details of this study also provided in the Experience and the Interview section above). Findings regarding the relationship between the personality variables and ratings on the PBIs were generally similar to the Cowney & Penno (1999) results in that four of the Big Five personality traits were not significantly related to PBIs. Specifically, Agreeableness, $r = -.01$; Conscientiousness, $r = .08$; Neuroticism, $r = -.05$; and Openness, $r = .06$ yielded non-significant correlations with PBI scores. The difference however, was that a significant positive correlation was found between Extraversion and PBI performance ($r = .30, p < .01$). One factor that potentially limits the generalizability of these findings to other past behavior interviews is that these interviews were conducted over the phone. Any differences in data collection and scoring result from phone administration of a PBI versus typical face-to-face administration are unclear. Similarly, the impact personality has on a past behavior interview conducted over the phone is unclear as well, but the possibility of phone administration confounding the relationship in some manner cannot be completely dismissed.

Finally, Krajewski et al. (2006), like many of the studies mentioned previously, compared the criterion- and construct-related validities of the PBI and SI. In this study, 157 applicants for “high-level managerial positions” within a large organization participated in an assessment

center in which multiple measures of job performance were measured. Assessment center measures pertinent to this discussion were a Past Behavior Interview, a personality measure, and a cognitive ability assessment. The 6-question PBI and accompanying scoring guide were developed from a “meticulous” job analysis and were administered by a two-person panel. There were four interviewers in all, each of whom was thoroughly trained prior to interview administration. Unlike the studies evaluating the personality/PBI relationship discussed above; Krajewski et al. (2006) chose to explore the relationship between PBI ratings and two specific subscales of the Conscientiousness trait of the Big Five personality scale. These two subscales—Achievement Orientation and Dominance—have been previously shown to be significantly related to job performance. It is relevant to mention that while 157 applicants participated in the assessment center overall, but for reasons not discussed in the article, only 84 completed personality measure, thus limiting the sample size. When PBI performance of these 84 applicants was correlated with score on the two personality subscales, small significant correlations were found. Specifically, the PBI correlated with Achievement Orientation $r = .22$ ($p < .05$) and with Dominance $r = .26$ ($p < .05$). Although these two notes are not tremendous limitations of the findings, utilizing a sample roughly half the size of the original leads to a bit less confidence in the PBI/personality relationship reported. Additionally, as Krajewski et al. (2006) point out, the correlation coefficients, although statistically significant, are considered “small” by standard conventions. The authors’ use of the specific subscales of conscientiousness does further the literature in that it provides insight into what might potentially be the cause of the relationship between conscientiousness and past behavior interviews if a relationship does ultimately exist. As Krajewski et al. (2006) acknowledge and is apparent from the studies discussed above,

research on the nature of the relationship between past behavior interviews and personality has yielded mixed results and must be considered inconclusive to this point.

Cognitive Ability and the Interview

In Mayfield's 1964 summary of the literature on the selection interview, the author hypothesized that intelligence was probably the trait most accurately measured by the interview; although the consistently poor analytical data led him to quickly follow up this supposition with the qualifying statement that it likely did not add incremental validity above cognitive ability tests alone. Aside from this speculation, the relationship between cognitive ability and interviews received little attention until the late 1980s when researchers began to call for construct validation studies on the employment interview (e.g., Harris, 1989). Two examples of the initial articles exploring the interview/cognitive ability relationship published in 1988 were Campion, Pursell, and Brown (1988) and Schmidt (1988). These "structured"—though not to the point of designating and utilizing one type of question—interviews showed small- to medium-sized correlations with cognitive ability. Schmidt (1988) and Campion et al. (1988) found the employment interview yielded *r*-values of .42 and .3 respectively when correlated with general cognitive ability.

Again, the lack of structuring the interviews by question type limits the generalizability of studies conducted during this era to the highly structured interviews of today. However, gradually researchers exploring the nature of the relationship between cognitive ability and the interview did begin to utilize interviews that controlled for question type, and today a considerable amount of subsequent research exists exploring the interview/cognitive ability relationship using past behavior interviews and/or situational interviews.

Of the body of research on the relationship between cognitive ability and PBIs, a few studies have found a significant relationship between score on a PBI and an individual's level of cognitive ability. In 1994, Campion et al. sought to compare the incremental validity of both a PBI and a SI above and beyond a cognitive ability measure. The concurrent validation study obtained scores from 70 pulp mill employees on the SI and PBI via a 30-question composite interview—15 situational, 15 past behavior questions—another example of a within-subjects design. Additionally, scores on nine different cognitive ability measures (3 verbal ability scales, 3 mathematical ability scales, 2 measures of perceptual ability, and 1 measure of perceptual ability) were collected. Each of the 7 interviewing managers, previously experienced in interviewing, received a 1-day training course before conducting interviews via a two-person panel. When evaluated independently, the ratings on the PBI significantly correlated ($r = .61, p < .05$) with a composite of the 9 cognitive ability scales.

A correlation as strong as .61 is certainly impressive; however, it must be interpreted cautiously in light of certain potentially influential limitations. First, while a sample size of 70 is not out of the question given the magnitude of the correlation, it is still relatively small. Again, the potential order effects resulting from a within-subjects design must be considered. Order effects are potentially even more relevant in this study considering the fact the 30-question composite interview must have been considerably longer than most structured interviews. Although no average time per interview was reported, the sheer number of questions combined with the fact that situational questions were always asked first and past behavior questions afterwards may have meant participants were more tired and less focused by the time the past behavior questions were asked, causing the participants to pay less attention to the particular type of past experience asked about by the interviewer. In addition, instead of recalling a situation

from long-term memory (which requires relatively little cognitive processing) interviewees were forced to blend multiple scenarios or create altogether new scenarios in an attempt to respond with a situation matching the interviewer's request (obviously a process requiring more cognitive processing).

Krajewski et al., (2006) also found significant, albeit small, positive correlations between cognitive ability and PBI scores. This study (described in the Personality and the Interview section above) of 157 high level managers found correlations between the PBI developed for their study and three different scales of cognitive ability: verbal comprehension, $r = .18$ ($p < .05$); verbal reasoning, $r = .20$ ($p < .05$); and numerical reasoning, $r = .20$ ($p < .05$). As was the case with the findings of Krajewski et al.'s (2006) exploration of the relationship between personality and past behavior interview performance, although the findings were statistically significant, the correlations were in fact quite small.

While the results from these two studies mentioned above cannot be overlooked, a vast majority of the research exploring the past behavior interview/cognitive ability relationship has reached contradictory conclusions finding essentially no support for the existence of a relationship between the two variables. The first study specifically studying the correlation between interviews composed only of behavior description questions was Motowidlo et al. (1992). Unlike many studies of the structured interview that compare and often combine the past behavior interview and the situational interview (which yields potentially confounding findings), Motowidlo et al.'s (1992) research exclusively explores the PBI in a series of five separate studies, four of which addressed the past behavior interview/cognitive ability relationship. In Study I; the authors developed a 7-question PBI from a thorough job analysis. A 3-tiered rating scale was developed with behavioral examples differentiating between low, medium, and high

responses. Nineteen interviewers—company recruiters who received a thorough 3-day interview training—administered the interviews to 107 applicants for entry-level management positions from 7 different telecommunication companies. The participants also were administered a cognitive ability measure with 4 subscales—quantitative reasoning, writing fluency, reading accuracy, and following directions. In Study I, overall score on the PBI was not significantly correlated ($r = .07, ns$) with composite score on the cognitive ability measure.

Study II replicated Study I; it involved administration of the same PBI developed in Study I and the same cognitive ability scales to 164 current managers from the same 7 telecommunications companies. Aside from the sample composition, the only notable difference in Study II was that the interviewers—this time 25 company recruiters—only received 1 day of training. This difference had little impact overall; the results were similar to those obtained in Study I. The correlation between the PBI and the cognitive ability measure was not significant, $r = -.09, ns$.

In Study III, the same PBI was administered to applicants for entry-level management positions at one of the telecommunications firms involved in Study I and II. Unlike the first two studies, a statistically significant correlation between the PBI and average cognitive ability score was found. However, the correlation was, in reality, quite small ($.17, p < .01$) and likely only reached significance due to the large sample size ($N = 875$).

Finally, Study IV involved the development of a different PBI for a different position within the telecommunication organizations. For this study, a 7-question PBI and behavioral rating scale were created for a marketing position using the same rigorous development procedures discussed in Study I. Apparently, 18 company recruiters (each of whom received one day of interview training) interviewed 176 marketing incumbents. However, Motowidlo et al.

(1992) states that scores on the cognitive ability measure were “available for 36 incumbents in the validation sample.” Although this inference could not be explicitly confirmed in the study, it appears the PBI/cognitive ability correlation in Study IV is based on a small sample of 36. Like Studies I and II, no significant relationship was found between this PBI and cognitive ability ($r = .13, ns$). Based on the 3 non-significant and 1 significant though quite small correlations with cognitive ability, and an overall average correlation (with each of the 4 studies weighted according to sample size) of .14, Motowidlo et al. (1992) concluded that PBIs were generally not correlated with cognitive Ability. The fact that the study replicated its findings, used both concurrent and predictive designs, and contained no major design limitations allows one to have a degree of confidence in the generalizability of the Motowidlo et al. (1992) findings.

As mentioned previously in the discussion of Pulakos and Schmitt (1995; described in detail in Experience and the Interview section above) the second study reported in this article involving 464 employees in a federal organization examined the relationship between past behavior interview performance and cognitive ability, which was assessed by measures of verbal reasoning, quantitative reasoning, fluid intelligence and crystallized intelligence subscales). As a result of the between-subjects study design, half (232) of the participants took the PBI. When correlated with the composite cognitive ability measure, no significant relationship between PBI and cognitive ability was found ($r = .09$). Another between-subjects research design study testing the PBI/cognitive ability performance relationship was Day and Carroll (2003). In this study (described in detail in Experience and the Interview section above), as in Pulakos and Schmitt (1995), when the interview scores from the half of the participant pool who took the PBI were correlated with cognitive ability scores, no significant relationship was found ($r = .10, ns$). As

mentioned in the introduction, the relationship between PBI ratings and cognitive ability has yielded a considerable amount of research.

Two further studies were Conway and Peneno (1999) and Huffcutt et al. (2001). These found similar results regarding the PBI and cognitive ability correlation. Conway and Peneno's (1999) study of applicants for university resident-assistant positions (reviewed in Experience and the Interview section above) collected cognitive ability scores for all 131 applicants who made it past the screening round. When those scores were correlated with the past behavior type questions from the second round interview, no significant relationship between the two was found ($r = -.10, ns$). Similarly, Huffcutt et al. (2001; discussed in Experience and the Interview section above) found a Pearson r of $-.09$ when the PBI performance of 93 district managers from a merchandise chain was correlated with a cognitive ability measure.

Finally, one other study should be discussed that sheds light on the nature of the relationship between cognitive ability and PBIs. Huffcutt et al. (1996) conducted a meta-analysis of the literature to that point which had explored the past behavior interview/cognitive ability relationship. The authors found 7 studies evaluating this relationship, including Motowidlo et al. (1992), Campion et al. (1994), and Pulakos and Schmitt (1995) discussed above; the other four studies appear to be from unpublished papers and dissertations. The meta-analysis of these 7 studies revealed a correlation of $.18$ between PBIs and cognitive ability. Obviously, one must take into consideration the fact that this research was conducted in 1996 and multiple studies (many of which are discussed above) have been conducted on this topic since its publication, which limits the generalizability of these findings. Future research would benefit from an updated meta-analysis of the PBI/cognitive ability relationship. Looking at the body of research on the relationship between cognitive ability and PBI score collectively, there seems to be

enough consistent evidence to allow one to draw the general conclusion that cognitive ability does not appear to have a significant impact on past behavior interview performance.

Current Study

The practical application of this research is that it will allow insight into the mechanisms that influence the most critical personnel selection decisions—the selection of senior personnel. Frequently considerable financial implications depend upon these decisions, and almost no quantitative analysis of selection procedures at this level exists. This research is designed to accomplish multiple objectives. First, several researchers have acknowledged the need for the identification of the constructs measured by structured interviews (e.g., Conway and Peneno, 1999; Posthuma et al. 2002); indeed, Huffcutt et al. (1996) called the identification of the constructs measured by structured interviews the “next major breakthrough” (page 470) in construct and structured interview research. Some research has attempted to identify those constructs that could be related to structured interviews—and the Past Behavior Interview (PBI) in particular (e.g., Conway & Peneno 1999; Day and Carroll, 2003). While a variety of constructs have been suggested, research has been predominantly inconclusive and as Taylor and Small (2002) suggest, further research is needed to establish the extent to which structured interviews converge with the broader underlying constructs with which they have been suggested to be correlated.

To that end, one of the purposes of this study is to offer further insight on four potential constructs—Experience, Motivation, Personality and Cognitive Ability that could be linked to PBI performance. Further, as the correlation between experience and motivation with PBIs has received little or no quantitative investigation in the past, this study will be able to contribute

unique and previously unknown information to the literature on these two constructs. Specifically, the relationship between experience and past behavior interviews has only been operationalized as a single, general construct in the past, typically represented by a specific number of years (e.g., years on the job) as an indicator of an individual's general "experience." In this study, experience is actually operationalized by a valid, reliable instrument designed to measure the depth and breadth on an individual's experience.

Additionally, the Leadership Experience Inventory (LEI; the scale used to measure experience in this study which is described in detail below) is comprised of four statistically grouped subscales which are operationalizations of the four meaningful components that make up the overall LEI rating (described in greater detail below). In previous research, subscales of the general constructs of cognitive ability (e.g., reasoning ability and verbal ability) and personality (e.g., extraversion, conscientiousness) have been explored in order to explain what specific elements of these constructs might be causing the observed relationship between the two variables. Similarly, the LEI provides specific information about the correlation of the elements underlying the PBI/experience relationship as the LEI allows for the correlation of the four specific subscales of the experience construct with PBI performance as well; this is information which was previously unavailable.

Similarly, Motivation will be quantified in a manner such that its subscales can be examined to shed light on the nature of the relationship with past behavior interviews; however, the more significant contribution to the scientific literature is that, as mentioned above, the extent of relationship between Motivation and PBI performance has never been statistically tested. This study will be the first to empirically examine the Motivation/PBI relationship that had previously only been theoretically supported. Although no specific directional a priori hypotheses will be

made in this study, this study is able to make meaningful contributions to the overall understanding about the nature of the relationships between motivation and experience and PBI performance with a post hoc analysis of data, as these relationships have never been explored below the general construct level, if at all.

Research Questions

Hypothesis 1a: It is hypothesized that experience, as measured by raw score on the LEI, will be significantly positively correlated with overall Past Behavior Interview rating. This hypothesis stems from two main sources. First, on the surface, the existing research on the relationship between experience and PBI performance seems to yield mixed results; two studies found no significant relationship between experience and PBI performance while two others did find significant results. However, consideration of additional information from the studies – noting those which are rigorously developed and highly controlled as well as those with flaws in study design or research methodology can (and in this case, do) lead some research findings to carry more weight than others. In this particular case there were multiple limitations observed in the studies that found no significant relationship between experience and past behavior interview (PBI) performance—from the severely limited range in participants years of experience in Pulakos and Schmitt (1995) to Huffcutt et al.'s (2001) use of a behaviorally anchored rating scale (BARS) with only 3 response levels. On the other hand, aside from the minor problem of Conway and Peneno (1999) analyzing a small-ish sample, the research finding a relationship between experience and PBIs was methodologically sound and did not possess any factors limiting the interpretation of results.

This hypothesis is based on theoretical evidence as well. While commenting on the inability of completely standardized interview questions administered verbatim to accurately

capture a variety of meaningful individual differences which have a real impact on job performance, Taylor and Small (2003) state, “differences in work experience and thus the situations previously faced by interviewees are likely to account for a proportion of true variance in job performance as prior work experience and job performance are at least moderately correlated,” (page 278). This statement about the impact of work experience has implications for Hypothesis 1a as well. Working from Day and Carroll’s (2002) logic, it is reasonable to expect that if amount of work experience is likely to account for a proportion of true variance in job performance, then work experience should also account for a proportion of the variance in quality of behaviors described in past behavior interviews, as those past behaviors are direct examples of past job performance. This sound logic, along with careful analysis of the existing PBI/experience literature, points to a meaningful relationship between the two.

Hypothesis 1b: It is hypothesized that the Personal and Career Related Experience super-factor of the LEI will yield a smaller relationship with overall past behavior interview performance rating than the other three LEI super-factors—General Management Experience, Overcoming Challenge and Adversity, and Risky and/or Critical Experiences. Each of the four super-factors of the LEI assesses previous work experience to some degree. Furthermore, a closer review of the specific types of past experiences assessed by the LEI indicate a majority of the various subscales and specific behaviors measured are practical, important areas of work which could easily be paralleled with competencies and the behaviors identified as elements critical to the successful performance of a particular job. It would not be surprising then for many of the elements measured in the LEI to be competencies/behaviors measured in a PBI.

All that being said, of the 4 super-factors in the LEI, the Personal and Career Related Experience super-factor seems to assess elements of experience less related to work in general

(e.g., “Extracurricular Activities” is one of the subscales) and of the behaviors which could be considered related to on-the-job activities, they seem to be more peripheral than the fundamental business experiences measured in other super-factors. In the limited time available, past behavior interview questions are highly focused on eliciting descriptions of the most important behaviors related to successful performance in the job at hand. Personal and Career Related Experience is expected to yield a smaller correlation with PBI performance then, because the types of behaviors and experiences measured in this competency are not the types of behaviors that would likely be deemed critical in performing a job well and therefore are not likely to be assessed in a PBI. On the other hand, the types of experiences measured in the remaining three super-factors seem to all potentially be considered critical competencies for any number of jobs and subsequently assessed in a PBI.

Hypothesis 2a: It is hypothesized that motivation, as measured by the Motivation super-factor of the Global Personality Inventory (GPI; the personality measure used in this study, described in detail below), will be significantly positively correlated with overall past behavior interview performance rating. No direct quantitative analysis supports this hypothesis; in fact, this hypothesis has not been directly analytically tested at all. This study will be the first to quantitatively explore the relationship between motivation and PBI performance. The hypothesis is based on Taylor and Small’s (2002) theoretical argument (discussed in detail above) that due to the fact the past behavior interview is a measure of typical (as opposed to maximal) performance, it is not only able to tap declarative and procedural knowledge and skills but also able to tap motivation, the third determinant of job performance and therefore should be significantly correlated to the motivation construct measured by the GPI.

Additionally this hypothesis is supported indirectly by research on the beneficial effects of various types of interview preparation. Specifically, an individual more motivated overall would be more likely to take the time to prepare for an interview via research on the organization (Caldwell and Burger, 1998) or undergoing some type of interview training (Mauer et al., 2001).

Hypothesis 2b: It is hypothesized that the Desire for Achievement subscale of the Motivation super-factor will display a significantly greater correlation with PBI ratings than either the Energy Level or Initiative subscales. This is hypothesized for two reasons. First, the Desire for Achievement subscale directly reflects an individual's motivation as it relates to work situations. It measures "the tendency to be ambitious in the advancement of one's career or position in organizational hierarchy," (page 6; Personnel Decisions International, 2001). Based on Taylor and Small's (2002) theory discussed earlier that motivation to perform at work is part of what is measured by the past behavior interview, the Desire for Achievement should be reflected in a score on a PBI. Research on the meaningful impact of various methods of preparation also was discussed previously, and the assumption was made that those superiorly motivated in general would be more likely to prepare for an interview and thus would be more likely to perform well. One can be even more confident in the likelihood that individuals high in Desire for Achievement will take steps to prepare for an interview, for these individuals possess "a continual desire to get ahead of where one is currently in work..." (PDI, page 6).

Hypothesis 2c: Ultimately, this hypothesis states that the 7 remaining subscales of the Big Five factor Extroversion will not add incremental validity to the PBI performance/Extroversion correlation beyond that of the 3 motivation subscales of the GPI (all of which fall under the Extroversion factor). However, several conditions must first be met before this hypothesis can be tested. First, a significant positive correlation between Motivation and PBI ratings, as

hypothesized in Hypothesis 2a, must be found. Further, when the GPI items are restructured around the Big Five personality scale, a significant positive correlation must be found between Extraversion and PBI ratings as well..

It should be noted that of the two studies identified evaluating the impact of Big Five personality traits on PBI performance, only one dimension— Extraversion—in one of the two studies (the Huffcutt et al., 2001 research), was found to be significantly correlated with PBI performance. If the conditions to test this hypothesis are met, this study could shed light on the mechanism through which extraversion impacts interview performance.

Hypothesis 3a: It is hypothesized that 8 of the 9 dimensions of personality measured as super-factors of GPI will not be significantly correlated with overall past behavior interview performance rating. As mentioned in Hypothesis 2a, the one personality dimension/GPI super-factor predicted to be significantly correlated with overall PBI performance rating is Motivation. Unlike the inconsistent findings of the research on the Experience/PBI relationship, more thorough analysis of the PBI performance/Personality quantitative research still yields inconclusive results and does not provide any hint of the significance or direction of the relationship between the two variables.

This hypothesis is based predominantly on the nature of the independent variable; that is, the specific structural components built-in to the actual past behavior interview designed to limit the influence of variables that could potentially confound ratings. For example, the two biggest factors minimizing the influence of extraneous variables such as personality on interview ratings are: 1.) the utilization of a critical incident job analysis to identify the competencies critical for job success and develop the actual questions and to assess those competencies; and 2.) utilization of behaviorally anchored rating scales (BARS). The job analysis is the element that focuses the

entire interview exclusively on job-related factors and job-relevant behaviors so that irrelevant variables such as personality are not introduced. The BARS clearly puts forth the only behaviors—job-relevant behaviors—which ratings are to be based upon (Campion et al., 1997). Additionally, another element of structure—interviewer training—can prevent personality factors from seeping in and influencing PBI ratings. Training makes interviewer aware of the various sources of bias—similarity, confirmatory, first impression bias, many which stem from effects of personality (Campion et al., 1997). Finally, I am even more confident in this hypothesis because individuals conducting the interviews in this particular study are highly trained consultants, many of whom have graduate degrees in psychology, years of interviewing experience, and are acutely aware of the dynamics of interpersonal interactions and the most subtle attempts at impression management.

Hypothesis 4a: It is hypothesized that no statistically significant correlation will be found between general cognitive ability, as measured by overall score on the Watson-Glaser, and overall past behavior interview performance rating. This hypothesis is based upon the literature reviewed above. Overall, there seems to have been a thorough investigation into this relationship yielding consistent enough results to support this hypothesis.

Exploratory Research Questions

Research Question 1: All of the prior research identified exploring the relationship between personality and past behavior interview performance utilizes the Big Five personality scale; while analysis of the GPI in its standard 9 factor structure is interesting and yields meaningful hypotheses and research questions, it would also be meaningful to evaluate the current data in terms of the Big Five factor model and compare the results of this research with

previous research along Big Five factors. As the GPI was in fact developed based on the Five factor model, the items can be restructured and evaluated along Big Five dimensions allowing for meaningful comparisons of this study to other Big Five research. The specific research question then is how do the results of this study compare to the results of other studies that have explored the relationship of the Big Five personality factors to past behavior interview performance?

Research Question 2: Experience has typically been operationalized solely in terms of the number of years a person has performed a task or group of tasks; it is likely that little attention has been paid to the components that “make up” experience as a construct. The LEI, however, is able to capture these components, or subscales of experience. The research question that presents itself then is which of these components of experience is the most correlated with past behavior interview performance? Which is the least correlated? Where do the other subscales fall in between? Are there any theoretical explanations for the order and variability in predictive ability?

Research Question 3: This research question is much the same as Research Question 2, except that the construct of interest is motivation and neither the construct as a whole nor the components have been evaluated in terms of their relationship with past behavior interview performance. The research question of interest is which of the three Motivation subscales has the strongest correlation with PBI performance, and which has the weakest correlation? Again, is there any theoretical explanation for these relationships?

CHAPTER 2

METHOD

Participants

Archival data from 1023 incumbents of, or applicants for, Business Unit Leader (BUL; often called Vice-President) positions from numerous organizations across the United States will be analyzed for this study. Typically the BUL level includes individuals who set policies and goals for companies (or at least a division within the company) typically greater than 500 employees in size. Between July of 2004 and May of 2007 these individuals participated in a two-day assessment center involving multiple assessments developed by a human resources consulting firm. Of the respondents, approximately half (48.1%, $n = 492$) either chose not to list age or did not have age data requested. Of the participants reporting age, the average age was 44.40 years old ($SD = 6.25$). It is understandable that those undergoing the assessment process might be hesitant to report their age for fear of it biasing ratings.

A majority of the sample (89.1%) reported gender information. Of those reporting, a majority, (69.7%, $n = 713$) were male. Of those reporting ethnicity information (24.5% did not report) a majority of the sample (91.6%, $n = 707$) were white, with others being African American (3.1%, $n = 24$), Hispanic/Latino (2.6%, $n = 20$), and Asian/Pacific Islander (1.3%, $n = 10$). As mentioned, there were various reasons individuals participated in the assessment center process; most (88.6%, $n = 906$) were for development purposes (i.e., identifying individual strengths and opportunities for growth). Other reasons for participation in the assessment center were for selection purposes (6.8%, $n = 70$) and for internal readiness purposes (i.e., how prepared the individual is to be promoted to the next level; 4.6% $n = 47$).

Additional demographic information was available about the sample as well. Most of the

sample (89.8%) spent their early years (i.e., “grew up”) in the United States. Further, of the 788 reporting data on highest level of education obtained, 15 participants (1.9%) reported less than a high school degree, 69 (8.8%) had only a high school degree, 355 (45.1%) reported obtaining an undergraduate degree, 298 (37.8%) had a post-graduate degree, and 51 (6.5%) had a doctoral/professional degree. Finally, participants reported a having a variety of job functions, with the most popular being “business/general management” (30.8%, $n = 321$), “finance/accounting” (13.3%, $n = 138$) and “marketing/advertising” (10.7%, $n = 111$).

Regarding the organizations from which the participants came, of the 824 participants reporting industry type, the largest proportions came from “transportation/freight and logistics” (20.6%), “wholesale/retail trade” (14.8%), “banking/financial services” (14.0%), “business services” (7.6%), and healthcare (5.0%). As for the sizes of these organizations, they ranged in size from less than 1,000 to 100,000 or more, with 214 (26.3%) of the 815 participants reporting information on the topic being employed in organizations of more than 100,000 or more.

Measures

Past Behavior Interview

The past behavior interview (PBI) developed for this study was created in a manner very similar to the methodical process described in the introduction to the PBI previously. The interview was developed from a thorough behavioral job analysis involving interviews with numerous subject matter experts (SMEs) and content experts. For example, business unit leader (BUL) incumbents, managers of business unit leaders such as CEOs, CFOs, and executive vice presidents, as well as individuals intimately familiar with the demands and requirements of the position such as experienced executive coaches and board members of corporations were all

involved in the identification of critical incidents.

It is accepted that the general leadership characteristics necessary for successful performance at a given level of responsibility consistently appear across organizations; therefore in order to gain a complete understanding of the characteristics necessary for success at the BUL level individuals from multiple organizations were involved in the job analysis process. Additionally, since the broad skills necessary for success remain relatively constant, once the competencies are identified, they can be meaningfully assessed in BUL incumbents and applicants from different organizations. More specifically, the same interview questions can be administered to participants from different companies.

Once the critical incidents were established, 16 competencies critical for success at the BUL level within an organization were identified. It was determined past behavior questions could be developed from the critical incidents to sufficiently gather data and measure 12 of the 16 competencies. The competencies measured by the PBI were: Think Strategically, Innovate, Display Global Perspective, Influence Others, Engage and Inspire, Build Talent, Ensure Execution, Drive for Results, Focus on Customers, Lead Courageously, Inspire Trust, and Adapt and Learn. Drive for Results was also measured in the interview although no specific questions were developed assessing this competency; it was assumed that this competency could be sufficiently rated from answers generated for other competencies (see Table 6 for a list of these competencies and their definitions). Multiple questions were designed for each competency (except for Drive for Results), and the interviewers were allowed to choose from the list of questions and ask them in varying order, depending on the flow and direction of the interview. Additionally, interviewers were encouraged to ask good probing and follow-up questions to ensure sufficient data was collected to accurately score each competency.

The interview was conducted one-on-one; the interviewer took detailed notes throughout the interview and referred back to them when scoring upon completion of the interview. The PBI was scored using a behaviorally anchored rating scale, which again is essentially a five-point, Likert-type rating scale with specific behavioral examples for the lowest, middle, and highest ratings for each key behavior/skill measured in every competency. The BARS was developed based on the same critical incident job analysis process as the questions were. As mentioned, upon completion of the interview, the interviewer uses the detailed notes to rate each of the specific behaviors identified under each competency using the behavioral anchors to accurately place the behaviors reported by the participant. After each of the specific behaviors for a competency has been rated on a 1 to 5 scale, an overall 1 to 5 rating for that competency is assigned based on the behavior ratings. After each of the 12 competencies has been rated in this manner, an overall categorical rating (e.g., below average, on par, above average, well above average) is given for the interview performance overall. In order to statistically compare performance on the past behavior interview with the other scales, an “overall past behavior interview rating” will be calculated as the 12 competency average. Each interview took approximately one hour and fifteen minutes to conduct.

While the interviewers’ main goal and focus was to elicit and accurately record participant descriptions of behaviors related to the competencies critical for success, in this PBI, it should be noted that interviewers also used this time to accomplish a few additional subtle, yet important goals. For example, the interviewers used this time to answer any questions the participant may have about the assessment process or alleviate any undue anxiety or stress. Additionally, interviewers sought to establish a rapport with the participant during this valuable time. Establishing rapport during the interview was especially important if the participant was

going through the assessment center for developmental purposes; in these instances the interviewer was also the one to deliver the results upon completion of the assessment center and give performance feedback. Furthermore, as the assessment center was set up (which will be discussed later), the interview is the only time the consultant had to interact with the participant before the integral and often very personal feedback session. To this point, the interviewer also worked to gather a general sense of past work experiences, the current work situation, as well as values and future goals of the participant; this type of information helps the consultant to make the feedback experience more meaningful and relevant in that the information is delivered in the context of current issues the participant is facing and what can be done with the feedback moving forward to improve and accomplish goals.

The interviewers were highly trained consultants from the firm which developed the past behavior interview and assessment center. Each possessed a graduate degree in business, psychology, or a related field; many possessed a doctorate in industrial/organizational psychology. Many of the interviewers had previous experience in interviewing as well as observing and recording behavior. The systematic training method employed allowed the interviewers to become intimately familiar with administering the PBI, eliciting and capturing behavior descriptions from interviewees, and scoring the interview based solely on the behaviors reported and their value according to the scoring guide. The training process consisted of the consultants first studying and learning the past behavior questions and associated competencies; then the behavioral modeling technique was utilized, where the consultant observed, or shadowed, other consultants conducting interviews. Often, training interviewers also utilize role-playing and practice interview training techniques to prepare to lead an interview. When the consultant learning the process felt comfortable enough to lead an interview with another

experienced interviewer in the room, the experienced consultant would observe and provide feedback and coaching on the interview. This supervision and feedback process would continue until the consultant could confidently execute the interview and maintain a set standard.

Experience—Leadership Experience Inventory

The LEI is a standardized self-report measure of leadership-related experience. It consists of 118 items used to interpret the nature and extent of managerial experience an individual possesses. The resulting data is organized into four super-factors: General Management Experiences, Overcoming Challenge and Adversity, Risky and/or Critical Experiences, and Personal and Career Related Experience. Each super-category consists of between four and ten subscales; a total of 23 subscales are measured in all (LEI super-factors and subscales are listed in Table 3). The items within a subscale yield a score for that particular subscale, and scores on the subscales within a super-factor are simply summed to calculate a score for each of the four super-factors. However, the scoring system used on the LEI is not altogether straightforward. Various response formats are used to collect multiple types of experience-related data on the LEI. On some of the items the participants respond on a 1 to 4 Likert-type rating scale indicating the actual number of times he or she has experienced a certain event that is described, where 1 indicates *never had the experience* and 4 indicates *had the experience 6 or more times*. On other items the participant is given a certain scenario and asked to indicate the number of times he or she has been involved in that particular type of situation in each three different levels of responsibility: as a “contributor,” “management,” or “lead strategist.” The participant indicates the number of times he or she has experienced each scenario given in each of the three aforementioned roles on a 1 to 3 Likert-type rating scale where 1 indicates *no experience with*

that type of situation in that role and 3 indicates *experienced this type of situation in that particular role 3 or more times*. Finally, in a third set of questions, the participant is asked to indicate the amount of experience he or she has with a particular scenario in terms of number of months via a Likert-type rating scale where 1 indicates *never experienced that type of activity* and 4 indicates *experienced that type of activity for 25 months or longer* (See Appendix A for examples of each). This data is differentially weighted and combined according to an algorithm which is proprietary and not available to the public. For feedback, the participant is shown graphically the percentile in which he or she falls in terms of possession of each dimension relative a specific norm group as well as the general population. It should be noted that an actual raw score—the overall sum of the four algorithm-deduced super-factors is also calculated—but never shown to the participant. This overall leadership experience score will be used in the statistical analyses. The LEI is typically administered online and un-timed, but usually takes approximately 35 minutes to complete.

In terms of predictive validity, the LEI has been shown to be significantly related to “career outcomes” indicators such as salary ($r = .18, p < .05$), and advancement potential ($r = .18, p < .05$), but the LEI was not significantly correlated with performance. Additionally, the LEI was shown to predict managerial competencies and career outcomes over and above managerial tenure alone. Since the LEI is the only measure of its kind, no convergent validity data is available.

Personality—Global Personality Inventory

The GPI was developed based on the Big Five factor personality model; this provides sound reliability evidence described below. Additionally, this allows the GPI subscales and items

to be meaningfully restructured along the Big Five dimensions (organization of the GPI subscales under the Big Five personality factors is depicted in Table 5).

The GPI is a cross-cultural measure of personality in a work-context, it contains 300 items which make up the 37 subscales; the 37 subscales are grouped into 9 super-factors. (super-factors and associated subscales are listed in Table 4). Each subscale contains from 7 to 10 items. On the actual instrument, respondents indicate the extent to which they agree or disagree with a statement using a 5-point Likert-type scale ranging from *strongly disagree* (1) to *strongly agree* (5). Each point on the Likert scale increases in value one-quarter of a point such that if the respondent chooses a “1” s/he receives 0 points, a “2” s/he receives .25 points, “3” s/he receives .50 of a point, up to the respondent choosing a “5” and getting one full point for that answer. It follows then, that the highest point value for any given subscale is equal to the number of items within that subscale; that is, if there were 8 items on the subscale and the participant chose a “5” indicating that s/he strongly agreed with each statement, then s/he would receive an 8 on the subscale. The participant receives a standardized score of 1 to 5 on each subscale representing the extent to which the individual possesses that trait. Each subscale is also graphically presented to show the percentile in which the participant falls in terms of amount of possession of that personality dimension relative to a particular norm group. The GPI is typically completed online and is un-timed.

The super-factors are only used for conceptual organization/grouping when interpreting feedback and therefore no actual score on the super-factors is calculated. For the analytical purposes of this study, super-factor scores were calculated by standardizing the subscale scores (necessary because some subscales contained more items than others and would have been more heavily weighted) and calculating an average of the subscales within each super-factor.

There is various evidence of validity and reliability of the Global Personality Inventory. For example, construct validity of the instrument was provided by a factor analysis which revealed a five-factor model underlying the GPI—paralleling the accepted Big Five factor model of personality—as well as facet-level loadings similar to previous Big Five analyses (GPI Technical Manual, 2006). Multiple studies have demonstrated validity of the GPI in terms of ability to meaningfully predict future job performance. Specifically, a study of 202 managers using an overall measure of job performance as the outcome variable demonstrated criterion-related validity coefficients of the nine overarching Performance Factors ranging from .17 to .32 (Personnel Decisions International, 2008). Additionally, when correlation coefficients were calculated with GPI items organized along the Big Five factor structure, all of the factors except Neuroticism were significantly positively correlated with a measure of overall job performance. Analysis of the dataset ($N = 198$) yielded correlation coefficients of .30, .22, .25, and .17 ($p < .05$) for Extraversion, Agreeableness, Conscientiousness, and Openness to Experience respectively.

The GPI has demonstrated sound reliability evidence as well. Analyses of a study involving 300 participants found a test-retest reliability of .78. Internal consistency reliability analyses revealed that reliabilities of the 37 facet scales ranged from .48 to .88. The average reliability was .71. The nine Performance Factor reliabilities ranged from .75 to .91, the average being .85 (PDI GPI Facts on Validity and Reliability, 2008). Sample GPI items are provided in Appendix B.

Motivation—Global Personality Inventory

Motivation is measured by a specific super-factor on the GPI, which is described in detail

below. The GPI Technical Manual (PDI, 2001) describes super-factor of Motivation as the measure of “the tendency to demonstrate motivated behavior that leads to successful work outcomes” (page 5). This includes characteristics such as consistently being active and energetic; the tendency to “take initiative in a proactive, rather than reactive manner; to have a strong drive, a desire for achievement; and to realize personally meaningful goals,” (page 5). Specifically, this construct is captured by 3 facet scales (subscales) of motivation: Energy Level, Initiative, and Desire for Achievement. Participants receive a standardized score from 1 to 5 on the Motivation super-factor as well as each subscale. Additionally, where the participant falls relative to his or her specific norm group and the general population is depicted graphically for each. Reliability analysis of this factor in 714 examinees revealed the overarching Motivation super-factor possessed a reliability of .89 (the second highest reliability of the 9 super-factors). The three facet scales, Energy Level, Initiative, and Desire for Achievement yielded reliabilities of .76, .79, and .77 respectively.

Cognitive Ability—Watson-Glaser Test of Critical Thinking

The Watson-Glaser test measures critical thinking skills, which includes the ability to: define a problem, select pertinent information for the solution of a problem, recognize stated and unstated assumptions, formulate and select relevant and promising hypotheses, draw valid conclusions and judge the validity of inferences. Four specific scales are measured: Inference, Recognition of Assumptions, Deduction, and Interpretation. It is a sound measure of an individual’s possession of the skills necessary for sound judgment and common sense problem solving. The 80-item instrument is typically administered on a computer and must be proctored; it is officially un-timed but the participant has an hour to complete the test. Although there is

only one right answer to each question and the response format remains the same in that the participant may only choose one of the possible answers on each item the response options change in that some items are Yes/No whereas other items provide four multiple-choice options. Participants must choose from a varying number of possible response options depending on the particular question. Upon completion the participant receives a raw score from 1 to 80, indicating number of items answered correctly. Item level performance is not available for this instrument.

The Watson-Glaser has shown convergent validity by demonstrating statistically significant correlations with the Stanford Achievement Test, overall student GPA, and the Wechsler Adult Intelligence Scale. The four scales assessed in the Watson-Glaser yielded internal consistency reliabilities from .69 to .85, as measured through split-half coefficients. Stability of scores, as measured through a test-retest coefficient, was .73 (Watson-Glaser Critical Thinking Appraisal Manual; 1980).

Procedure

The participants are emailed the Global Personality Inventory (GPI) and the Leadership Experience Inventory (LEI) ahead of time to be completed as pre-work prior to arrival. The rest of the assessments, including the Past Behavior Interview (PBI) and the Watson-Glaser Critical Thinking test are administered over a two-day period in what is called an assessment center. An assessment center is an intense, highly-structured and thorough assessment process in which the individual participates in multiple assessment procedures designed to objectively measure, through a variety of methods and perspectives, the key competencies identified in the job analysis necessary for success. While the GPI, LEI, Watson-Glaser and PBI are the assessment center components relevant to this study, as mentioned, additional assessments—including

multiple role-plays and an in-basket exercise—were administered as well. Every measure was scored by a different assessor who was not involved in, nor aware of the participant's performance on, other assessments. Again, the goal of this comprehensive process is to provide a holistic picture of the participant's abilities—via multiple individual's viewpoints and utilizing a variety of methods of collecting data.

It is not surprising that it is logistically challenging to schedule participants so that each assessor only sees the participant once; especially considering 7 different assessors are required per participant. The consulting firm conducting the assessment center typically preferred to conduct the PBI first on day one considering the nature of the additional goals of the interview as previously discussed (i.e., explain any questions about the process the participant may have, calm any apprehensions or fears, establish a rapport for the feedback session); however, as a result of the often tight scheduling depending on number of participants in a particular assessment center, this was not always the case. Occasionally the past behavior interview had to be scheduled for another time during the two day process. This was not expected to meaningfully impact interview data. On the other hand, the remaining measures administered during the assessment center (role plays, in-basket activity, cognitive ability measures, etc) were regularly scheduled for different times such that the times and order of assessments for one participant may be, and usually is completely different from the times and order of assessments for another participant. This serves two purposes: first it allows schedules to be arranged in such a way that every participant is seen by a different assessor in each measure and it combats the likelihood of order effects.

Analyses

First, descriptive statistics were calculated for the sample evaluated in this study. Frequencies and percentages for gender and ethnicity as well as purpose for assessment center participation (participants who were undergoing the assessment center for developmental purposes versus selection purposes) are reported. In addition, means and standard deviations for age were calculated. Frequencies and percentages are reported for the type of organization from which the participants come (e.g., manufacturing, retail, medical, non-profit, etc) will be reported if available. Finally, time frame during which data was collected is also reported.

Next, descriptive statistics for the dependent variable, the past behavior interview, are reported. Means, standard deviations, skewness, and kurtosis and were calculated for all 12 competencies as well as overall PBI rating. Again, the overall PBI rating is the mean of the ratings (ranging from 1 to 5) assigned for each of the 12 competencies. Due to variations in PBI item selection—the interviewer is allowed to choose from a list of possible interview questions depending on the participant and direction of the interview and is not required to identify which question was used—it was determined that item-level reliability analysis cannot be calculated. Descriptive statistics for the measures of the independent variables are presented as well. Specifically, means, standard deviations, skewness, kurtosis, and Chronbach's alpha reliability analysis are reported for each super-factor and subscale measured by the GPI and LEI, with the Motivation super-factor and its three subscales being reported separately as it represents a specific independent variable. These descriptive statistics were also calculated for the overall raw score on the LEI (again, the LEI raw score is the overall composite score on the LEI which is not available to the participant but is used in the current study for statistical analysis). Finally,

descriptive statistics are provided for scores on the Watson-Glaser, but as item-level data is not available on this measure, item-level reliability information is not be available.

For Hypothesis 1a, Pearson's product-moment correlation analysis was conducted between overall raw score on the LEI and overall past behavior interview rating in order to assess the strength of the relationship. The correlation coefficients were corrected for unreliability in the predictor and criteria which may attenuate the relationship.

To test Hypothesis 1b, first Pearson's product-moment correlation analyses was conducted between overall past behavior interview rating and each of the four super-factors of the LEI: Personal and Career Related Experience; General Management Experience; Overcoming Challenge and Adversity; and Risky and/or Critical Experiences. Following this, a Z-test for differences in correlations between Personal and Career Related Experience and each of the three remaining LEI subscales was conducted.

To test Hypothesis 2a, Pearson's product-moment correlation analysis was conducted between overall score on the Motivation super-factor of the GPI and overall past behavior interview rating in order to evaluate the correlation between the two variables. The correlation coefficients were corrected for unreliability in the predictor and criteria which may attenuate the relationship.

To test hypothesis 2b, Pearson's product-moment correlation analyses was conducted between: overall past behavior interview rating and each of the three subscales of the Motivation super-factor of the GPI: Desire for Achievement; Energy Level, and Initiative. Subsequently, a Z-test for differences in correlations between Desire for Achievement and Energy Level as well as Desire for Achievement and Initiative was conducted.

To test Hypothesis 2c, first Person's product-moment correlation analyses was conducted between overall past behavior interview performance and the GPI super-factor of Motivation and its 3 subscales: Energy Level, Initiative, and Desire for Achievement. Assuming a significant positive relationship was found between Motivation, its subscales, and PBI ratings, additional analyses were conducted once the GPI is restructured along Big Five dimensions. Once the GPI items have been restructured along the Big Five dimensions according to the PDI GPI Technical Manual (2001; see Table 5), the data was reanalyzed and Pearson's product-moment correlation analysis was conducted between score on the Extraversion factor of the Big Five and overall PBI rating. If a significant relationship was found between Extraversion and PBI rating, hierarchical multiple regression was performed to assess the variance attributed to the various subscales. Based on the hypothesis, the subscales were entered in the following order: the 3 motivation subscales: Desire for Achievement, Energy Level, and Initiative, then the remaining seven Extraversion subscales: Adaptability, Competitiveness, Desire for Advancement, Influence, Risk-Taking, Sociability, and Taking-Charge. Each step was entered using a stepwise (statistical) procedure in order to provide each type of item the statistical relationship with the criteria. No significant change in R^2 after entering the additional Extroversion subscales indicated that the remaining subscales did not explain any additional unique variance.

To test Hypothesis 3a, Pearson's product-moment correlation analyses were conducted between overall past behavior interview rating and each of the 9 super-factors of the GPI in order to evaluate the relationship between PBI performance and each of the GPI subscales. The correlation coefficients were corrected for unreliability in the predictor and criteria which may attenuate the relationship.

To test Hypothesis 4a, Pearson's product-moment correlation analysis was conducted between overall PBI rating and overall score on the Watson-Glaser to identify the correlation between cognitive ability and past behavior interview rating. The correlation coefficients were corrected for unreliability in the predictor and criteria which may attenuate the relationship.

In order to explore Research Question 1, once the GPI was realigned along the Big Five personality factors, the data were reanalyzed and Pearson's product-moment correlation analyses were conducted between the Big Five factors of personality and overall PBI rating. These results were compared to the results obtained by Huffcutt et al. (2001) as well as Conway and Peneno (1999). The results of the correlational analysis obtained from this study were graphically depicted in a table along with the results of the other two studies identified that explored this relationship as well in order to determine if the findings of this study are similar to the prior research findings.

In order to respond to Research Question 2 and explore the order of the correlations of the 4 LEI super-factors with overall past behavior interview rating, Pearson's product-moment correlation analyses were conducted between overall past behavior interview score and each of the 4 super-factors of the LEI. Pearson's product-moment correlation analysis with past behavior interview score. These correlation coefficients were arranged in order and evaluated in light of theoretical considerations

As in Research Question 2, in order to respond to Research Question 3 and provide the order of the correlations of the 3 Motivation subscales with PBI performance, first Pearson's r was evaluated between each of the three subscales and PBI ratings and then the correlation coefficients were arranged in order and evaluated in light of theoretical considerations.

CHAPTER 3

RESULTS

After a brief discussion of the descriptive statistics of the past behavior interview, results are reported in order by each hypothesis. Descriptive statistics for the measures involved in each hypothesis are provided in the beginning of each section. Following this, the research questions are explored in order as well.

Descriptive statistics for the PBI such as means, standard deviations, skewness and kurtosis for each of the competencies measured and overall PBI performance can be found in Table 7. The competency that received the highest average rating was Drive for Results (average rating: 3.66, $SD = .67$); participants also received high scores for Inspire Trust (average rating: 3.51, $SD = .56$) and Lead Courageously (average rating: 3.517, $SD = .572$). Participants received the lowest ratings on Display Global Perspective (average rating: 2.99, $SD = .543$). It should be noted that the Display Global Perspective competency had the smallest number of cases (that is, was rated the fewest number of times; $n = 696$, 67.9%); this was due to the fact that some participants were employed at organizations only operating within the United States and company officials decided not to have this competency measured in the assessment. Each competency has at least a few missing data points; one reason for a missing competency rating is occasionally an interviewer might not be able to collect sufficient information about the competency and its associated behaviors during the interview to provide a fair, accurate rating, and may choose to leave it blank instead. Additionally, occasionally organizations will send employees through the assessment center whose positions do not completely match the Business Unit Leader position and therefore the organization may choose not to have the candidate/participant be rated on a particular competency.

Experience and Past Behavior Interview Performance

Overall score on the LEI, the one of-a-kind assessment of an individual's leadership experience, was the measure of experience used to test the experience-related hypotheses. To provide a better understanding of this instrument descriptive statistics of the LEI are summarized in Table 8. The values listed as the means for the super-factors and subscales of the LEI are widely-varied and seemingly random. As discussed previously, a proprietary scoring algorithm is used to process item responses and calculate scores for the subscales which are summed to produce values for the super-factors. While such an algorithm undoubtedly allows for a consideration of a variety of factors and precise weighting of those factors according to their relative importance, the fact that no information is available to the public on how the item responses are weighted or even on the scales the resulting subscale values come from prevents much insightful analysis of the data. Considering this, the minimum and maximum score obtained and median scores for the subscales and super-factors are also reported in Table 8 to provide additional context for the data.

Some interesting information can be gleaned through analysis of the descriptive statistics of the LEI. For example, careful observation of the values of the subscale means reveals that those subscales which are more integral to business world success have higher means and generally appear to have a greater range of possible scores and a higher ceiling; that is, the maximum score achieved was higher than those subscales which are not as integral to professional success. For example, subscales such as Start-up Business and Extracurricular Activities produced mean scores of 17.56 and 38.17 respectively, and show top scores of 38.38 and 65.50. On the other hand, subscales such as Highly Critical/Visible Assignments or Initiatives and Interpersonally Challenging Situations—which tap the kinds of experiences

critical to business success had mean scores of 145.51 and 187.38 respectively, and the maximum scores obtained were 240.88 and 297.53. The fact that subscales tapping experiences more essential to success in the business world have higher averages and higher possible maximum scores means that these subscales are more heavily weighted in the calculation of the super-factors (and overall LEI score) since these scores are simply the sum of the subscales (and overall LEI score is the sum of the super-factors).

Alpha internal consistency reliability figures are also presented in Table 8 and should be briefly mentioned. Three of the four super-factors, General Management Experience, Overcoming Challenges and Adversity, and Risky and/or Critical Experiences demonstrate very strong internal consistency; producing alpha internal reliability coefficients of .88, .94, and .90, respectively. The remaining LEI super-factor, Personal and Career Related Experiences yields a considerably lower reliability coefficient of .64. Independently, this reliability coefficient would be considered completely appropriate and acceptable. However, the fact that the other LEI super-factors and the LEI as a whole (the entire instrument yields a reliability coefficient of .88) demonstrate such high internal reliability that the .64 alpha coefficient of Personal and Career Related Experience is noteworthy.

Hypothesis 1a stated that overall leadership experience score on the LEI would be significantly positively correlated to overall performance on the PBI. Table 9 shows a statistically significant relationship between overall LEI experience and overall PBI score ($r = .20, p < .01$), thus confirming the hypothesis. Table 9 also shows that Hypothesis 1b, which predicted that the LEI super-factor Personal and Career Related Experience would yield a significantly smaller relationship with overall PBI rating than the other three LEI super-factors. While the correlation between Personal and Career Related Experience and PBI performance (r

= .17, $p < .01$) was slightly smaller than the correlations between the remaining LEI super-factors and overall PBI score (General Management Experience, $r = .19$; Overcoming Challenge and Adversity, $r = .19$ and Risky and/or Critical Experiences, $r = .18$; $p < .01$ for each), Z tests for significant differences in correlations between the Personal and Career Related Management subscale and each of the other three LEI subscales revealed that none of the differences were statistically significantly.

Personality and Past Behavior Interview Performance

The next set of hypotheses involved the Global Personality Inventory; descriptive statistics for this measure are provided in Table 10. It is relevant to note that the means of each of the GPI super-factors are listed as “.00”. This is explained by first considering the fact that typically, the GPI super-factors do not receive an actual numeric score; they just serve as general grouping categories for which the subscales are organized to help participants understand and interpret results. Unlike the LEI, the GPI was not created in such a way that subscale totals can simply be summed “as-is” to yield a meaningful score for the super-factor. As discussed previously, the scale upon which each GPI subscale score is based varies depending on the number of items in the subscale; therefore in order to create a score for each super-factor from the average of the subscales within each super-factor weighting each subscale equally, subscale scores first had to be standardized by converting them to z scores. Z scores fit the standard normal distribution, which has a mean of 0 and a variance and standard deviation of 1 (Howell, 2002). With each subscale score transformed to a score that fits a standard distribution with a mean of 0, obviously the mean of those standardized subscales—the mean super-factor score—is 0 as well.

The alpha-reliability coefficients for the 9 GPI super-factors—ranging from .55 to .83—are listed in Table 10 as well. These coefficients seem to generally be on par with reliability data reported in the technical manual (GPI Technical Manual, 2001), although the .55 and .59 reliabilities calculated for Individual Work Orientation and Collective Work Orientation respectively are somewhat below the .85 and .75 alpha coefficients reported for those super-factors in the GPI Technical Manual (2001). While the reliability figures found for Individual Work Orientation and Collective Work Orientation are notably smaller than the reliabilities reported in the technical manual, they can still be considered to represent reliability

Motivation and Past Behavior Interview Performance

Hypotheses 2a through 2c specifically address one particular super-factor of the GPI—Motivation. Hypothesis 2a, which stated that Motivation would be significantly correlated to overall PBI performance was supported, as Motivation correlated significantly with overall PBI performance ($r = .19, p < .01$; Table 11). Hypothesis 2b stated that the Desire for Achievement subscale of Motivation would be significantly more strongly correlated with overall PBI performance than the remaining two subscales. Although Table 11 shows that the Desire for Achievement/PBI performance ($r = .18, p < .01$) correlation was stronger than Energy Level/PBI performance relationship ($r = .15, p < .01$) and the Initiative/PBI performance relationship ($r = .17, p < .01$), Z tests for significant differences in correlations revealed that none of the differences between the correlations were significantly different.

Hypothesis 2c was dependent upon the results of certain analyses in the study in order to fully be tested. The first condition, that overall Motivation as well as each of its subscales be significantly correlated with PBI performance was met as discussed previously and is depicted in

Table 11. The next condition, that upon restructuring GPI subscales under the Big Five factor structure, the Big Five dimension Extraversion must be significantly correlated with overall PBI performance, was met as well ($r = .13, p < .01$; see Table 14). Upon meeting these conditions, it was hypothesized that the 7 remaining subscales of the Big Five factor Extraversion would not add incremental validity to the Extraversion/PBI performance correlation beyond that of the three Motivation subscales of the GPI (Energy Level, Initiative, and Desire for Achievement). Hierarchical multiple regression was performed to test this hypothesis. The 3 Motivation subscales were entered in step one, and the remaining 7 subscales falling under the Extraversion factor were entered in Step 2. Due to sound theoretical evidence supporting a relationship between Motivation and PBI performance, it was hypothesized that the motivational elements of the Extraversion factor could be responsible for any significant relationship between Extraversion and PBI performance found in this study. As Table 12 shows, the three Motivation subscales added significant incremental validity to the Extraversion/overall PBI correlation ($\Delta R^2 = .039, p < .01$), and adding the remaining seven Extraversion subscales did not provide any significant incremental validity beyond that provided by the Motivation subscales.

Personality and Past Behavior Interview Performance

Hypothesis 3a stated that 8 of the 9 super-factors of the GPI (Motivation excluded) would not be significantly related to overall performance on the PBI. Data on the correlations between overall PBI rating and GPI super-factors is found in Table 13. As can be seen in the table, contrary to the hypothesis, a statistically significant relationship was found between PBI performance and each of the 8 GPI super-factors at the $p < .01$ level, although only one of the correlations was above .15 (Thinking, $r = .18$). The rest of the correlation coefficients were .15

or smaller: Facilitating Leadership, $r = .15$; Self-Management, $r = .15$; Interpersonal, $r = .14$; Collective Work Orientation, $r = .11$; Derailing Leadership, $r = -.11$; Individual Work Orientation, $r = .09$; and Planning & Execution, $r = .08$.

Cognitive Ability and Past Behavior Interview Performance

Hypothesis 4a posited that no statistically significant relationship would be found between cognitive ability as measured by overall score on the Watson-Glaser and overall PBI performance. Descriptive analyses of the performance of the 1,023 participants who took the Watson-Glaser revealed that with a possible scoring range from 1 to 80, the mean score was 66.36 with a *SD* of 6.9. Reliability information could not be calculated as item-level data was not available. The hypothesis that no significant relationship existed was confirmed; PBI performance and Watson-Glaser score were not significantly correlated ($r = -.03$, *ns*).

Research Question 1 inquired about the relationship between the GPI and past behavior interview performance when the items and subscales are reorganized under the Big Five personality dimensions, and then sought to compare those correlations with the findings of similar studies that have explored the relationship between PBI performance and the Big Five personality factors (i.e., Conway and Peneno, 1999; Huffcutt et al., 2001). When the GPI items and subscales were regrouped along the Big Five factors (again, see Table 5 for GPI subscale restructuring) and Pearson correlation analyses were re-run, the coefficient values were statistically significant. The correlations were: Ageeableness, $r = .14$; Conscientiousness, $r = .09$; Extraversion, $r = .19$; Neuroticism, $r = .13$; and Openness to Experience, $r = .14$ (all values significant at $p < .01$). These correlations and the correlations calculated between the Big Five

personality factors and overall PBI performance in Conway and Peneno (1999) and Huffcutt et al. (2001) are given in Table 14.

Research Question 2 sought to dive deeper into the nature of the relationship between experience and personnel selection interviews and uncover new information about the specific elements of experience that influence performance on the past behavior interview. The goal of this question was to explore which components of leadership experience were most strongly correlated with PBI performance and perhaps offer some hypothetical explanations for the reasons behind those relationships which could be tested at a later time. However, as the correlations between overall PBI performance and each of the LEI super-factors and were virtually identical (the correlations were .17, .18, .19, and .19), the opportunity for speculation on causes of differential correlations between the four LEI super-factors did not present itself. However, as no previous research has been conducted on the extent to which various components or subscales of experience influence the experience/PBI performance correlation, any information that sheds light on the nature of this relationship is beneficial.

Research Question 3 also sought to contribute new information to the literature by exploring in more depth the nature of the relationship between Motivation and PBI performance, a relationship which to this point had only been speculated upon. Unfortunately, as with the LEI super-factors, all 3 Motivation subscales were very similar in the strength of the correlation with past behavior interview performance (subscale correlations were .15, .17, and .18); therefore no one element of motivation was obviously more correlated with interview performance than any other. As was the case with the experience construct, since no previous research has been conducted on the correlation of various components of motivation (or motivation as a whole, for that matter) with PBI performance, this information is valuable to the field.

CHAPTER 4

DISCUSSION

The purpose of this study was to evaluate the relationship of a past behavior interview with several individual difference variables. Construct validation of this sort is important in order to help researchers and human resources professionals alike understand exactly what factors are being measured in the interviews they develop. This chapter discusses the results of each of the hypotheses tested in this study and their implications, as well as limitations of this study, and suggestions for future research.

Interpretations

The first set of hypotheses was oriented around the relationship between Experience and PBI performance. What made the examination of these constructs unique in this particular study was the instrument with which Experience was measured. While nearly every previous exploration of the Experience/PBI performance relationship in the past has simply used some version of an imprecise, unvalidated “number of years” measure (for example, work tenure, or number of years in graduate school) to assess experience, the LEI is a validated, reliable self-report measure designed specifically to assess leadership experience. Additionally, the four super-factors and 23 subscales within the LEI allow for an in-depth analysis of the correlation and a greater understanding of the specific components of the experience construct involved in the correlation with past behavior interviews.

The first hypothesis (Hypothesis 1a) initially set forth to establish a significant relationship between overall experience and PBI performance. To test this hypothesis, Pearson’s product-moment correlation analysis was conducted between overall LEI score and overall PBI

rating. The analysis confirmed the hypothesis; a statistically significant correlation was found between overall LEI score and average performance on the LEI ($r = .20, p < .01$).

Hypothesis 1b sought to provide new insight on the Experience/PBI relationship by exploring the content and organizational structure of the LEI subscales and super-factors and developing some ordered hypotheses about the relative strengths of the correlations of the super-factors with overall PBI performance based on consideration of the groupings of the super-factors and the nature of the PBI and the 12 competencies it measures. Specifically, because the Personal and Career Related Experience super-factor seemed to generally tap the types of experiences that were not as critical to success at the Business Unit Leader level and were therefore not as likely to be the types of experiences questions in the PBI would be devoted to, it was hypothesized that Personal and Career Related Experience would yield significantly smaller correlations with PBI performance than the remaining super-factors. Each super-factor correlated significantly with PBI score: General Management Experience, $r = .19$; Overcoming Challenge & Adversity, $r = .19$; Risky & Critical Experiences, $r = .18$; and Personal and Career Related Experiences, $r = .17$ (all values significant at $p < .01$); however, analyses revealed that none of the correlations were significantly different.

Despite the lack of support for Hypothesis 1b considerable valuable information was produced from this analysis. First of all, although the differences between the correlations were not statistically significant, the practical significance for the values lies in that the correlations were ordered in the hypothesized direction—Personal and Career Related Experiences did produce the lowest correlation with overall PBI performance; there could still be value in the theory upon which the hypothesis was based.

Practical significance of the data is also evident when it is considered that information on the factors that make up motivation and the extent to which they themselves correlate with PBI performance was previously unavailable; that being said, any data on the correlation between PBI performance and the super-factors comprising experience is valuable, regardless of whether or not the differences between them are statistically significant.

As Day and Carroll (2003) note, research findings on the relationship between experience and PBIs have been fairly mixed. Attempting to draw any kind of general conclusions about the true state of the Experience/PBI performance relationship must involve the consideration of a variety of factors. First, calculating the correlation of experience and past behavior interview performance with LEI measuring of the Independent Variable allows for at least some increased confidence in the accuracy of the correlation coefficient—despite its effect size—especially considering the alternative is a correlation coefficient with a small effect size calculated from an experience measure such as “number of years on the job.” Taking into considering the statistically significant correlation found between experience and PBI performance ($r = .20, p < .01$) in this study (yet still being mindful of the “small” effect size; Cohen, 1992), and previous research conducted on the topic, which as discussed earlier, generally points to the existence of a relationship, it seems likely that at least a small, positive significant correlation exists between experience and overall past behavior interview performance.

The second set of hypotheses evaluated the extent of the relationship between motivation and PBI performance. Testing these hypotheses also provided a unique contribution to the field as no research had directly tested the existence of a correlation between motivation and PBI performance. As detailed in the introduction, Taylor and Small (2002) argued theoretically for the existence of a statistically significant positive relationship between the two and even

provided what they considered indirect statistical evidence by suggesting that the statistically significant difference in the predictive validity of past behavior interviews over situational interviews found in the meta-analysis conducted in the study was attributable to the fact that past behavior interviews are able to meaningfully tap an individual's motivation level while situational interviews are not.

To accurately measure motivation, the Motivation super-factor of the Global Personality Inventory was isolated and treated as a unique construct. As no items on the GPI cross-load, there was no concern over measuring Motivation in this manner. Hypothesis 2a tested the zero-order correlation between overall Motivation and overall PBI rating. Correlation analysis revealed a statistically significant relationship between Motivation and PBI performance, $r = .19$, $p < .01$. As this was the first time the relationship between Motivation and performance on a past behavior interview has been statistically calculated, it is not a valuable use of time to even think about the true state of this relationship without considerable replication of this data. That being said, considering the fact that no other data—especially no contradictory data—exists, one cannot help but be optimistic about the existence of some degree of positive relationship, despite the small effect size of the coefficient. The fact that this correlation coefficient has not been previously calculated again makes clear the practical significance of such a statistic far outweighs its statistical significance. Finally, as was the case with having the LEI as the measure of experience in the hypotheses above, using a valid, reliable measure of Motivation provides at least some additional confidence in the accuracy of the statistic.

Hypothesis 2b attempted to make use of the multi-dimensional measure of the Motivation scale evaluated in this set of hypotheses and provide a more in depth understanding of the relationship between Motivation and PBI performance. As in Hypothesis 1b, careful

consideration of the content and key competencies of the Motivation subscales led to a hypothesis that one particular subscale would yield significantly higher correlations with overall PBI rating than the remaining two subscales.

The Desire for Achievement subscale directly measures an individual's motivation as it related to the workplace and career advancement. Based on this, Day and Carroll's (2003) theory that work motivation is a significant part of PBI performance, and that individuals high in Desire for Achievement are inevitably more likely to prepare for an interview which can lead to improved interview performance, Hypothesis 2b stated that Desire for Achievement would yield significantly higher correlations with PBI performance than the remaining subscales. Results showed Desire for Achievement was in fact more highly correlated with PBI performance ($r = .18, p < .01$) than both Energy Level ($r = .17, p < .01$) and Initiative ($r = .15, p < .01$), although further analyses revealed none of these differences were statistically significant.

Again, the practical significance of these coefficients far outweighs their statistical significance; similar to Hypothesis 1b, the Desire for Achievement subscale did not produce statistically significant correlations over the remaining two subscales, but again, the correlations were predicted in the right direction. The practical significance lies in the potential utility of the theory upon which Hypothesis 2b was based. The correlations have additional practical significance in that no other research has explored the correlation between past behavior interview performance and Motivation subscales.

Hypothesis 2c was set up to test if the motivation elements within the Big Five factor Extraversion were responsible for a majority of the variance accounted for in the Extraversion/PBI performance relationship. Multiple hierarchical regression was used to test this hypothesis, with the three motivation subscales being entered into the regression equation in the

first step, and the remaining 7 subscales being entered in step 2. This confirms the hypothesis that the three motivation subscales are in fact responsible for a majority of the variance accounted for in the correlation between Extraversion and PBI score, and that the 7 other subscales do not add any incremental variance. This finding has implications for research on the relationship between Extraversion and PBI performance; it puts forth at least the possibility that the Extraversion/PBI performance relationship, whatever the strength may be, could be attributable to aspects of motivation within the Extraversion construct. This topic will be revisited momentarily.

Hypothesis 3a tested the relationship between past behavior interview and the personality factors of the GPI. Surprisingly, all 8 factors tested yielded significant correlations with PBI performance. That being said, the correlations were quite small, and evaluation of the effect sizes in terms of the amount of variance explained (R^2) reveals that the practical significance of the correlations was minimal (Table 13). The highest percent of variance accounted for was by the Thinking super-factor, which only accounted for 4% of the variance. The remaining super-factors accounted for 3% of the variance or less. Despite the fact that the relationships were statistically significant, the results of this study combined with previous research lead to the tentative conclusion that little-to-no meaningful relationship exists between personality and past behavior interview performance.

The only hypothesis test that went completely as expected was hypothesis 4a. It was hypothesized that no significant relationship would be found between score on the Watson-Glaser and past behavior interview performance, which testing confirmed. Although previous research had reached obtained similar findings, this study confirmed that even at the Business

Unit Leader level, no meaningful relationship exists between cognitive ability and performance on a past behavior interview

Research Question 1 had the GPI restructured along the Big Five dimensions and the correlations between PBI performance and the Big Five factors were calculated and depicted in Table 14 along side two previous studies which had studied the relationship as well. Although all of the correlations between the Big Five factors and PBI performance in this study were statistically significant, the practical significance of the values is they all have small effects sizes and similar to the values found in the two previous studies, which between the two, only found only one significant correlation between PBI performance and a dimension of the Big Five. Quite interestingly, Huffcutt et al. (2001) found a significant relationship between PBI performance and Extraversion; and based on the results of the analyses conducted earlier in this study, one possible explanation for the significant result found in Huffcutt et al. (2001) is that the motivational elements of the Extraversion dimension were responsible for the significant correlation with PBI score.

Research Question 2 and Research Question 3 unfortunately were not able to be fully explored due to the highly similar correlations of the subscales of the LEI and Motivation with PBI performance. As the subscales were clearly measuring different aspects of Motivation and Experience, it is quite puzzling as to how the subscales' correlations with Past Behavior interview performance would group so tightly. Regardless, as stated earlier, as no prior research had been conducted on the correlations of the components of the experience construct with PBI score nor had any statistical research at all been conducted on the relationship between motivation and PBI performance, these findings still have considerable practical significance.

Limitations

One limitation of this research was that data was only collected for past behavior interview. As there is equal interest in the construct validity of the other form structured interview popular today—the situational interview, this study would have been able to contribute important information through construct validating the situational interview as well. Additionally, this study would have benefited by having access to some measure of the participants' job performance in order to evaluate the criterion validity of the past behavior interview.

Future Research

Future research on the past behavior interview should adhere to the structure guidelines set forth by Campion et al. (1997) and report as much information about the structure of the interview as possible to allow for organization of research findings and meta-analysis in the future. Additionally, future research should investigate the decision making process of the interviewee when choosing which past situation to relay in the interview. Obviously, the incident the participant chooses to share could impact ratings.

Table 1

Overview of Campion, Palmer, & Campion's (1997) Fifteen Elements of Interview Structure

Elements of Structure	Description
1. Base questions on a Job Analysis	Questions should be developed based on a thorough job analysis (often using the critical incident method); this is one of the most fundamental elements of a structured interview.
2. Ask exact same questions of each candidate	One of the most basic elements of structure; this can vary from where every interviewee must be asked the exact same question to where no two candidates get the same question.
3. Limit prompting, follow-up questioning, and elaboration on questions	Can range from no follow-up questions being allowed at all, to specific follow-up probing questions being allowed, to the interviewer asking any follow-up questions thought necessary to gather information.
4. Use better types of questions	Situational Questions versus Past Behavior Questions—Situational-type questions are future-oriented, i.e., they ask items such as, “How would you handle a situation if...” whereas past behavior-type questions are past-oriented, i.e., they ask items like, “Tell me about a time when you had to...”
5. Use longer interview or larger number of questions	This can refer to the actual question content or the overall number of questions being asked. Campion et al (1997) argue that longer questions and or interviews would be more valid because they elicit more information about the participant; however, this must have a ceiling effect upon reaching a point where less valuable answers are received due to participant fatigue.
6. Control ancillary information	Advanced knowledge of influential information such as cognitive ability or personality measure results have been shown to have significant impacts on interviewer ratings, so access to this information beforehand is discouraged; other examples of ancillary information are recommendations, resumes, previous interviews and performance on other assessment center measures (e.g., in-baskets, role plays).
7. Do not allow questions from the candidate until after the interview	Hypothetically this could lead to redirecting an interview in unpredictable ways, though little research has focused on this topic..

(table continues)

Table 1 (continued).

Elements of Structure	Description
8. Rate each answer or use multiple scales 673	Ratings can be given for each answer to an interview question or ratings can be held until the end and the dimensions being rated can be rated considering answers to multiple questions.
9. Use detailed anchored rating scales	Behaviorally anchored rating scales (BARS) are behavioral examples used to illustrate the various qualities of responses in answers; they reduce ambiguity and semantic differences.
10. Take detailed notes	Note-taking is thought to enhance structure by reducing memory decay and helping interviewers avoid primacy and recency effects
11. Use multiple interviewers	This can range from using one interviewer to a panel of interviewers involved in the same interview to multiple interviewers conducting the interviews separately.
12. Use same interviewer across all candidates	While interviewer ability has been shown to differ, using a highly structured interview and scoring guide greatly reduces impact of interviewer variability.
13. Do not discuss candidates or answers between interviews	This is related to ancillary information influencing interview scores as thoughts, insights, impressions from others could influence future interviewer ratings.
14. Provide extensive interviewer training	This is the most common way to improve interviews
15. Use statistical rather than clinical prediction	Research has found that objective, statistical methods of weighting and aggregating data is more effective than subjective methods used in clinical settings

Table 2

Assessment Methods' Ability to Measure Elements of Performance as Given by the Performance Determinants Model and Type of Performance Demonstrated

Measures of Job Performance	Elements of Performance Determinants Model Measured	"Maximal" or "Typical" Performance
Job knowledge Tests	DK	Maximal
Work Sample Tests	DK, PKS	Maximal
Supervisor ratings of Perf.	DK, PKS, M	Typical or Maximal
Situational Judgment Tests	DK, PKS	Maximal
Past Behavior Interview	DK, PKS, M	Typical
Situational Interview	DK, PKS	Maximal
In-Baskets	DK, PKS	Maximal

Note: DK = Declarative Knowledge; PKS = Procedural Knowledge and Skills; M = Motivation

Table 3

Super-factors and Subscales of the Leadership Experience Inventory (LEI)

Super-factor	Sub-scale
General Management Experiences	Strategy development Project management and implementation Business development and marketing Business growth Product development Start-up business Financial management Operations Support functions External relations
Overcoming Challenge and Adversity	Inherited problems and challenges Interpersonally challenging situations Downturns and/or failures Difficult financial situations Difficult staffing situations
Risky and/or Critical Experiences	High-risk situations Critical negotiations Crisis management Highly critical/visible assignments or initiatives
Personal and Career Related Experience	Self development Development of others International/cross-cultural Extracurricular activities

Table 4

Super-factors and Subscales of the Global Personality Inventory (GPI)

Super-factor	Sub-scale
Thinking	Thought Agility
	Innovation/Creativity
	Thought Focus
	Vision
Planning and Execution	Attention to Detail
	Work Focus
Facilitating Leadership	Taking Charge
	Influence
Derailing Leadership	Ego Centered
	Manipulation
	Micro-Managing
	Intimidating
	Passive-Aggressive
Interpersonal	Sociability
	Consideration
	Empathy
	Trust
	Social Astuteness
Motivation	Energy Level
	Initiative
	Desire for Achievement
Self-Management	Adaptability
	Openness
	Negative Affectivity
	Optimism
	Emotional Control
	Stress Tolerance
	Self-Confidence
	Impressing
Self-Awareness/Insight	
Individual Work Orientation	Independence
	Competitive
	Risk-Taking
	Desire for Advancement
Collective Work Orientation	Interdependence
	Dutifulness
	Responsibility

Table 5

Global Personality Inventory Subscales Restructured Along the Big Five Personality Factors, with Original GPI Super-factor

Big Five Personality Factor	GPI Subscale	GPI Super-factor
Agreeableness	Consideration	Interpersonal
	Empathy	Interpersonal
	Interdependence	Collective Work Orientation
	Openness	Self-Management
	Thought Agility Trust	Thinking
Conscientiousness	Attention to Detail	Planning and Execution
	Dutifulness	Collective Work Orientation
	Responsibility	Collective Work Orientation
	Work Focus	Planning and Execution
Extraversion	Adaptability	Self-Management
	Competitiveness	Individual Work Orientation
	*Desire for Achievement	Motivation
	Desire for Advancement	Individual Work Orientation
	*Energy Level	Motivation
	Influence	Facilitating Leadership
	*Initiative	Motivation
	Risk Taking	Individual Work Orientation
Sociability	Interpersonal	
Neuroticism	Taking Charge	Facilitating Leadership
	Emotional Control	Self-Management
	Negative Affectivity	Self-Management
	Optimism	Self-Management
	Self-Confidence	Self-Management
Openness to Experience	Stress Tolerance	Self-Management
	Independence	Individual Work Orientation
	Innovation/Creativity	Thinking
	Social Astuteness	Interpersonal
	Thought Focus	Thinking
Trait Composites	Vision	Thinking
	Ego-Centered	Derailing Leadership
	Impressing	Self-Management
	Intimidating	Derailing Leadership
	Manipulating	Derailing Leadership
	Micro-Managing	Derailing Leadership
	Passive-Aggressive	Derailing Leadership
Self-Aware	Self-Management	

Table 6

12 Competencies Measured in the Past Behavior Interview and Their Definitions

Competency	Definition
Think Strategically	Applies appropriate strategic logic to decisions and initiatives in one's area.
Innovate	Champions new ideas and initiatives and creates an environment that supports innovation.
Display Global Perspective	Establishes and promotes effective business operations across multiple countries and/or regions and coordinates appropriately with the broader global business.
Influence Others	Presents a compelling case for ideas and initiatives via an appropriately chosen strategy.
Engage and Inspire	Articulates and inspires commitment to a plan of action aligned with organizational mission and goals.
Build Talent	Ensures the availability and development of the talent needed to meet current and future organization goals.
Ensure Execution	Plans, executes, and improves work processes to ensure achievement of business goals.
Drive for Results	Demonstrates and fosters a sense of urgency and strong commitment to achieving goals.
Focus on Customers	Ensures the delivery of exceptional customer service.
Lead Courageously	Takes personal and organizational risks to do what is right and achieve organizational success, and supports others who do so.
Inspire Trust	Gains the confidence and trust of others through principled leadership and sound business ethics.
Adapt and Learn	Works to learn and develop from experience and from others, deals effectively with ambiguity and change, and adapts appropriately to new situations.

Table 7

Descriptive Statistics for the Past Behavior Interview Competencies and Overall Rating

Competency	<i>n</i>	%	<i>n</i> Missing	% Missing	Mean	Std.Dev.	Skewness	Std. Error Skewness	Skewness/ Std Error Skewness	Kurtosis	Std. Error Kurtosis	Kurtosis/ Std Error Kurtosis
Think Strategically	1004	98.1%	19	1.9%	3.319	.531	-.142	.077	-1.844	.013	.154	.084
Innovate	996	97.3%	27	2.7%	3.335	.516	.278	.077	3.610	.157	.155	1.013
Display Global Perspective	696	67.9%	327	32.1%	2.997	.667	-.329	.093	-3.538	.737	.185	3.984
Influence Others	1016	99.3%	7	.7%	3.226	.506	-.009	.077	-3.538	.379	.153	2.477
Engage & Inspire	1010	98.7%	13	1.3%	3.211	.481	.049	.077	.636	.248	.154	1.610
Build Talent	983	96.1%	40	3.9%	3.097	.523	-.117	.078	-1.500	.441	.156	2.827
Ensure Execution	1003	98.0%	20	2.0%	3.383	.484	.139	.077	1.805	-.178	.154	-1.156
Drive for Results	1015	99.2%	8	.8%	3.685	.543	.226	.077	2.935	-.206	.153	-1.356
Focus on Customers	965	94.3%	58	5.7%	3.448	.505	.209	.079	2.646	.247	.157	1.573
Lead Courageously	1008	98.5%	15	1.5%	3.517	.572	.044	.077	.571	-.220	.154	-1.429
Inspire Trust	1016	99.3%	7	.7%	3.531	.558	.302	.077	3.922	.122	.153	.797
Adapt & Learn	1014	99.1%	9	.9%	3.325	.473	.102	.077	1.325	.116	.153	.758
Overall Rating	1019	99.6%	4	.4%	3.350	.308	.157	.077	2.039	.788	.153	5.150

Note: Total *N* = 1023; possible range for competency ratings 1.0 to 5.0 (actual ratings are whole and half numbers only)

Table 8

Descriptive Statistics for the Super-factors and Subscales of the Leadership Experience Inventory (LEI)

Super-factor	Subscale	Mean	Min/ max	Median	Std. Dev.	Alpha Rlby	Skewness	Std. Error Skewness	Skewness/ Std Error Skewness	Kurtosis	Std. Error Kurtosis	Kurtosis/ Std Error Kurtosis
Overall L-ship Experience		1592.67	888.18/ 2582.60	1575.68	294.89	.88	.313	.076	4.12	-.125	.153	-0.82
General Mgmt Experiences		606.25	292.75/ 983.43	601.53	127.70	.94	.197	.076	2.59	-.336	.153	-2.20
	Strategy Dvlpmnt	82.66	39.50/ 138.00	81.63	19.55		.185	.076	2.43	-.501	.153	-3.27
	Project Mgmt & Implementation	104.17	47.30/ 160.25	103.88	23.62		-.049	.076	-0.64	-.620	.153	-4.05
	Business Development & Marketing	42.09	20.00/ 77.90	40.65	13.27		.330	.076	4.34	-.737	.153	-4.82
	Business Growth	45.03	20.70/ 77.65	44.45	12.81		.177	.076	2.33	-.835	.153	-5.46
	Product Development	19.76	8.10/ 32.40	19.95	6.63		-.022	.076	-0.29	-1.148	.153	-7.50
	Start-up Business	17.56	9.80/ 38.38	16.55	5.56		.939	.076	12.36	.762	.153	4.98
	Financial Mgmt	91.02	41.50/ 147.65	90.93	19.52		.047	.076	0.62	-.399	.153	-2.61
	Operations	70.32	29.23/ 105.18	71.43	15.37		-.215	.076	-2.83	-.604	.153	-3.95
	Support Functions	38.32	16.10/ 56.40	38.78	7.94		-.187	.076	-2.46	-.531	.153	-3.47
	External Relations	95.33	44.50/ 165.45	94.05	22.12		.348	.076	4.58	-.093	.153	-0.61

(table continues)

Table 8 (continued).

Super-factor	Subscale	Mean	Min/ max	Median	Std. Dev.	Alpha Relty	Skewness	Std. Error Skewness	Skewness/ Std Error Skewness	Kurtosis	Std. Error Kurtosis	Kurtosis/ Std Error Kurtosis
Overcoming Challenge & Adversity		416.55	206.80/ 679.80	414.08	77.49	.90	.321	.076	4.22	-.061	.153	-0.40
	Inherited Problems & Challenges	45.87	23.73/ 81.68	45.18	9.58		.505	.076	6.64	.336	.153	2.20
	Interpersonally Challenging Situations	187.38	94.90/ 297.53	186.40	31.23		.198	.076	2.61	-.019	.153	-0.12
	Downturns and/or Failures	33.75	15.60/ 60.50	32.83	9.17		.398	.076	5.24	-.396	.153	-2.59
	Difficult Financial Situations	70.41	35.70/ 125.73	68.88	16.98		.356	.076	4.68	-.299	.153	-1.95
	Difficult Staffing Situations	79.14	35.60/ 129.98	78.43	15.90		.221	.076	2.91	-.250	.153	-1.63
Risky and/or Critical Expers		364.43	190.20/ 621.80	357.93	80.03	.91	.338	.076	4.45	-.222	.153	-1.45
	High Risk Situations	70.12	33.80/ 119.55	69.05	15.69		.262	.076	3.45	-.309	.153	-2.02
	Critical Negotiations	106.07	52.65/ 188.35	104.20	26.17		.370	.076	4.87	-.230	.153	-1.50
	Crisis Mgmt	42.74	21.60/ 73.83	42.10	10.11		.366	.076	4.82	-.319	.153	-2.08
	Highly critical/visible assignments or initiatives	145.51	72.60/ 240.88	144.25	30.93		.252	.076	3.32	-.271	.153	-1.77
Personal & Career Related Experience		205.45	126.35/ 310.78	202.85	29.72	.64	.351	.076	4.62	.004	.153	0.03
	Self-Development	63.91	34.40/ 93.40	64.30	9.28		-.076	.076	-1.00	-.239	.153	-1.56
	Development of Others	61.61	25.80/ 77.60	62.60	8.57		-.927	.076	-12.20	1.186	.153	7.75
	International/ Cross- Cultural	41.76	23.20/ 89.65	38.63	15.09		.802	.076	10.55	-.168	.153	-1.10
	Extracurricular Activities	38.17	19.50/ 65.50	37.15	8.38		.421	.076	5.54	-.320	.153	-2.09

Note: Subscale scores are calculated from proprietary algorithm; Min/Max and median scores are given to allow for additional context and meaningful interpretation of values.

Table 9

LEI Super-factors and Overall LEI Score Correlations with PBI Overall Average

	1	2	3	4	5	6
1. PBI Overall Average	-	.19	.19	.18	.17	.20
2. General Management Experience		-	.86	.93	.60	.97
3. Overcoming Challenge & Adversity			-	.90	.62	.94
4. Risky/Critical Experiences				-	.59	.70
5. Personal/Career Related Experiences					-	.68
6. Overall LEI Score						-

Note: $N = 1019$. All values significant at the $p < .01$ level.

Table 10

Descriptive Statistics for the Super-factors and Subscales of the Global Personality Inventory (GPI)

Super-factor	Subscale	Mean	Std.Dev.	Alpha Reliability	Skewness	Std. Error Skewness	Skewness/ Std Error Skewness	Kurtosis	Std. Error Kurtosis	Kurtosis/ Std Error Kurtosis
Thinking		.000	.804	.81	-.256	.078	-3.28	2.825	.156	18.11
	Thought Agility	7.034	.963		-.590	.078	-7.56	2.567	.156	16.46
	Innovation/ Creativity	6.475	1.144		-.327	.078	-4.19	.578	.156	3.71
	Thought Focus	5.280	.780		-.138	.078	-1.77	1.424	.156	9.13
	Vision	6.537	.978		-.231	.078	-2.96	1.335	.156	8.56
Planning & Execution		.000	.873	.69	-.349	.078	-4.47	.693	.156	4.44
	Attention to Detail	5.314	1.288		-.206	.078	-2.64	.049	.156	0.31
	Work Focus	6.381	1.150		-.581	.078	-7.45	1.087	.156	6.97
Facilitating Leadership		.000	.895	.75	-.195	.078	-2.50	1.309	.156	8.39
	Taking Charge	7.608	1.038		-.342	.078	-4.38	1.274	.156	8.17
	Influence	6.406	.964		-.163	.078	-2.09	.527	.156	3.38
Derailing Leadership		.000	.711	.75	.521	.078	6.68	3.410	.156	21.86
	Ego Centered	2.646	.875		.024	.078	0.31	.043	.156	0.28
	Manipulation	2.868	1.049		.211	.078	2.71	.745	.156	4.78
	Micro-Managing	1.730	.784		.483	.078	6.19	1.243	.156	7.97
	Intimidating	1.826	.793		.634	.078	8.13	1.833	.156	11.75
	Passive-Aggressive	1.718	.811		.476	.078	6.10	1.589	.156	10.19
Interpersonal		.000	.712	.73	-.247	.078	-3.17	2.848	.156	18.26
	Sociability	6.343	1.476		-.546	.078	-7.00	.179	.156	1.15
	Consideration	7.561	1.104		-.273	.078	-3.50	1.104	.156	7.08
	Empathy	4.858	.812		-.244	.078	-3.13	.602	.156	3.86
	Trust	5.512	.745		-.697	.078	-8.94	3.733	.156	23.93
	Social Astuteness	5.324	.816		-.041	.078	-0.53	.550	.156	3.53

(table continues)

Table 10 (continued).

Super-factor	Subscale	Mean	Std.Dev.	Alpha Reliability	Skewness	Std. Error Skewness	Skewness/ Std Error Skewness	Kurtosis	Std. Error Kurtosis	Kurtosis/ Std Error Kurtosis
Motivation		.000	.866	.83	-.496	.078	-6.36	2.649	.156	16.98
	Energy Level	6.507	1.068		-.293	.078	-3.76	1.035	.156	6.63
	Initiative	6.372	.937		-.169	.078	-2.17	1.001	.156	6.42
	Desire for Achievement	6.341	.883		-.679	.078	-8.71	2.500	.156	16.03
Self-Mgmt		.000	.530	.71	.057	.078	0.73	1.263	.156	8.10
	Adaptability	5.487	.871		-.190	.078	-2.44	.632	.156	4.05
	Openness	4.768	.893		-.192	.078	-2.46	.663	.156	4.25
	Negative Affectivity	1.227	.736		1.046	.078	13.41	4.789	.156	30.70
	Optimism	6.838	1.045		-.565	.078	-7.24	1.414	.156	9.06
	Emotional Control	4.900	.982		-.552	.078	-7.08	.436	.156	2.79
	Stress Tolerance	5.119	1.194		-.092	.078	-1.18	-.199	.156	-1.28
	Self-Confidence	5.392	.712		-.305	.078	-3.91	1.399	.156	8.97
	Impressing	3.755	.725		.006	.078	0.08	.201	.156	1.29
	Self-Awareness/Insight	7.278	.891		-.489	.078	-6.27	3.298	.156	21.14
Individual Work Orientation		.000	.650	.55	.035	.078	0.45	.417	.156	2.67
	Independence	3.270	.843		.169	.078	2.17	.772	.156	4.95
	Competitive	4.590	1.094		-.083	.078	-1.06	-.197	.156	-1.26
	Risk-Taking	6.256	1.131		-.174	.078	-2.23	.297	.156	1.90
	Desire for Advancement	3.898	1.014		-.087	.078	-1.12	-.069	.156	-0.44
Collective Work Orientation		.000	.748	.59	-.461	.078	-5.91	2.569	.156	16.47
	Interdependence	5.439	.936		.008	.078	0.10	.241	.156	1.54
	Dutifulness	6.170	.882		-.414	.078	-5.31	.171	.156	1.10
	Responsibility	5.995	.748		-.828	.078	-10.62	2.846	.156	18.24

Note: N = 987.

Table 11

Overall Motivation and Motivation Subscales Correlation with PBI Overall Average

	1	2	3	4	5
1. PBI Overall Avg	-	.17	.15	.18	.19
2. Energy Level		-	.60	.63	.86
3. Initiative			-	.65	.87
4. Desire for Ach				-	.88
5. Overall Motivation					-

Note: $N = 987$. All values significant at the $p < .01$ level.

Table 12

Hierarchical Regression of Motivation Subscales and 7 Remaining Subscales Under the Big Five Factor Extraversion

Step and Antecedent	R	R^2	ΔR^2	β
Step 1. Motivation subscales	.196**	.039	.039**	
Energy Level				.085*
Initiative				.026
Desire for Ach				.111*
Step 2. Other Extvn subscales	.228	.052	.013	
Adaptability				-.032
Competitiveness				-.069
Desire for Advcmnt				.047
Risk Taking				.085*
Sociability				.020
Influence				.077
Taking Charge				.054

Note: $N = 987$. * $p < .05$; ** $p < .01$.

Table 13

GPI Super-factors Correlation with PBI Overall Average

	1	2	3	4	5	6	7	8	9	R ²
1. PBI Overall Avg	-	.18	.08	.15	-.11	.14	.15	.09	.11	
2. Thinking		-	.40	.62	-.24	.59	.65	.35	.51	.03
3. Planning/Execution			-	.34	-.11	.38	.45	.13	.56	.006
4. Facilitating Leadership				-	-.05	.55	.56	.52	.44	.02
5. Derailing Leadership					-	-.37	-.29	-.30	-.42	.01
6. Interpersonal						-	.65	.18	.60	.02
7. Self-Management							-	.31	.60	.02
8. Individ Work Orient								-	.05*	.008
9. Collectve Work Orient									-	.01

Note: N = 983. All values significant at the .01 level unless indicated by an asterisk, indicating nonsignificance.

Table 14

Big Five Factor Correlations with PBI Performance from Curren Study, Conway and Peneno (1999) and Huffcutt et al. (2001)

	GPI Items Regrouped under Big Five Factors N = 983	Conway and Peneno (1999) N = 179	Huffcutt et al. (2001) N = 93
1. Agreeableness	.14*	-.02	-.01
2. Conscientiousness	.09*	.00	.08
3. Extroversion	.19*	.10	.30*
4. Neuroticism	.13*	.07	.05
5. Openness to Exper	.14*	.10	.06

Note: * $p < .01$.

APPENDIX
SAMPLE QUESTIONS

Sample LEI Questions

Example of first type of items:

Please indicate how many times you have experienced the events the following events listed below:

- Was responsible for resolving discrepancies

1—never had that experience

2—experienced the event 1 or 2 times

3—experienced the event 3 or 4 times

4—experienced the event 6 or more times.

Example of second type of items:

“Please indicate the number of times you have experienced the following events in each of the roles below:”

- “I was involved in a project and or team that included participants from a number of organizational units or functions throughout the organization”

Contributor,--participates by providing support or assistance

Management—manages the relevant work effort or maintains key relationships

Lead strategist—sets the overall direction or has the overall responsibility

1—no experience with this situation in this role

2—experienced this situation in this role 1 to 2 times

3—experienced this situation in this role 3 or more times

Example of third type of items:

“Please indicate the amount of experience you have (in number of months) in each of the following activities:”

- “worked on an international experience”

How much experience at this activity do you have?

1—never experienced this type of activity

2—1 to 12 months of experience with this type of activity

3—13 to 24 months of experience with this type of activity

4—experienced the activity for 25 months or longer

Sample GPI Items

(Rated on a 1 to 5 Likert-type scale, 1 being *strongly agree* to 5 being *strongly disagree*)

Scale	Sample Items
Consideration	I like to little things for people to make them feel good
Trust	I believe other people are usually honest with me
Interdependence	I tend to put group goals first and individual goals second
Reliability	I can be relied on to do what is expected of me
Adaptability	For me, change is exciting

REFERENCES

- Ayres J. & Crosby, S. (1995). Two studies concerning the predictive validity of the personal report of communication apprehension in employment interviews. *Communication Reports, 12*, 145-151.
- Barrick, M.R. & Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.
- Caldwell, D.F. & Burger, J.M. (1998). Personality characteristics of job applicants and success in screening interviews. *Personnel Psychology, 51*, 119-136.
- Campion, M.A., Pursell, E.D., & Brown, B.K. (1988). Structured interviewing: Raising of the psychometric properties of the employment interview. *Personnel Psychology, 41*, 25-42.
- Campion, M.A., Campion, J.E., & Hudson Jr., P.J. (1994). Structured interviewing: A note on incremental validity and alternative question types. *Journal of Applied Psychology, 79*, 998-1002.
- Campion, M.A., Palmer, D.K., & Campion, J.E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50*, 655-702.
- Cohen, J. (1992). Fuzzy methodology. *Psychological Bulletin, 112*, 409-410.
- Conway, J.M & Peneno, G.M. (1999). Comparing structured interview question types: Construct validity and applicant reactions. *Journal of Business and Psychology, 13*, 485-506.
- Day, A.L. & Carroll, S.A. (2003). Situational and patterned behavior description interviews: A comparison of their validity, correlates, and perceived fairness. *Human Performance, 16*, 25-47.
- Fletcher, C. (1987). Candidate personality as an influence on selection interview assessments. *Applied Psychology: An International Review, 36*, 157-162.
- Guion, (1976). Recruiting, selection, and job placement. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 77-82). Chicago, IL: Rand McNally.
- Harris, M.M. (1989). Reconsidering the employment interview: A review of recent literature and suggestions for future research, *Personnel Psychology, 9*, 691-726.
- Hollingworth, H.L. (1922). The personal interview. In *Judging human character* (pp. 60-73). New York: D Appleton & Co.
- Hovland, C.I. & Wonderlic, E.F. (1939). Prediction of industrial success from a standardized interview. *Journal of Applied Psychology, 23*, 537-546
- Howell, D.C., & Crockett, C., et al. (Eds.). (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Wadsworth.

- Huffcutt, A.I., Roth, P.L., & McDaniel M.A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity, *Journal of Applied Psychology*, 81, 459-473.
- Huffcutt, A.I., Weekley, J.A., Wiesner, W.H., DeGroot, T.G., & Jones, C. (2001). Comparison of situational and behavior description interview questions for higher level positions. *Personnel Psychology*, 54, 619-644.
- Jackson, D.N., Peacock, A.C., & Holden, R.R. (1982). Professional interviewers' trait inferential structures for diverse occupational groups. *Organizational Behavior & Human Performance*, 29, 1-20.
- Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67, 577-580.
- Janz, T. (1989). *The employment interview: Theory, research, and practice* (pp. 158-168). Thousand Oaks, CA: Sage Publications.
- Janz, T., Hellervik, L., & Gilmore, D.C. (1986). *Behavior description interviewing: New, accurate, cost effective*. Upper Saddle River, NJ: Prentice-Hall.
- Keith, R.D., Engelkes, J.R., & Winborn, B.B. (1977.) Employment-seeking preparation and activity: An experimental job-placement training model for rehabilitation clients. *Rehabilitation Counseling Bulletin*, 21, 159-165.
- Krajewski, H.T., Goffin, R.D., McCarthy, J.M., Rothstein, M.G., & Johnston, N. (2006). Comparing the validity of structured interviews for managerial level employees: Should we look to the past, or focus on the future? *Journal of Occupational and Organizational Psychology*, 79, 411-432.
- Latham, G.P., & Saari, L.M. (1984) Do people do what they say? Further studies of the situational interview. *Journal of Applied Psychology*, 69, 569-573.
- Latham, G.P., Saari, L.M., Pursell, E.D., & Campion, M.A. (1980). The situational interview. *Journal of Applied Psychology*, 65, 422-427.
- Latham, G. P. & Skarlicki, D.P. (1995). Criterion-related validity of the situational and patterned behavior description interviews with organizational citizenship behavior. *Human Performance*, 8, 67-80.
- Locke, E.A. (1968). Effects of knowledge of results, feedback in relation to standards, and goals on reaction-time performance. *American Journal of Psychology*, 81, 566-574.
- Mauer, T.J., Solamon, J.M., Andrews K.D., & Troxtel, D.D. (2001). Interviewee coaching, preparation strategies, and response strategies in relation to performance in situational employment interviews: An extension of Mauer, Solamon, and Troxtel (1998). *Journal of Applied Psychology*, 86, 709-717

- Mauer, T.J., Solamon, J.M., & Troxtel, D.D. (1998). Relationship of coaching with performance in situational employment interviews. *Journal of Applied Psychology*, 83, 128-136.
- McCloy, R.A., Campbell, J.P., & Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology*, 79, 493-505.
- Mayfield, E.C. (1964). The selection interview: A re-evaluation of published research. *Personnel Psychology*, 17, 239-260.
- McMurray, R.N. (1947). Validating the patterned interview. *Personnel*, 23, 263-272.
- McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., & Maurer, S. (1994). The validity of employment interview: A comprehensive meta-analysis. *Journal of Applied Psychology*, 79, 599-617.
- Motowidlo, S.J., Carter, G.W., Dunnette, M.D., Tippins, N., Werner, S., Burnett, J.R., & Vaughn, M.J. (1992). Studies of the structured behavioral interview. *Journal of Applied Psychology*, 77, 571-587.
- Orpen, C. (1985). Patterned behavior description interviews versus unstructured interviews: A comparative validity study. *Journal of Applied Psychology*, 70, 774-776.
- Paunonen, S.V., Jackson, D.N., & Oberman, S.M. (1987). Personnel selection decisions: Effects of applicant personality and the letter of reference. *Organizational Behavior and Human Decision Processes*, 40, 96-114.
- Personnel Decisions International. (2001). *Technical manual of the Global Personality Inventory*. Minneapolis, MN: Author.
- Personnel Decisions International (2006). *Technical manual of the Leadership Experience Inventory*. Minneapolis, MN: Author.
- Personnel Decisions International. (2008). *Global Personality Inventory: Facts on validity and reliability*. Minneapolis, MN: Author.
- Posthuma, R.A., Morgeson, F.P., Campion, M.A. (2002). Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time. *Personnel Psychology*, 55, 1-81.
- Pulakos, E.D. & Schmitt, N. (1995). Experience-based and situational interview questions: Studies of validity. *Personnel Psychology*, 48, 289-308.
- Robertson, I.T. & Makin, P.J. (1986). Management selection in Britain: A survey and a critique. *Journal of Occupational Psychology*, 59, 45-57.
- Schmidt, F.L. (1988). The problem of group differences in ability test scores in employment selection. *Journal of Vocational Behavior*, 33, 272-292.

- Schmitt, N. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. *Personnel Psychology*, 29, 79-101.
- Scott, W.D. (1915). The scientific selection of salesmen. *Advertising and Selling*, 25, 5-6.
- Smith, P.C. & Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47, 149-155.
- Spiegel, W.R. & James, V.A. (1958). Trends in recruitment and selection practices. *Personnel*, 35, 42-48.
- Stevens, W. & Tornatzky, L. (1976). The effects of a job-interview skills workshop on drug-abuse clients. *Journal of Employment Counseling*, 13, 156-163
- Taylor, P.J. & Small, B. (2002). Asking applicants what they would do versus what they did do: A meta-analytic comparison of situational and past behaviour employment interview questions. *Journal of Occupational and Organizational Psychology*, 75, 277-294.
- Ulrich, L. & Trumbo, D. (1965). The selection interview since 1945. *Psychological Bulletin*, 63, 100-116.
- Wagner, R. (1949). The employment interview: A critical summary. *Personnel Psychology*, 2, 17-46.
- Walters, L.C., Miller, M.R., & Ree, M.J. (1993). Structured interviews for pilot selection: No incremental validity. *International Journal of Aviation Psychology*, 3, 25-38.
- Webster, E.C. (1964). *Decision making in the employment interview*. Montreal, Quebec, Canada: Eagle.
- Weekley, J.A. & Guier, J.A. (1987). Reliability and validity of the situational interview for a sales position. *Journal of Applied Psychology*, 72, 484-487.
- Wonderlic, E.F. (1937). Personnel as a control function. *Personnel*, 14, 31-41.
- Wright Jr., O.R. (1969). Summary of research on the selection interview since 1964. *Personnel Psychology*, 22, 391-413.