Graduate Theses and Dissertations

Graduate School

2007

Construct validity of situational judgment tests: An examination of the effects of agreeableness, organizational leadership culture, and experience on SJT responses

Jonathan Adam Shoemaker
University of South Florida

Construct Validity of Situational Judgment Tests:
An Examination of the Effects of Agreeableness,
Organizational Leadership Culture, and Experience on SJT Responses

by

Jonathan Adam Shoemaker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
College of Arts & Sciences
University of South Florida

Major Professor: Michael Brannick, Ph.D.
Tammy Allen, Ph.D.
Walter Borman, Ph.D.
Marcia Finkelstein, Ph.D.
Joseph Vandello, Ph.D.
Interdepartmental Chair: Michael Reid, Ph.D.

Date of Approval
April 11, 2007

Keywords: practical intelligence, tacit knowledge, low-fidelity simulation, organizational culture, path-goal theory, participative leadership, work experience, management, construct explication, job knowledge, personality, agreeableness, personnel selection

Dedication

This research and obtaining my Ph.D. has been a longer process than I ever expected. I would like to thank all my friends and colleagues at the University of South Florida, as well as all the professors who made me think, and all the undergraduates that made me feel like I was making a difference.

This paper is dedicated to my wonderful wife Amy Rothman for her support and partnership with this project, during our years in Florida, and always. You are my favorite! I hope I write something more interesting for you next time. Finally, thanks to Beth "Nana" Rothman for helping me submit this paper and "fanning" the PC.

Acknowledgements

I would like to thank my committee for their participation and support throughout this project. I must especially thank my major professor, Dr. Michael Brannick, who was patient and informative to the "Nth" degree even on days when I was only performing at N-1. Some of my best memories of graduate school will be the theoretical discussions we had about how this project might turn out and what it might all mean.

I must also take this opportunity to thank my colleagues at Watkins Motor Lines, Inc. for their support of and participation in the pilot portion of this project, my colleagues at Verizon Wireless (notably Ashley Gray and Jon Canger) for their support of and participation in the experimental portion of this project, and the students of the University of South Florida for participating in this project at every level.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

iv

Construct Validity of Situational Judgment Tests:

An Examination of the Effects of Agreeableness,

Organizational Leadership Culture and Experience on SJT Responses

Jonathan Adam Shoemaker

ABSTRACT

Numerous factors are likely to influence response patterns to situational judgment tests, including agreeableness, leadership style, impression management, and job and organizational experience. This research presents background information and research on situational judgment tests and several constructs hypothesized to influence situational judgment test responses. A situational judgment test and manipulations to influence response patterns were developed and piloted with a small sample of management professionals and undergraduate students. Larger samples of management professionals and undergraduate students participated in the experimental research. Participants were asked to imagine that they are applying for a job. Each participant was presented with background information about a fictitious company, describing a company as either highly Participative/Supportive or highly Directive/Achieving in its leadership culture. A third description provided no information about leadership culture to serve as a control. Participants responded to a situational judgment test consisting of some commercially developed items and some new items. Then participants responded to an inventory comprised of items that measure the factors hypothesized to influence response patterns,

specifically Agreeableness and Experience. Significant differences in response patterns

were determined to be attributable to the Agreeableness and Experience variables, and

the Leadership Culture manipulations, as well as the interaction between Experience and

the Leadership Culture manipulations. No significant differences were clearly

attributable to the Agreeableness by Leadership Culture interaction. The ramifications of

these findings are discussed and recommendations for future research are presented.

Construct Validity of Situational Judgment Tests:

An Examination of the Effects of Agreeableness,

Organizational Leadership Culture and Experience on SJT Responses

The ultimate purpose of selection testing is to provide a prediction of performance that is accurate and inexpensive. Accuracy is important because well-designed tests should be scientifically developed and validated to fairly exclude unqualified candidates and to select those that will perform best. Ideally, tests should be highly accurate in the settings for which they were developed, and they should be easily transferable to new settings. Affordability is desirable because for applied purposes, it is difficult to implement a high-priced, cumbersome test that requires a major commitment of resources. Situational judgment tests (SJTs) may fit both criteria. SJTs are face valid; that is, the stimuli and responses required by these tests appear to be related to activities that are performed on the job. High face validity coupled with good predictive validity and the relatively inexpensive nature of such tests has made SJTs a popular selection technique. However, the construct(s) being measured by situational judgment tests has received very little attention, and our understanding of what the tests actually measure is very shallow.

This research strives for better understanding the nature of the constructs that affect responses to situational judgment tests. It was hypothesized that multiple factors influence responses to situational judgment tests. Among these are: Agreeableness, as

1

described by McCrae and Costa (1985) in their Five Factor Model, The Leadership Culture of an organization, specifically in terms of manipulations based on Path-Goal Theory, as proposed by House (1971), motives of the test-taker (intent to present a favorable impression), and job experience. There is already great deal of evidence that situational judgment tests are related to cognitive ability, experience, and even personality. However, it is still unclear what other constructs or variables may provide incremental variance in relation to SJT responses, and how all these constructs work together to affect response choices. This study proposed that a nomological representation of responses to situational judgment items, specifically those that address dealing with subordinates, should also include agreeableness on the part of the respondent, impression management on the part of the respondent, the leadership characteristics of the organization, and job experience, as well as interactions between these constructs.

This paper begins by presenting background information and research on situational judgment tests themselves, their reliability and validity, and how they are used. Then the paper describes those constructs hypothesized to influence SJT responses and the rationale for their inclusion in the current research. First, Agreeableness (and its opposite, Antagonism) is described. Next, the concept of an organizational leadership culture will be discussed, with emphasis on Path-Goal Theory (House, 1974). Some attention will be paid to the importance of impression management in the workplace – specifically, respondents' attempts to display congruence between their own style of

2

leadership and the organization's leadership culture. Finally, the concept of experience will be explored.

A rudimentary explanation of the study is described below so that the literature review and hypotheses that follow can be placed in context. Participants from two settings were included in the research: the professional or "Manager" sample included management employees, while the novice or "Student" sample included college students. Participants were asked to imagine they are applying for a job. Each participant was presented with one of three sets of background information about a fictitious company. One set of information clearly indicated that the company is highly Participative and Supportive in its leadership culture. Another set of information clearly indicated that the company is highly Directive and Achievement-Oriented in its leadership culture. The third set of information contained no information about the leadership culture and acted as a control. Participants then responded to a situational judgment test. Finally, participants responded to an inventory that included items that address personality and experience, as well as manipulation check items.

Simply put, participants were expected to respond to the situational judgment test differently based on the variables mentioned above: notably, agreeableness on the part of the respondent, impression management on the part of the respondent, the leadership characteristics of the organization (as operationalized by the manipulations of background information), and experience in the job of manager.

3

Situational Judgment Tests

What is a "Situational Judgment Test"?

A situational judgment test is essentially a work sample test that presents job-specific problems to which a test taker must choose from among several possible solutions. The concept of using a work sample as an assessment tool has been around since the beginnings of the field of personnel psychology. Work samples have actually demonstrated the highest validity for any individual assessment method, even higher than that of cognitive ability tests (Schmidt & Hunter, 1998).

Tests that present written situations to simulate actual work were first used in selection in the 1920's (McDaniel & Nguyen, 2001). The first instrument to be classified as a situational judgment test was developed in the late 1950's to help select supervisors (Mowry, 1957). Situational judgment tests are a natural progression from more elaborate simulation techniques such as assessment centers and other work simulations. Because the situational fidelity is compromised when the respondent is not physically in the situation, SJTs have been referred to as "low-fidelity simulations," a term coined by Motowidlo, Dunnette, and Carter (1990). Situational judgment tests have been criticized as "a rather extreme attempt to streamline assessment centers" (Borman *et al.,* 1997, p. 314). However, other authors suggest that situational approaches to selection techniques are one of the most significant developments in selection research (Robertson & Smith, 1989).

Situational judgment tests continue to be popular because they appear to be highly work-related, more so than many personality or cognitive ability tests. Therefore, they

4

demonstrate better face validity and more favorable reactions from respondents (Ryers & Connerly, 1993). SJTs have also been shown to have less adverse impact than tests of cognitive ability (*e.g.* Oswald, et al., 2004; Callinan & Robertson, 2000; Chan & Schmidt, 1997; Weekley & Jones, 1997; Motowidlo & Tippins, 1993).

SJTs are characterized by a stem/response format (McDaniel & Nguyen, 2001). Each item begins with a stem that presents a work-related situation. Then, a series of response options are presented. Stems and responses can vary in fidelity, length and complexity (McDaniel & Nguyen, 2001). Typical SJTs are presented and completed in paper and pencil format, although a few SJTs use a video format in which respondents watch videotaped scenarios and choose from a set of videotaped or written responses (Weekley & Jones, 1997). Most situational judgment tests are developed for specific companies or for specific jobs or job families (Hanson & Ramos, 1996). There are few commercially available SJTs. Two notable exceptions are "The *ProveIt!* Manager," published by Kenexa, Inc., and the "Supervisory Skills Inventory" (SSI), published by gNeil, Inc., though only portions of these two tests include situational judgment items. Items from these tests were included in this research and information about them is elaborated below.

SJTs often consist of a smaller number of items than are commonly seen on cognitive ability or personality instruments. Many SJTs consist of as few as 20 item stems (Hanson & Ramos, 1996). This is primarily because every item requires the respondent to read detailed situational stems before responding.   Response choices

5

usually occur along a continuum, ranging from behavior(s) considered most effective to those considered least effective.

SJTs are most often developed from critical incidents (McDaniel & Nguyen, 2001). Next, response options are created when subject matter experts and/or novices unfamiliar with the job generate lists of effective and ineffective reactions to the described situation. Finally, scoring keys are developed, either rationally, by asking experts to rate the effectiveness of each response, or empirically by having participants take the test and comparing their scores to some external criterion, such as performance (Hanson & Ramos, 1996).

A widely accepted technique for situational judgment test instructions is to ask for both an effective and ineffective response, as introduced by Motowidlo, Dunnette and Carter (1990). This method provides information not only on effective performance, but also on the ability to avoid those most severely ineffective behaviors (Hanson & Ramos, 1996). Instructions for situational judgment tests may also address a respondent's potential or actual behavior. For example, one of the most popular scoring formats, as proposed by Motowidlo, Dunnette and Carter (1990) is for respondents to indicate which response choices are the most effective and which the least effective. Motowidlo and McDaniel (2005) referred to this scoring format as "knowledge instructions," as contrasted with "behavioral tendency instructions," that ask respondents to indicate which response choice is most like them and which is least like them. Other researchers have called these opposing instructional formats "should do" and "would do," respectively (Ployhart & Ehrhart, 2003). Findings suggest that "would do" or behavioral

6

tendency instructions show better reliability and validity than "should do" or knowledge instructions (Ployhart & Ehrhart, 2003). However, it is also likely that "should do" instructions are more appropriate for addressing maximal performance such as what would be expected from job applicants (Ployhart & Ehrhart, 2003). "Should do" and knowledge instructions appear to allow less dissimulation, and do not inflate scores when respondents are motivated to fake (McDaniel & Nguyen, 2001), such as they might be when applying for a job. The instructional format used also has implications for how much responses are influenced by cognitive ability vs. non-cognitive traits. SJTs with knowledge instructions tend to be more highly correlated with general cognitive ability; SJTs with behavioral tendency instructions tend to be more highly correlated with non-cognitive traits such as Conscientiousness and Agreeableness (McDaniel, Grubb & Hartman, 2003).

Situational judgment tests can vary widely in length, instructions and stem/response format. Regardless of the format, SJTs are intended to predict performance; however it remains unclear exactly what construct(s) SJTs truly are measuring.

There has been a rift in the literature as some researchers suggest situational judgment is a truly unique construct, also referred to as tacit knowledge (Sternberg *et al*., 1995) or procedural knowledge. Other researchers consider situational judgment tests to be merely another method of testing already well-established constructs such as job knowledge, work experience, or even general cognitive ability (Stevens & Campion, 1999; Chan & Schmidt, 1997; Schmidt & Hunter, 1993).

7

Sternberg and his associates (1995) suggested that SJTs measure a unique construct that is not related to general cognitive ability, or even job knowledge, as it is traditionally operationalized. Wagner (1987) called this construct tacit knowledge, a practical intelligence that is never described formally or taught directly within an organization. According to Sternberg's triarchic theory of intelligence, practical intelligence exists outside of the traditional sphere of general cognitive ability (Sternberg, 1985). Tacit knowledge is typically procedural and goal oriented (Sternberg & Wagner, 1993). Further, tacit knowledge is acquired without formal instruction from others (Sternberg *et al.*, 1995). Sternberg has equated practical or tacit knowledge with "street smarts," "learning the ropes" and "common sense" (Wagner & Sternberg, 1991, p. 1).

One major drawback to tacit knowledge is plain from its very name: unspoken, unofficial information transfer is very difficult to measure. The result is that researchers still write about tacit knowledge as a theoretical construct while also describing it as synthetic, intuitive, and not easy to operationalize (Styhre, 2004).

Chan and Schmidt (1997) argued that a test of situational judgment is simply a method of measuring multiple job-relevant skills and abilities. Studies of biodata, interviews and assessment centers have demonstrated that these are methods of measuring common constructs, not constructs in and of themselves (Schmidt & Rothstein, 1994). The same is almost certainly true of SJTs.

The balance of opinion now suggests that whether tacit knowledge exists or not is immaterial. Current theory considers the situational judgment test to be a style of measurement that taps numerous constructs, not a single, specific construct. However,

8

little research has been performed to show how the constructs measured by SJTs are affected by other constructs (Motowidlo & McDaniel, 2005).

Reliability of Situational Judgment Tests

Several researchers have developed unique situational judgment instruments. Most report reliability coefficients in the form of internal consistency. This statistic estimates the correlation that would be observed if the examinees took another test 'just like this one' and the correlation between the (often hypothetical) alternate forms was computed. Table 1 shows internal consistency reliabilities (Cronbach's *alpha*) for SJTs used in research. A few studies reported additional measures of reliability, in lieu of, or in addition to, internal consistency. The additional reliability coefficients included in the table are inter-rater reliability, test-retest reliability, and alternate forms reliability.

Table 1: Reliability Estimates of Situational Judgment Tests in Various Studies

Study	N	N of SJT Items	Reliability Coefficient(s)
Lee, Choi & Choe (2005)	498	16	= .13-.64*
Oswald *et al.* (2004)	634	57	= .85
Ployhart *et al.* (2003)	5325	10	= .46-.62*
Ployhart & Ehrhart (2003)	84; 23	5	= .36; test-retest = .63
Chan & Schmidt (2002)	160	8	= .73; alternate forms = .76
Clevenger *et al.* (2001)	412	39	= .63-.82*
Motowidlo, Dunnette & Carter (1990)	252	58	Mean inter-rater Reliability = .95
Weekly & Jones (1999)	1884	34	= .73
Clause *et al.* (1998)	377	33	Alternate forms = .70-.77*

*Ranges of reliability are shown if more than one sample or form was used in the research

As illustrated above, internal consistency reliability coefficients have ranged from poor to acceptable, according to the current standard of acceptable reliability for use in research (.70; Nunnally & Bernstein, 1994). The principles of test theory state that a larger number of items will increase reliability. However, one of the tests with the fewest

9

items and the smallest number of participants still demonstrated acceptable internal consistency reliability (Chan & Schmidt, 2002).   This may be because the items addressed one specific construct instead of numerous situational judgment constructs. It appears possible to construct a situational judgment test with acceptable internal consistency with as few as 15 to 30 items, especially if that test addresses a single construct of interest (*e.g.*, dealing with subordinates).

Validity of Situational Judgment Tests

The validity of a test shows the degree to which the test is useful in light of the inferences that the test giver wishes to make. Simply put, a situational judgment test is our best guess about actual situational judgment, or how respondents might respond in a real life situation. While no test is a completely accurate reflection of behavior, tests can be shown to be significant predictors of behavior and/or be related to important factors that affect behavioral outcomes.   Psychologists refer to test validity as being demonstrated both externally and internally. External validity refers to what the test can meaningfully predict in a practical sense. External validity is chiefly demonstrated through criterion-related validity, where a *criterion* is some real world measure of performance that can be shown to meaningfully relate to how a respondent performs on a test. Internal validity refers to the nature of the test itself: which construct(s) the test measures, what the content domain of the test is, and how respondents perceive the test.

Criterion-Related Validity

Most research investigating criterion-related validity has used highly subjective criteria, such as multi-faceted performance ratings from supervisors. A meta-analysis by

10

McDaniel, Morgeson, Finnegan, Campion, and Braverman (2001) showed that SJTs have a mean uncorrected validity of .26 with performance. Numerous studies have shown a moderate correlation between SJT responses and supervisor performance ratings (Chan & Schmidt, 2002; Clevenger *et al.*, 2001; Weekley & Jones, 1997; Motowidlo & Tippins, 1993; Motowidlo, Dunnette & Carter, 1990) and turnover (D'Alessio, 1994). Table 2 shows several studies that reported correlations between SJT scores and subjectively rated job performance.

Table 2: Relationships Between SJT Scores and Supervisor Ratings of Performance

Study	N	Coefficient (r)
Chan & Schmidt (2002)	160	.30*
Clevenger *et al.* (2001)	412	.21* (avg.)
Motowidlo & Tippins (1993)	165	.31*
Motowidlo, Dunnette & Carter (1990)	252	.30**
Weekly & Jones (1997)	1471	.35**

*Significant at $p < .05$, **significant at $p < .01$

Oswald *et al.* (2004) similarly determined that SJT scores correlated with self-ratings of college performance in a sample of college freshmen. Stevens and Campion (1999) developed an SJT that addresses teamwork. They determined that their instrument was moderately correlated with ratings of teamwork performance ($r = .32$) and overall performance ($r = .37$), though their instrument was highly redundant with the measure of general cognitive abilities that they used in the same study ($r = .95$). Weekley and Jones (1999) determined that their situational judgment instrument was correlated only weakly ($r = .19$, *n.s.*) with performance ratings, despite the fact that their previous research showed more predictive validity.

11

Construct and Content Validity

Construct validity refers to an instrument's ability to measure an unobservable idea or construct. Constructs must be defined through the characteristics of the variables that are used to measure them. The most common way to demonstrate the construct validity of an instrument is to compare it to other, well-established measures of the construct. Content validity refers to how completely an instrument addresses the entire realm of characteristics that make up a construct. The "realm of characteristics" is more appropriately called the content domain. A substantial portion of SJT research has suggested that SJT responses are little more than measures of general cognitive ability (g).

Cognitive Ability

Weekly and Jones (1999) showed that SJTs are significantly related to cognitive ability with an average weighted correlation of .45, however, their sample of Yale undergraduates and managerial employees certainly had a severe restriction of range on the cognitive ability measure. McDaniel et al. (2001) provided meta-analytic evidence that SJT responses have an average corrected correlation of .39 with general cognitive ability. This finding shows the clear import of cognitive ability in understanding SJT responses; it also shows that there is more to SJT responses than just (g). For example, McElreath and Vasilopoulos (2002) reported that "most likely" and "least likely" responses to SJT items have a different relationship with (g): least likely SJT scores had a stronger relationship with cognitive ability than did most likely SJT scores. This may be because what *should not* be done is typically very clear to respondents with higher

12

cognitive abilities, although what *should* be done is not always as obvious.   Further, Motowidlo, Dunnette and Carter showed that SJT scores did not correlate with aptitude test scores (1990). Numerous studies have shown that SJTs provide incremental validity over and above cognitive ability tests (Chan & Schmidt, 2002; Clevenger *et al.*, 2001; McDaniel *et al.*, 2001). The incremental validity provided by SJTs affords the notion that there is more to the content domain of the SJT than just general cognitive ability.

The content domain of situational judgment response may be incomplete because too much emphasis is placed on (g). While cognitive ability is certainly an important factor in understanding situational judgment response, considering only cognitive ability ignores the possible influence of Agreeableness, multifaceted levels of experience, and other variables.

Interpersonal Traits and Skills

McDaniel and Nguyen (2001) performed a meta-analysis in which they determined that Agreeableness correlated with SJT responses (mean *r* = .25 over 12 studies). SJT scores have also shown to be correlated with "interpersonal skills," "communication skills," & "negotiation skills" as rated by interviewers (Motowidlo, Dunnette & Carter, 1990).

There has also been some exploration of how responses to SJTs may be influenced by personality, not by levels of a given trait specifically, but by how effective different levels of expression of that trait are perceived to be in a given situation. For example, Motowidlo and colleagues (2006) coined the term ITP or implicit trait policy, suggesting that personal levels of Agreeableness could influence responses, but also that

13

the amount of importance given to Agreeableness in decision-making would influence

how favorably participants viewed high Agreeableness versus low Agreeableness

response options in terms of effectiveness. Motowidlo and colleagues (2006) showed

that personality traits do have some influence on situational judgment; notably,

procedural knowledge scores were significantly correlated with agreeableness scores $r =$

.25 ($p < .01$), and procedural knowledge scores were significantly correlated with

implicit trait policy for Agreeableness $r = .73$ ($p < .01$).

Borman and his colleagues (1991) reported that job knowledge mediates the

relationship between cognitive ability and performance. Other researchers have agreed

that job knowledge (as operationalized by job experience) appears to have a positive

relationship with SJTs (McDaniel & Nguyen, 2001). Weekly and Jones (1999) found

that SJTs are correlated with overall job experience ($r = .23$), but not necessarily with

company tenure ($r = .02$).

Face Validity

One of the main reasons that Situational judgment tests are highly valued is

because they typically show a high degree of face validity. Face validity may be the most

important kind of validity in terms of creating an instrument that encourages valid

responses from the test taker. Face validity is a theoretical term that refers to how the

instrument appears, especially to the respondent. Situational judgment tests offer a high

degree of face validity because respondents perceive the items on the tests to be highly

related to the duties they will perform on the job. A test of general cognitive ability may

not be perceived in the same way, because the items on a cognitive ability test, though

14

predictive of performance outcomes, may be perceived as being outside the context of realistic job performance (Callinan & Robertson, 2000). Likewise, the perception that a test is an opportunity to show how well they can perform on highly job-related tasks leads to more applicant motivation and better engagement in the test (Callinan & Robertson, 2000).

Generalizability of Research on SJTs

Situational judgment tests appear most valid for the jobs and organizations for which they were originally developed. Even though similar situations may occur in different organizations, differences in organizational goals, culture or values may require unique scoring keys (Hanson & Ramos, 1996). Research on the generalizability of SJTs across organizations would be a great contribution to the literature (Hanson & Ramos, 1996). It is widely accepted that a high level of cognitive ability is desirable for success in every organization. However, it is the basis of this research that SJTs cannot fairly be generalized across organizations because different organizations require different degrees of characteristics such as Agreeableness and leadership style.

The body of empirical research offers numerous recommendations for future research using situational judgment tests. Weekly and Jones (1999) and McDaniel and colleagues (2001) suggested that a nomological net of the constructs that SJTs measure should be developed, and recommended measuring specific constructs with SJTs by developing items directly related to those constructs. These authors also suggested that SJTs are most likely to have different nomological nets (thus, address different constructs) if they are based on unique aspects of job content in different jobs; for

15

example, dealing with subordinates would probably have a different nomological net than closing the sale with new clients, even though both types of items may appear on the same situational judgment test. McDaniel and Nguyen (2001), Clevenger and his colleagues (2001), Chan and Schmidt (2002), and Motowidlo and his colleagues (2006) echoed the recommendation that future research should develop strategies to target specific constructs within the context of situational judgment items. Ployhart and Ehrhardt (2003) suggested that more research is needed on the psychological processes that people employ to complete SJTs. It has also been suggested that experience should be measured more specifically to better understand the role it plays in SJT response (Chan & Schmidt, 2002; Clevenger *et al.*, 2001). Finally, Oswald and colleagues (2004) called for a better understanding of how impression management affects responses to situational judgment items. The current research moves in these directions first by narrowing the focus of an SJT to issues of supervisors dealing with subordinates and second by elaborating the nomological net. Specific variables to be included in the nomological net are described next.

<center>Agreeableness</center>

What is Agreeableness?

As early as the 1960's, personality researchers had developed a rudimentary "Five Factor" model of human personality (Norman, 1963; cited in McCrae & Costa, 2003). More than 20 years later, McCrae and Costa (1985) first proposed the "Big Five" model that is widely accepted today. The five factors that Costa and McCrae hypothesized are:

<center>16</center>

Extroversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience.

Agreeableness is described as selflessness, concern for others, trust and generosity of sentiment (McCrae & Costa, 2003). The antithesis of Agreeableness is referred to as Antagonism or tough mindedness (McCrae & Costa, 2003). It should be noted that a high degree of Agreeableness is not always desirable; in some cases, less Agreeableness, or even Antagonism is advantageous (*e.g.*, for prosecuting attorneys, soldiers in combat conditions, or simply committee members who do not wish to take on additional responsibilities).

McCrae, Costa, and Busch (1986) described the Agreeable and Antagonistic personality through comparison with a set of personality aspects called the California Q-Set. They described a highly Agreeable individual as: sympathetic, considerate, warm, compassionate, arousing, liking and behaving in a giving way. They described a highly Antagonistic individual as: critical, skeptical, showing condescension, pushing limits and expressing direct hostility. Both of these types may be advantageous in certain settings. These type descriptions were used to aid in creating instructional manipulations for this research.

Costa, McCrae and Dye (1991) specified six facets that fall under the factor of Agreeableness. These six facets are: Trust, Straightforwardness, Altruism, Compliance, Modesty, and Tender-mindedness. Each of these has implications for use in the workplace, and were addressed specifically in measures used for this research.

17

Agreeableness as Antecedent to Performance

For almost as long as the Five Factor Model of personality has been established, researchers have explored personality variables as predictors of performance. Tett and colleagues reported in their meta-analysis of confirmatory personality and performance studies that the correlation between Agreeableness and performance was the highest of all the Big 5 personality variables (average $r = .326$; Tett, Jackson & Rothstein, 1991). This is in contrast to a similar meta-analysis performed the same year that showed little relationship between Agreeableness and performance, even in jobs that seem to require a high degree of sociability (Barrick & Mount, 1991)[1]. Further research later conceded that Agreeableness (among other personality variables) is a better predictor of job performance in highly autonomous jobs, such as management, than in less autonomous jobs, though the magnitude of the reported correlations remained small (Barrick, Mount & Judge, 2001; Barrick & Mount, 1993).

Agreeableness and SJTs

Although evidence of a significant relationship between Agreeableness and performance is slim, some research has supported the potential relationship between Agreeableness and responses to situational judgment items.

McDaniel and Nguyen (2001) showed that SJTs are correlated on average with measures of Agreeableness ($r = .25$), Neuroticism ($r = .31$), and Conscientiousness ($r = .27$). Cucina, Vasilopoulos and Leaman (2003) suggested that "Best" SJT responses

---

[1] Tett *et al.*, (1991) argued that this was because Barrick and Mount neglected to consider absolute values in their analyses, effectively causing significant negative and positive values to "cancel each other out." Barrick, Mount and others continue to assert that Tett *et al.*'s analytic methods were mathematically incorrect.

reflect an individual's typical behavioral preferences (which should be a function of personality) and are most likely to be relevant in highly autonomous situations. Similarly, "Worst" responses would be most relevant to situations in which individuals' personalities have less influence on their behavior and they are not able to support the behaviors they would consider "Best". Cucina, Vasilopoulos and Leaman (2003) also found higher correlations between personality and situational judgment when using measures of narrower constructs.

Clearly those personality factors that play a role in situational judgment response require further empirical clarification. Another construct relevant to management that calls for additional research is leadership.

Leadership

Why Leadership?

Situational Judgment tests are tools that are most often used to select managers and supervisors (Hanson & Ramos, 1996). It is therefore important to consider how characteristics of leadership may affect SJT responses. Leadership is defined as the process of influencing other group members to achieve organizational goals (Greenberg & Baron, 1997). Organizations may or may not make a distinction between leaders and managers: leaders are responsible for the vision of the company while managers are responsible for the implementation of that vision (Greenberg & Baron, 1997). This distinction seems somewhat arbitrary: although some management employees do not play a role in setting organizational goals, a manager is often perceived as the leader of his or her work group and must contribute as a leader on a smaller scale. This is especially true

19

when the organization is so large that employees have little to no contact with the leader of the company, when the organization is so small that leaders must also take on the duties of management, or in organizations that encourage participation in goal-setting and decision-making from all levels of the company.

Numerous theories of leadership have been set forth over the last 100 years of research (Yukl, 1994). Leadership theory began with the idea that great leaders are born, not made, called the "great man theory" – that particular traits set leaders apart from ordinary people (Locke, 1991). Although most researchers have dismissed the "great man theory," support for a solely trait (or dispositional) theory of leadership continues in recent literature (e.g., House, Shane & Herold, 1996). Current trends suggest that leader traits are important, but that successful leadership is also influenced by situational variables. Called contingency theories, these models suggest that leaders must adapt their behavior based on the organizational environment (Schriescheim, Tepper & Terault, 1994). Modern contingency leadership theories also typically address the relationship between leaders and their subordinates (Yukl, 1994). One of the more prominent modern contingency theories of leadership is Path-Goal theory.

Path-Goal Theory of Leadership

The Path-Goal theory of leadership addresses how formal supervisors affect the motivation and satisfaction of their subordinates (House, 1996). It is a dyadic theory, meaning that it concerns itself with individual relationships between a supervisor and each subordinate (House, 1996). When House first proposed Path-Goal theory (1971), he was concerned only with establishing that supervisors served the needs of subordinates in

20

two different ways: supervisors 1) helped subordinates to follow a course toward an outcome (the "path-goal" component), and 2) satisfied subordinate needs. He adopted the constructs of Consideration and Initiating Structure that were originally advanced by the Ohio State Leadership Studies (Stogdill & Coons, 1957). Consideration is defined as the amount of concern and empathy a leader shows toward subordinates (Judge, Piccolo & Ilies, 2004). Initiating Structure refers to how much a leader defines roles for self and subordinates, sets goals, and is achievement-oriented (House, 1996). Not surprisingly, Consideration has been consistently shown to correlate more strongly with subordinate satisfaction, while Initiating Structure has been shown to correlate more strongly with performance and effectiveness (Judge, Piccolo & Ilies, 2004). One major drawback of this two-factor leadership theory is that Consideration and Initiating Structure were initially supposed to be orthogonal, but numerous studies have shown that there is a significant correlation between them (Judge, Piccolo & Ilies, 2004; Fleishman, 1995; Bass, 1990).

Later, House and Mitchell (1974) modified Path-Goal theory into the traditional four-factor model that is known today. Directive leadership is described as providing structure and expectations to subordinates; essentially, telling subordinates what they are supposed to do. Participative leadership is described as consulting with subordinates and taking their opinions and suggestions into account when making decisions; in effect, this approach allows subordinates to help decide what is to be done. Both Directive and Participative leadership are components of path-goal clarifying behavior (Evans, 1996). Supportive leadership is described as creating a supportive work environment and being

21

concerned for the welfare of subordinates. Achievement-oriented leadership is described as setting challenging goals and emphasizing performance excellence. Both Supportive and Achievement Oriented leadership are components of satisfying subordinate needs (Evans, 1996). It is important to stress that although these characteristics would be demonstrated with marked behavioral contrasts, they are not polar opposites and are not mutually exclusive – it is possible for a leader to be both Supportive and Achievement-Oriented, for example (Greenberg & Baron, 1997). Further, Path-Goal theory also states that it is up to the leader to motivate subordinates through appropriate use of the components described above and that the effectiveness of this motivation is contingent on the degree of structure present in the work being performed (House, 1996).

House later reformulated his own theory again, using Path-Goal theory as a springboard for Charismatic Leadership theory and most recently, the Path-Goal theory of Work Unit Leadership (House, 1996).

Path-Goal Theory in Research

Path-Goal theory has enjoyed moderate support in empirical research. A meta-analysis performed by Judge, Piccolo and Ilies (2004) cited numerous studies that found little to no relationship between contingency theories of leadership and outcome variables. However, the same study analyzed several hundred correlations between consideration/initiating structure factors (the bases for Path-Goal theory) and leadership outcomes such as subordinate satisfaction and leader job performance, and reported

22

moderate average correlations for both of the above factors with numerous leadership outcomes (Consideration .48, Initiating Structure .29; Judge, Piccolo & Ilies, 2004).

A handful of recent studies have shown that specific elements of the Path-Goal model (particularly participative leadership) influence subordinate outcomes. Coleman (2004) determined that managers with more cooperative beliefs and ideals about organizational power relations were more likely to engage in participative leadership behavior than those with more competitive beliefs and ideals about organizational power relations. This effect was enhanced by the use of subliminal priming (quickly showing words related to competitive or cooperative beliefs on a computer screen) so that competitive priming reduced the participative leadership behaviors of even those with more cooperative beliefs about organizational power.

Oshagbemi (2004) demonstrated that older managers tend to use significantly more participative leadership behaviors than younger mangers. However, differences in the amount of directive leadership behaviors between older and younger managers were not statistically significant.

Somech (2003) studied demographic differences between leaders and subordinates and showed that differences in age, gender, and level of education between the leaders and subordinates decreased the amount of participative leadership that was exhibited. These effects diminished over time, with the exception of dissimilar genders, which intensified over time. In other words, the longer a demographically dissimilar leader and subordinate worked together, the more likely the leader would display participative leadership, except in the case that the leader and subordinate were of

23

opposite genders; in this case, the leader was less likely to display participative leadership after a longer relationship.

Kahai, Sosik and Avolio (1997) showed that participative and directive leadership provided diverse results from subordinates in an electronic meeting situation. Under participative leadership conditions, subordinates provided more suggested solutions, were more supportive, and were less critical of the situation, than under directive conditions.

There is no research that considers the effects of Path-Goal characteristics on situational judgment responses. However, Somech (2003) suggested that the relationship between participative leadership and organizational culture and structure should be explored in future research. Path-Goal theory is a useful way to think about leadership from the perspective of organizational culture. Path-Goal theory was not used in this research to predict subordinate performance at different levels of task structure. Instead, the four characteristics identified by Path-Goal theory were used to create fictitious leadership cultures to demonstrate how organizational leadership culture can affect situational judgment responses. Specifically, it is likely that respondents will attempt to display congruence between their reported behavior (measured by responses to a situational judgment test) and the fictitious organizational leadership cultures that were presented in this research. The desire to display such congruence is due to social desirability, or more specifically, impression management, or the ability to "fake good."

24

Impression Management and Faking

What are Impression Management and Faking?

Impression Management is the attempt to present a positive impression of oneself to someone else (Ones & Vishwesvaran, 1998). A "positive impression," in the context of selection, refers to the representation that best fits into organizational norms; "someone else" refers to the person making the selection decision. There are numerous studies that show how impression management is employed during selection interviews (*e.g.*, Kristof-Brown *et al.*, 2002; Silvester *et al.*, 2002; Vishwesvaran *et al.*, 2001). It has similarly been written that responses to personality assessments are self-presentations, not self-reports (Hogan, 1998). That is, the response that a person provides to an instrument, especially when the person is highly motivated to succeed on the instrument (such as a candidate for employment), is likely to reflect how the person would like to be seen, rather than how they truly are. It is likely the inference can be extrapolated to situational judgment tests: respondents are likely to choose the behavior that they perceive is most acceptable to the organization, rather than the behavior they would truly exhibit. This type of intentional distortion is better known as faking.

There is good evidence from past research that participants are motivated to "fake good" even when they are asked to do so only for research purposes (*e.g.*, McFarland & Ryan, 2000). Likewise, individuals can fake selectively, that is, on only certain parts of selection instruments (Dalen, Stanton & Roberts, 2001). The same research showed that the amount of information presented to an individual has little influence on how much they choose to fake (Dalen, Stanton & Roberts, 2001).

25

Situational judgment tests are about judgment; thus they appear to tap an ability that is expected to be relatively enduring. As previously mentioned, such tests typically correlate with measures of cognitive ability. The nature of the questions also appears to tap maximal rather than typical performance, especially when the stem is written to address Best and Worst responses to the situation (Hooper, Cullen & Sackett, 2006). However, many SJTs (including the one used in the current research) concern issues of social conflict. Therefore, there is good reason to believe that individual differences outside of analytical judgment, such as beliefs regarding norms, customs, personal values and experience, personality, and ideas about the testing organization's values may all influence the choice of the best and worst response to SJT items. In fact, this research expected that all motivated respondents would respond to a situational judgment test with some degree of impression management or faking, limited in only two circumstances: when extreme differences between respondents' individual values and the organizational values presented create cognitive dissonance, *or* when not enough information about organizational values is presented for respondents to form any impression. These circumstances were controlled in the study through the use of manipulation checks.

Experience

What is Experience?

When industrial and organizational psychologists refer to "experience," they usually mean a simple measure of time on the job, often the sum of all time spent on similar jobs with different organizations. A seminal meta-analysis performed by Quinones, Ford and Teachout (1995), showed that out of 22 studies of experience

26

completed in the previous 20 years, all but 2 used time on the job or time with the company as their experience measures. Further, the analyses for 15 out of the 22 studies were computed at the level of job tenure (7 were computed at the level of organizational tenure).

The variable "experience" is a common way to operationalize job knowledge. More experienced workers are typically expected to perform at higher levels, make fewer mistakes, and require less supervision (Greguras, 2005). Popular belief suggests that job knowledge is directly related to time on the job, although it appears that experience may be asymptotically related to job knowledge (McDaniel, Schmidt, & Hunter, 1988), or even parabolically related to job knowledge (Sturman, 2003). That is, job knowledge does not continue to increase with time on the job, but either reaches asymptote after a certain amount of time (when little or no job knowledge and skills remain to be learned), or begins to decrease after a certain amount of time (when job knowledge and skills are no longer state of the art).

The word "experience" usually refers to job experience, and thus job knowledge; however, there are different kinds of work-related experience that an individual can have. For example, time with one particular organization, even across numerous and unrelated jobs, could be an operationalization of *organizational* experience (or "organizational socialization," see Quinones, Ford and Teachout, 1995). Organizational experience may be related to what Sternberg (1995) referred to as tacit knowledge: the individual understands organizational rules, and formal and informal procedures, even if the "rules" are never explicitly stated (Sturman, 2003). Management in companies that promote

27

from within can be assumed to have a moderate to high level of organizational

experience, since workers would be promoted to management after some years of service.

On the other hand, managers who are recruited externally will probably have a low level

of organizational experience when they are new to their management positions. These

externally recruited managers may have more trouble adjusting and performing than

those with more organizational experience.

Similarly, time in a particular type of organizational culture, even across

numerous and unrelated organizations, might be considered *culture* experience. That is,

the individual understands what broad goals, values, and ideals are emphasized, on both

formal and informal levels. It follows, as above, that when new management is sought,

managers with highly incongruous cultural experience in their previous organizations will

have more trouble performing and adjusting than those managers with more congruous

cultural experience. Although there is very little research on the aforementioned types of

experience, it was hypothesized that at least job experience may have important effects

on SJT responses.

Experience as Antecedent to Performance

Job experience (and thus job knowledge) is one of the most widely recognized

predictors of job performance (Kolz, McFarland, & Silverman, 1998). McDaniel,

Schmidt and Hunter (1988) found a mean correlation of .32 between job experience and

performance in their meta-analysis across multiple occupations.

28

Experience and SJTs

SJTs are highly related to job knowledge (and thus experience), which presents a problem in terms of testing inexperienced respondents who may have little or no experience in the situations as presented (Weekley & Jones, 1999). This is likely more of a concern on a SJT of technical, or "hard" skills, which are typically very job-specific, than with SJTs of "soft" skills (such as dealing with subordinates) that may be commonly used across numerous jobs.

Quantifying Experience

It is difficult to assign value to experience in terms of months or years in a position. Several theories of expertise do exist. Most commonly, experience is broken down into three to five linear stages proceeding from novice to expert (*e.g.*, Anderson, as cited in Genberg, 1992; Dreyfus & Dreyfus, 1985). However, developing expertise may not be a linear process. Although job experience and organizational experience are highly correlated with time on the job, little research exists on how much *time* at work separates "novices" from "experts" (Genberg, 1992). Daley (1999) states, "competent professionals have usually been in practice three to five years" (pp. 134-135). It must be noted that this research is focused on expertise for nurses, and what Daley calls a "competent professional" would fall about in the middle of the linear novice-to-expert model proposed by Dreyfus and Dreyfus (1985). It is difficult to find recommendations on quantifying expertise for managers, in part because the duties of a manager can be quite different across many different organizations. For the purposes of this research,

29

both job and organizational experience were categorized into five distinct groups based roughly on Dreyfus and Dreyfus' (1985) model.

<div align="center">Research Hypotheses</div>

The nomological network that follows in Figure 1 is presented for the purposes of hypothesis testing only. The model is intended to illustrate overall hypothetical relationships between the constructs described above and to show that they may affect responses to situational judgment tests. It is not intended as a structural equation model.

Figure 1: A Nomological Network of Constructs Affecting SJT Response
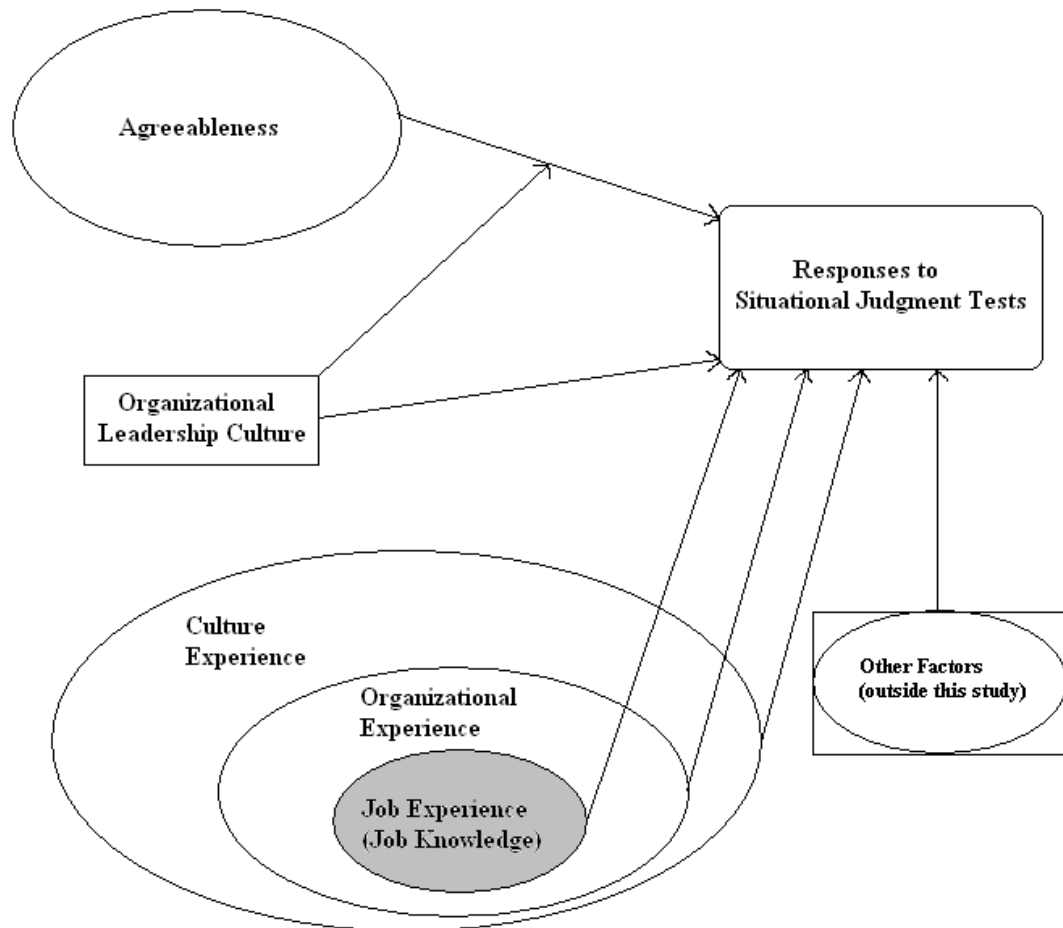
This model suggests that Personality and Experiential constructs, in addition to attempts to impression manage, affect responses to situational judgment tests. The model also demonstrates the contribution of organizational leadership culture, and how it can indirectly influence the effects of Agreeableness. The Experiential portion of the model shows how job experience may be thought of as nested within organizational experience, which is similarly nested within culture experience. Job experience is shaded to indicate the dominance of this construct in the literature on experience. However, the model in Figure 1 argues that both organizational and culture experience may also contribute individually to SJT responses. Finally, it is realistic to assume that there are other, unexplored constructs that play a role (and that may affect the Personal and Experiential constructs), either on an individual or an organizational level. For example, as stated above, there is ample evidence that cognitive ability can influence responses to SJT items. Such constructs are beyond the scope of this research; their presence is included only to demonstrate that the model is not a complete explication of what factors affect situational judgment responses.

The following hypotheses are the central questions of this research. More complex hypotheses are illustrated graphically to clarify the expected effect.

The Effects of Agreeableness on SJT Scores

Hypothesis 1: SJT scores will be affected by the personality variable of Agreeableness regardless of leadership culture manipulation. That is, respondents with a relatively higher level of Agreeableness will choose different responses to situational judgment

31

items than those with a relatively lower level of Agreeableness, regardless of which organizational leadership culture is presented.

The Effects of Organizational Leadership Culture Manipulations on SJT Responses

Hypothesis 2: Responses to situational judgment items are influenced by the target's perception of the leadership culture of the organization.

Corollary 2a: A motivated test taker who has knowledge of the testing organization's leadership culture will choose those responses that best reflect the test taker's understanding of the leadership culture so that scores will differ between strong Participative/Supportive and strong Directive/Achieving cultures.

Corollary 2b: A motivated test taker who is presented with no information about the leadership culture of the testing organization will choose responses without being influenced by leadership culture so that scores will differ between a Neutral condition and the strong conditions described above.

Corollary 2c: Item-level responses will differ across disparate leadership cultures.

The Interaction of Agreeableness and Organizational Leadership Culture Manipulations on SJT Scores

Hypothesis 3: The relationship between Agreeableness and SJT scores will be different in discrete leadership culture conditions such that High Agreeableness shows a positive relationship with SJT responses in some conditions (Participative/Supportive and Control), and a negative relationship with SJT responses in other conditions (Directive/Achieving), as indicated in Figure 2.

32

Figure 2: Agreeableness X Leadership Characteristics Interaction Effects on SJT Scores
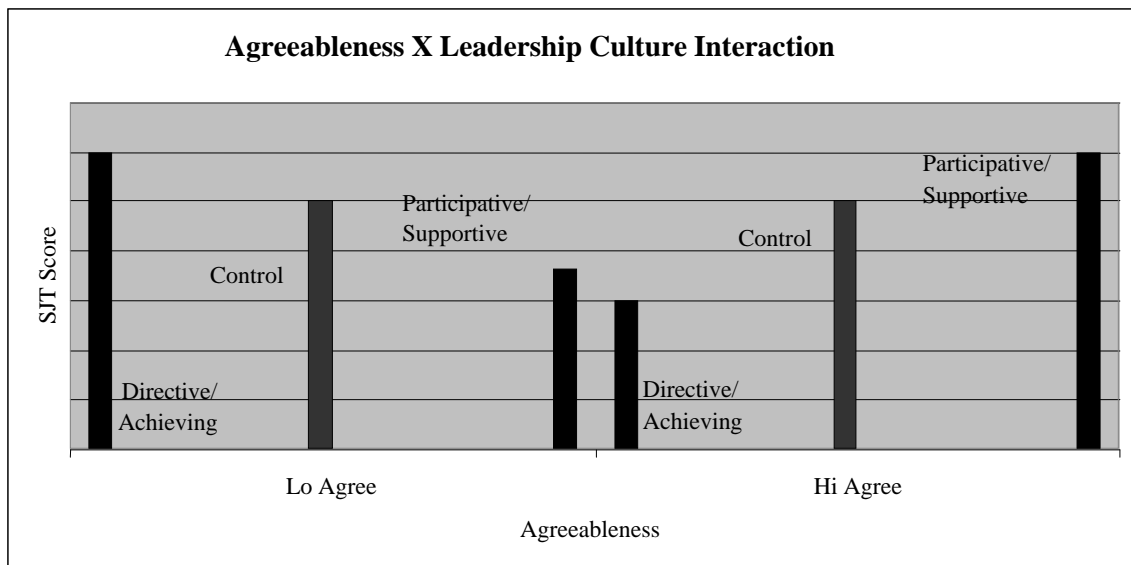
The Effects of Experience on SJT Scores

Hypothesis 4: SJT scores are influenced by Job Experience such that scores on an SJT

will be higher for experienced vs. inexperienced participants, regardless of organizational

manipulation.

Essentially, experienced managers likely have a deeper understanding of how to

respond to the different demands of discrete organizations and organizational cultures,

and will make judgments appropriately.

The Interaction of Job Experience and Organizational Leadership Culture Manipulations

on SJT Scores

Hypothesis 5: SJT scores under different organizational manipulations will be moderated

by job experience. That is, when a relatively low degree of job experience is present,

information on organizational leadership culture may have a smaller effect on judgment.

This hypothesis may hold especially true in the Directive/Achieving condition.

33

Inexperienced participants will likely not attach as much significance to the cultural

information presented in the manipulation as experienced participants, and will thus score

lower in the Directive/Achieving condition. The same effect may also occur in the

Participative/Supportive condition, but it is not expected to be as pronounced (please

refer to Figure 3).

Figure 3: Experience X Leadership Characteristics Interaction on SJT Response
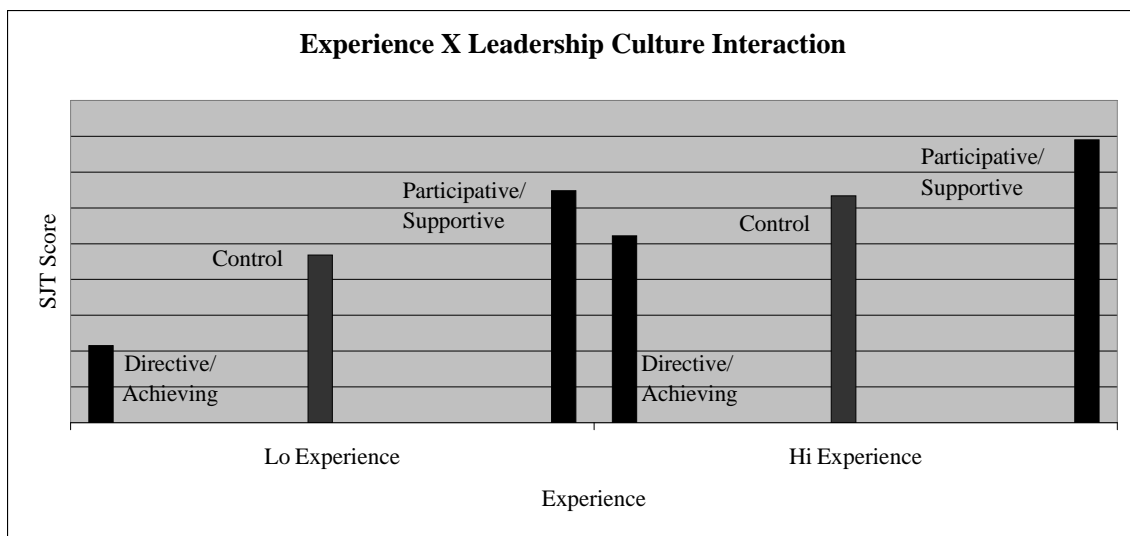
Method

Participants

Two samples were recruited for the experimental study. The first sample

consisted of retail sales managers and assistant retail sales managers recruited from a

major national telecommunications company. Retail sales managers are line employees

who are responsible (within this company) for day-to-day operations of retail stores,

where cellular handsets and peripherals are sold, customer accounts are processed, and

customer service issues are handled on a person-to-person basis. Assistant retail sales

34

managers have very similar duties and differ from "full" managers only in their level of experience or tenure with the company. Retail stores usually have a single retail sales manager or a single assistant retail sales manager; larger stores may have a retail sales manager and an assistant retail sales manager. Retail sales managers are typically responsible for an individual retail location of 8-10 employees (Jonathan Canger, Associate Director of Staffing and Talent Acquisition, October 4, 2006, *personal communication*). District Managers throughout the company (226) were asked to provide the names of five retail sales managers or assistant retail sales managers within their districts to participate in this research. Approximately 149 District Managers responded (response rate = 66%), resulting in a list of 745 potential participants. Study information was emailed to the participants with a link to participate in the research online. A total of 386 managers and assistant managers responded to the request to participate, and 258 completed the entire research instrument (overall response rate = 35%). The sample consisted predominantly of retail managers (N of Managers = 229, or 89%; N of Assistant Managers = 23, or 9%, 6 employees, or 2% did not respond to this item). Demographic data on experience for this sample is presented in Table 3.

The second sample in the main study consisted of 138 undergraduates from a large southeastern university. Students from introductory psychology courses volunteered to participate in the research in partial fulfillment of course experimental participation requirements. Although data on major was not collected, because introductory psychology courses are core requirements, the students within likely represented a diverse mixture of backgrounds and potential majors (*i.e.,* not all the

35

students were psychology majors). Student data were controlled for any part- or full-

time management experience. Demographic data on experience for this sample are

presented in Table 3.

Table 3: Comparative Experience for Managers vs. Students

Management Experience	Managers		Students	
	Frequency	Percent	Frequency	Percent
0 years (None)	0	0%	70	50.7%
Less than 1 year total	10	3.9%	22	15.9%
1 to 3 years total	43	16.7%	26	18.8%
3 to 5 years total	38	14.8%	10	7.2%
More than 5 years total	166	64.6%	10	7.2%
Total N	257*		138	
Standard Deviation	0.91		1.28	

*1 professional participant did not respond to the experience items

Procedure

Each participant responded to a series of situational judgment items as described

in the introduction to this research (see Appendix B). Three separate sets of instructions

were developed for both the student sample and the management sample. Each set of

instructions consisted of a description of a hypothetical organization and a fabricated

email about the organization from a fictional acquaintance within that organization. All

participants responded to the same situational judgment items after being exposed to one

of the three conditions:

1) The information provided in the "Participative/Supportive" condition suggested

that the primary goal of management in this particular organization is to be highly

36

supportive of subordinates even at the expense of profitability (see Appendix A). For example, subordinates are encouraged to contribute to and correct their supervisors, while supervisors are expected to provide a high amount of job-related and personal support to their subordinates.

2) The information provided in the "Directive/Achieving" condition suggested that the primary goal of management in this particular organization is to focus on profitability even at the expense of subordinate support (see Appendix A). That is, subordinates are discouraged from contributing to or correcting their supervisors, while supervisors are instructed place organizational goals ahead of the personal and professional needs of subordinates.

3) No information was provided in the "Control" condition so that respondents were able to draw their own conclusions without any kind of direct influence (see Appendix A).

Participants were assigned to one of the three manipulation conditions based on their date of birth. The first item on the survey asked participants to report their date of birth (day of the month only), and program logic took participants immediately to the correct manipulation condition, then on to the main survey.

The manipulations were evaluated in a brief pilot study to ensure that they were being interpreted by participants in a manner consistent with Path-Goal Theory and the intent of the research. Please refer to the section of this dissertation entitled "Preliminary Analyses: Pilot Testing the Manipulations of Organizational Leadership Culture" in the "Results" section for supporting information.

37

Participants were asked to imagine they are applying for a job that they are very motivated to get, the idea being that they would therefore try their best to model response sets that would be acceptable to the company to which they are applying. Upon completing the situational judgment measure (Appendix B), participants were asked to complete a secondary inventory (see Appendix C) consisting of personality items, experience items, and manipulation checks (to determine whether they were dissimulating as instructed during the situational judgment portion of the research). Participants were instructed that they should now answer honestly and no longer attempt to fit their responses to the information they read at the beginning of the research.

All participants responded to the research measures electronically. Participants were emailed a web address (URL) to access an online "survey-hosting" website that displayed one of the instructional manipulations followed by the research instruments described. Data was downloaded from the website's database after collection, and coded and analyzed using Microsoft Excel and SPSS 11.5.

Measures

Situational Judgment Inventory: The *SJT-280*

Although managers certainly have additional job responsibilities, it is argued that one of the most important and characteristic duties of a manager is to supervise subordinates. Therefore, the primary measure was a set of situational judgment items that all address dealing with subordinates (see Appendix B). These items were derived from several sources. A total of 26 situational judgment items were included. This author created nine items specifically for use in this research. Eight items were taken

38

from the ProveIt® Manager by Kenexa, Inc. Two were taken from the Supervisory Skills Inventory (SSI™) by gNeil, Inc. Seven were based on items from an inventory created by the Personnel Decisions Research Institute for non-commissioned officers in the U.S. Army [because this measure was specifically created for Army officers, wording was edited to make the items more appropriate for a corporate setting. The nature of the items was not changed] (Hanson & Borman, 1992). All items are used in this research with the permission of the test development companies (Kenexa, Inc., gNeil, Inc., and the Personnel Decisions Research Institute). All three instruments are highly researched selection instruments that include situational judgment items on numerous management-related topics; with the exception of the PDRI measure, these tests are commercially available for private use.

Effectiveness of a particular response was expected to be subjective and highly culture-and experience- dependent. Therefore, six separate scoring keys were initially created for each of the three conditions described above, with both an experienced and a novice sample. This was necessary because what constitutes effective performance is thought to be different in organizations with disparate leadership cultures like those described above. Further, level of experience was expected to contribute to interpreting the effectiveness of responses.

Scoring keys were developed through a pilot study of novice and experienced raters. Undergraduate students at a major southeastern university created the novice scoring keys for the SJT Inventory by rating the Effectiveness of *every* response choice on a 1-4 scale (Note that this rating procedure was more exhaustive than the Most/Least

39

Effective rating procedure used in the main study). Upper-level management at a major southeastern freight and shipping company created the experienced scoring keys in similar fashion.

The experienced (Manager) and novice (Student) ratings were converted to analogous scores and analyzed to demonstrate the item difficulty and reliability of the instrument. Although alphas for the Manager items were acceptable (range of = .80 - .85), the alphas for two of the three Student keys were slightly lower than the standard of .70 (Participative = .78; Directive = .66; Control = .64). Further, all keys showed numerous negative item-total correlations that were difficult to interpret. The finalized scoring keys were created to evaluate the responses of participants in each condition; however, due to potential reliability issues and other concerns described later, an additional keying method was developed. For supporting information on the development and abandonment of the initial scoring keys, please refer to the portion of this dissertation entitled "Preliminary Analyses: Pilot Data" in the "Results" section.

Participants in the main research sample were asked to choose one Most Effective and one Least Effective response choice for each item (see Appendix B). This response format was chosen because it is more likely to address maximal performance (also referred to as "should do" or "could do" performance), which is what would be expected from a job applicant (Ployhart & Ehrhart, 2003). Furthermore, it was expected that responding with Most Effective and Least Effective would avoid any cognitive dissonance that might result from asking participants to choose the responses that are "Most Like" and "Least Like" them, due to the likelihood that participants are

40

dissimulating. That is, since participants in some conditions were expected to impression manage their responses, it was expected that separating the participant from the response might lead to answers that were better tailored to fit the manipulation. The revised key method assigned scores to participants based on frequency of response endorsement. Each participant was assigned two scores for each item: one score for the response endorsed as Most Effective and one score for the response endorsed as Least Effective. The scores were equal to the proportion of all experimental participants (from both Student and Manager samples) who endorsed those responses. The frequency-based key method is described further in the section entitled, "Preliminary Analysis: Creating a New Frequency-Based Key Using Experimental Data" in the Results. The frequency-based method allowed for all three conditions and both levels of experience to be scored on the same key while still retaining individual differences in response patterns. This in turn allowed for more sophisticated statistical analyses such as analyses of variance and covariance to be performed on the whole dataset.

Secondary Inventory: The *SPE-30*

A secondary inventory was used to measure the Big 5 personality factor of Agreeableness, which was expected to moderate responses to the Situational Judgment Items. Ten items that measure agreeableness were taken from the IPIP website of public domain test items available for research purposes (*http://www.ipip.ori.org*). These items relate to Agreeableness, and address Costa, McCrae and Dye's (1991) facets of Altruism and Tender- mindedness. However, four additional items were written by this author to address the other Agreeableness facets (Trust, Straightforwardness, Compliance, and

41

Modesty) that appear be related to the experimental construct of dealing with subordinates, but that did not appear to be addressed by the IPIP items.

The secondary inventory included demographic items about management interest and experience, as well as a question about category of industry for the student sample. The industry categories for this question were based on the U.S. Department of Labor Bureau of Labor Statistics Standard Occupational Codes (SOC), taken from the BLS website (*http://www.bls.gov*). Categories that were not likely to include management positions, and the category "Management Occupations" were excluded. These items were used to control for management experience and interest, specifically in the undergraduate sample. These items also included a categorical variable of management experience for the incumbent manager sample. Based on the work of Dreyfus and Dreyfus (1985) described earlier, the experience variable included five distinct categories. Management experience on the part of the students was compared to management experience for the professionals to ensure that there were significant differences between the two experience conditions; this analysis is presented in the "Results" section of this research.

Finally, this inventory included manipulation check items to ensure that participants are responding within the provisions of the instructional manipulations of organizational culture. The items on this inventory were scored on a 5-point Likert scale, with response options ranging from "I Very Much Disagree" to "I Very Much Agree". The secondary inventory is included as Appendix C.

42

Preliminary Analysis: Pilot Testing the Manipulations of

Organizational Leadership Culture

Before the manipulations were ever presented to the pilot or experimental

samples, it was important to ensure that the participants would interpret the

manipulations as intended by the experimenter. A short scale was developed to measure

reactions to the manipulations. The Manipulation Pilot Inventory (see Appendix D)

included 9 items based on characteristics of leadership styles as proposed by Path-Goal

Theory (House & Mitchell, 1974). A total of 14 undergraduate students were exposed to

one of the three manipulation conditions, and then asked to evaluate whether the

company in the description that they read matched important characteristics of

Participative or Directive leadership. It was expected that there would be significant

differences in how participants judged each manipulation condition based on leadership

characteristics. A one-way ANOVA with accompanying post-hoc Tukey tests were

computed, and results for individual leadership characteristics appear in Table 4.

Table 4: Interpretations of Leadership Characteristics Between Manipulation Conditions

	Employees…			Leaders…		
Condition	Are Told What is Expected	Are Part of Decisions	Are Asked for Suggestions	Are Approachable	Are Concerned	Have High Performance Expectations
Participative	**4.20**(a)	**5.00**(a)	**4.80**(a)	**4.60**(a)	**4.60**(a)	**4.60**(a)
Directive	**4.80**(a)	**1.40**(b)	**1.80**(b)	**1.60**(b)	**1.80**(b)	**4.80**(a)
Neutral	**3.25**(b)	**4.00**(c)	**4.00**(a)	**4.00**(a)	**3.75**(a)	**3.00**(b)

Conditions that were significantly different ($p < .05$) are designated by different letters (a, b, c); mean response to each characteristic on a 1-5 scale is presented in **boldface**.

These data suggest that participants are likely to interpret the manipulation

conditions as different on numerous leadership characteristics proposed by Path-Goal

43

Theory. Responses to two leadership characteristics items were not significantly

different, although the effects were in the expected direction: "…employees are told how

to perform…" and "…leaders set challenging goals…" suggesting that these two

characteristics were not interpreted as typical of one particular leadership style, or that

this characteristic was not addressed strongly enough in the manipulation. However, the

significant differences observed are a close fit with the experimental purpose of the

manipulations. Finally, responses to the item on the inventory that addressed whether

participants would like to work in the environment described were not significantly

different across conditions. This is meaningful because it suggests that the environment

described in the Directive condition was not seen as objectionable, potentially decreasing

response bias from participants in that group.

Preliminary Analysis: Pilot Data

Upper-level managers from a major southeastern freight and shipping company

and undergraduate students from a major southeastern university were recruited to pilot

the situational judgment instrument and to develop a scoring key (see Appendix E).

These data were obtained chiefly to determine that the situational judgment items

included in this research had an acceptable level of reliability. A total of 118 participants

(30 across the three professional/manager conditions, 88 across the three novice/student

conditions) were used to create six unique keys. Each participant viewed one of the

instructional manipulations described previously before responding to the situational

judgment items. Participants were informed that their responses would be used to create

a key for a new situational judgment test, and asked to rate **every** response choice on a 1-

44

4 scale (where 1 = A Very Effective Response, 2 = A Somewhat Effective Response, 3 = A Somewhat Ineffective Response and 4 = A Very Ineffective Response).

To create the keys and to measure reliability, scores for each response choice were averaged across all participants in that condition. These means were summed to create an item-level score for each individual based on the response choices they designated as Very Effective minus the response choices they indicated were Very Ineffective (since participants could designate multiple response choices as Very Effective or Very Ineffective, item-level scores were calculated based on the following formula: means $_{\text{very effective}}$ /N means –   means $_{\text{very ineffective}}$ /N means). This method allowed for the creation of analogous scores across all three conditions and both levels of experience. To avoid the likelihood that the range of responses and thus variance would be different across the different levels of the independent variables, every participant was scored on each of the six keys and those six scores were added together to create a composite score from all six keys (referred to in this research as the Summed Six Key Method). However, at this point, a minor setback in data collection occurred.

Preliminary Analysis: Creating a New Frequency-Based Key Using Experimental Data

The professional sample in the pilot study, taken from the freight and shipping company became unavailable at the completion of pilot data collection, and was not available to participate in the experimental research. Therefore, because of the small sample size of the pilot group, because of an unusual number of negative item-total correlations, and to ensure the appropriateness of keyed responses across organizations, a new key was created based on the response frequencies of the experimental data itself.

45

Each participant was given two scores for each item: one score for the response he or she endorsed as Most Effective and one score for the response he or she endorsed as Least Effective. The scores were equal to the proportion of all experimental participants (both Student and Manager samples) who endorsed those responses. For example, if Participant X chose Response "a" as Most Effective for Item 1, and Response "a" was chosen by 38.6% of all respondents, Participant X would receive a score of .386 for Most Effective response for Item 1. Most Effective and Least Effective scores were summed for each item and then item scores were added together to create a composite score for each participant on the entire instrument. The following formula illustrates how composite scores were created using this method: (Most Effective$_{freq}$ + Least Effective$_{freq}$). This method allowed for all three conditions and both levels of experience to be scored on the same key while still retaining individual differences in response patterns. Using the experimental data to create the Frequency-Based Key may be expected to result in multicollinearity and artificial inflation of scores compared to a new sample. However, use of this method of keying the data is justified because of the relatively large N (396), and because any differences observed between experimental conditions cannot be explained by intercorrelation. Although the maximum hypothetical score using this key is 52, the maximum obtainable score using this key was 30.2, the minimum obtainable score using this key was 2.9 (since every response choice that was endorsed at least once must have a non-zero score and no response choice was endorsed at 100%). The alpha for the new, frequency-based key was .76. A single item had a weak negative item-total correlation, which did not strongly affect the reliability.

46

Therefore, all items were retained in the SJT. Finally, although the Six Key Method

was not used in further analyses, it is interesting to note that scores using the Frequency-

Based Key Method and the Summed Six Key Method were highly correlated [$r(395) =$

$.954, p < .05$].

Experimental Results: Manipulation Check

Experimental participants responded to three items used as manipulation checks

in the secondary inventory. These items were included to address whether participants

answered differently than normal because they were asked to act as though they really

wanted the job, and whether the company description and email presented before the SJT

provided clues about how to answer in order to get the job (the verbatim items are

included in Appendix C). A composite score was created based on participants'

responses to these three items. A two-way ANOVA demonstrated that while there was

no significant difference between managers' and students' responses to these items, the

responses from the Control condition were significantly different from the responses

from the Participative/Supportive and Directive/Achieving conditions [$F(2) = 10.99$,

p<.001]. The mean composites for the 3 manipulation check items are shown in Table 5.

Table 5: Manipulation Check Composite Scores Among 3 Conditions

Condition	Mean	Standard Deviation	N
Participative/Supportive	9.06	2.36	114
Directive/Achieving	9.41	2.77	141
Control	7.90	2.31	141

These data suggest that the manipulations of Organizational Leadership Culture

were working as intended; participants in the Control Condition were not as highly

47

influenced by the information presented as the other groups; that is, Control participants were more likely to answer as they normally would (instead of impression managing to get the job), and they were less influenced by the company description and email that was presented.

Experimental Results: Hypothesis Testing

A General Linear Model was created that included main effect terms for Agreeableness, the Leadership Culture manipulations, and Experience, as well as 2-way interactions between Agreeableness and the Leadership Culture manipulations and Experience and the Leadership Culture manipulations. For the convenience of the reader, the research hypotheses are restated below as the data are presented.

Hypothesis 1 stated that scores on the situational judgment test would be affected by the personality variable of Agreeableness regardless of leadership culture manipulation. This hypothesis was tested by the significance of the main effect for Agreeableness. A composite score was created for Agreeableness by summing responses to the 14 Agreeableness items on the secondary inventory (please refer to Appendix C). Reliability analyses of the Agreeableness scale were conducted first to ensure that the Agreeableness items had acceptable internal consistency. Alpha was acceptable at .71 for the Agreeableness measure. Analyses of Variance were conducted to determine whether Agreeableness varied across Experience (Manager vs. Student) or Leadership (Participative/Supportive vs. Directive/Achieving vs. Control) conditions. No significant differences across conditions were found, suggesting that Agreeableness (as tested) varied consistently throughout the experimental sample. Since participants were assigned

48

to condition based on their date of birth, it is unlikely that any systematic variation occurred based on Experience or Leadership Culture Condition.

Analysis of covariance (ANCOVA), using Experience Levels and the Leadership Culture manipulations as fixed factors and the Agreeableness composite score as a covariate indicated a significant main effect for Agreeableness [$F(2, 387) = 4.162$, $p = .042$] in predicting SJT scores. This analysis demonstrates differences in Effectiveness of responses. Means and standard deviations are shown in Table 6.

Table 6: Means, Standard Deviations and Score Ranges for Agreeableness

Mean Score	Standard Dev.	Observed Min – Max	Potential Min - Max
57.1	5.74	34-69	14-70

Discriminant analysis was also performed at a response level to determine if mean Agreeableness composite scores were significantly predictive of the choice of item response for each item. This analysis demonstrates differences in response choices regardless of Effectiveness. A total of 7 out of 52 analyses were significant, or just over 13%. Hypothesis 1 was supported.

Hypothesis 2 stated that scores on the SJT and item level responses on the situational judgment items are influenced by the target's knowledge of the leadership culture of the organization, as demonstrated by the presentation of distinct leadership culture manipulations (versus control). The first part of the hypothesis was tested by the significance of the main effect for the Leadership Culture manipulations [$F(2, 387) = 4.53$, $p = .011$]. Post-hoc Tukey tests were conducted on this result and demonstrated significant differences only between the Directive/Achieving condition and the Control condition. Descriptive statistics are presented in Table 7.

49

Table 7: SJT Score Means and Standard Deviations by Manipulation

Leadership Manipulation	N	Mean	Standard Deviation
Participative/Supportive	114	23.34	2.87
Directive/Achieving	141	22.72	3.37
Control	141	23.91	2.83

Therefore, knowledge of Organizational Leadership Culture, notably in a Directive/Achieving culture may have a significant effect on SJT response. The hypothesis was tested at a response level by comparison of response frequencies through Chi-square analysis. Chi-square coefficients and *Phi* values were computed for each item to test whether the frequency distributions differed across the three culture manipulations for each item. Chi-square coefficients were significant for 22 out of 52 analyses, or 42%, suggesting that item-level responses differed between the culture conditions. Differences in item-level responses among the manipulated conditions may suggest that participants tend to answer differently across conditions, regardless of whether their behaviors would be considered effective (*e.g*., different ineffective behaviors may be endorsed in different conditions). Significant Chi-square and *Phi* values are presented in Appendix F. Hypothesis 2 was supported.

Hypothesis 3 stated the relationship between Agreeableness and SJT scores are different in discrete leadership culture conditions such that High Agreeableness shows a positive correlation with SJT score in some conditions (Participative/Supportive and Control), and negative correlation with SJT responses in other conditions

50

(Directive/Achieving). This hypothesis was tested by examining the significance of the interaction term between Agreeableness and Culture.

The ANCOVA showed no significant interaction between the Agreeableness covariate and Leadership Culture Condition [$F(2, 387) = 1.33$, $p = .266$]. However, correlational analyses between Agreeableness composite scores and SJT scores demonstrated a small but significant positive correlation in the Participative/Supportive Condition, and no correlation in the Directive/Achieving and Control Conditions. Means and standard deviations as well as correlations between Agreeableness and Leadership Culture Condition are shown in Table 8.

Table 8: Means, Standard Deviations and Correlations Between Agreeableness & SJT Score by Condition

Condition	N	Agreeableness Mean Score	Agreeableness Standard Dev.	Correlations
Participative/Supportive	114	57.51	5.03	.191 (*p* = .041)
Directive/Achieving	141	56.27	6.26	.015 (*n.s.*)
Control	141	57.59	5.69	.029 (*n.s.*)

These small but potentially important differences in the relationship between Agreeableness Scores and SJT scores in specific Leadership Culture conditions are illustrated graphically below in Figure 4. However, Hypothesis 3 was not supported.

51

Figure 4: Slopes of Agreeableness * SJT Score Relationships by Leadership Culture
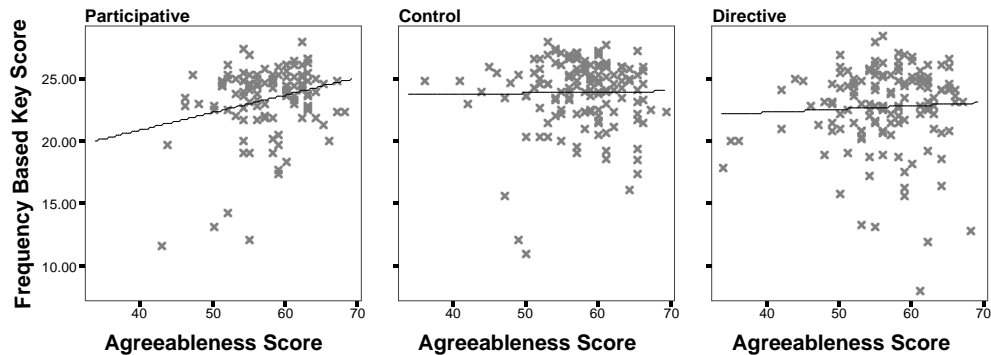
Hypothesis 4 stated that SJT scores are influenced by Job Experience such that scores on an SJT would be higher for experienced vs. inexperienced participants, regardless of organizational manipulation. This hypothesis was tested by the significance of the main effect for Experience.

First, it was important to be sure that the Manager and Student groups clearly differed in level of individual experience, based on the information presented in Table 3 above. An independent samples t-test was used to demonstrate that the experience level of the Management group was statistically different from that of the Student group [$t(393) = 21.1$, $p < .001$]. The analysis of covariance showed a significant main effect for Experience [$F(2, 387) = 102.22$, $p < .001$]. Descriptive statistics for the Experience variable are provided in Table 9. Hypothesis 4 was supported.

Table 9: Mean SJT Scores and Standard Deviations by Experience Status

Experience Status	Mean SJT Score	Standard Deviation
Student	21.35	2.98
Manager	24.39	2.58

52

Hypothesis 5 stated that scores on the situational judgment test under different organizational manipulations would be moderated by job experience. This hypothesis was tested by examining the significance of the interaction between Experience and the three Leadership Culture manipulations. The ANCOVA showed a significant interaction between Experience and Leadership Culture [$F(2, 387) = 3.804$, $p < .023$]. Mean SJT scores and standard deviations are provided in Table 10. Boxplots provide a graphical representation of the interaction in Figure 5.

Table 10: Mean SJT Scores and SDs by Experience and Leadership Condition

Status	Condition	Mean SJT Score	St. Dev.
Student	Participative/Supportive	21.93	3.02
	Control	21.27	3.24
	Directive/Achieving	20.97	2.69
Manager	Participative/Supportive	24.11	2.52
	Control	25.15	1.45
	Directive/Achieving	23.79	3.30

53

Figure 5: Student & Manager SJT Scores Across Conditions

Leadership Culture
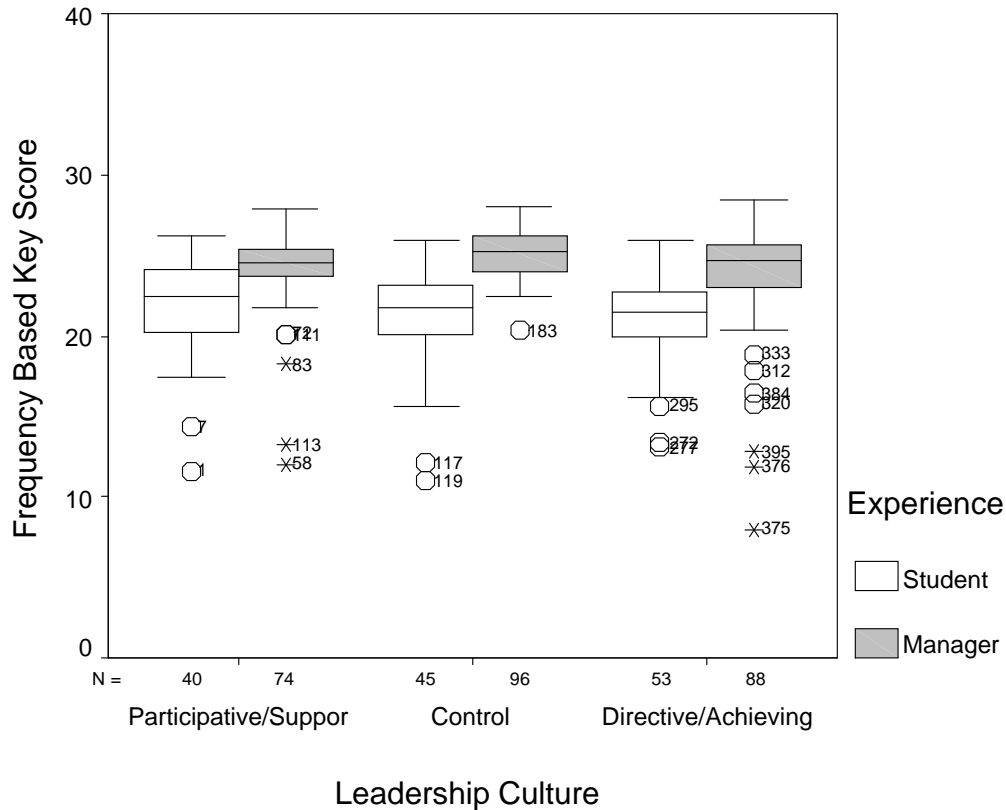
This hypothesis was also tested at a response level by comparison of response frequencies through Chi-square analysis. Chi-square coefficients and *Phi* values were computed for each item within the Student and the Manager sample to test whether the frequency distributions differed across the three culture manipulations for each item. Chi-square coefficients were significant for 28 out of 104 analyses (52 Student items and 52 Manager items), or 27%, suggesting that item-level responses differed between the culture conditions. Differences in item-level responses among experience status and the manipulated conditions may suggest that participants tend to answer differently across experience status and condition, regardless of whether their behaviors would be considered effective (*e.g.*, different ineffective behaviors may be endorsed in different

54

conditions). The tables of response frequencies are presented in Appendix G; Chi-

square and *Phi* values are presented in Appendices H and I. Hypothesis 5 was supported.

The analysis of covariance (ANCOVA) that was used to analyze the three main

effects and the two interactions is included for reference as Table 11.

Table 11: ANCOVA Model of Main Effects and Interactions

Source	Type III Sum of Squares	Degrees of Freedom	Mean Square	F	Sig.
Corrected Model	982.98(a)	8	122.87	17.19	.000
Intercept	1409.07	1	1409.07	197.13	.000
Experience	730.67	1	730.67	102.22	.000
Leadership Culture	64.76	2	32.38	4.53	.011
Agreeableness	29.75	1	29.75	4.16	.042
Experience * Leadership Culture	54.38	2	27.19	3.80	.023
Agreeableness * Leadership Culture	18.97	2	9.49	1.33	.266
Error	2766.30	387	7.15		
Total	219200.62	396			
Corrected Total	3749.28	395			

a R Squared = .262 (Adjusted R Squared = .247)

Discussion

This research was intended to be a preliminary step in critically examining

unexplored constructs at both a personal and situational level that explain variance in

responses to situational judgment tests, as recommended by earlier research (Ployhart &

Weekley, 2006). There are limitless factors and constructs that may contribute to

explained variance for this type of test, not only on a composite level, but also on an item

level or even a response level. The construct of Experience has been widely researched

in the past, Agreeableness has received some attention; both were considered deserving

of a second look in the context of SJT response. The "new" construct of Organizational

55

Leadership Culture was included in this research because of this author's first hand experience with commercial selection test publishers. It is not uncommon for commercial test publishers to base passing scores on "off the shelf" or "canned" situational judgment tests (and other kinds of tests) on norms that may or may not be appropriate for every organization, especially due to differences in organizational cultures.

This research suggests that Experience might play an important role in shaping responses to situational judgment tests. The result must be interpreted with caution because there are variables that are confounded with the operational definition of Experience that were not controlled in this experiment (*e.g.*, age, education, and the organizational and cultural definitions of experience explained earlier in this research).

The Organizational Leadership Culture and Agreeableness constructs may also provide a small contribution to SJT response. However, the constructs explored did not fit the model as expected, and undeniably the contributions of Leadership Culture and Agreeableness were small in comparison to the effect of Experience. One of the reasons that Organizational Leadership Culture may not have shown a robust significant effect is that participants may have seen the culture manipulations as transparent. Those participants in the Control condition actually had the highest scores on the SJT, suggesting that maybe the manipulations caused participants in those conditions to carefully consider their way of thinking about the job. Alternatively, it is possible that the Directive/Achieving condition was the most difficult of the three to interpret and

56

impression-manage to fit, hence this condition showed lower mean scores than the other conditions, as put forward in Hypothesis 5.

Agreeableness showed a small significant effect on SJT responses. This finding is in keeping with previous research and suggests that learning more about how personality traits are related to judgment will be useful in construct explication of situational judgment (*e.g.*, Motowidlo, Hooper & Jackson, 2006; McDaniel & Nguyen, 2001). This study suggests that high Agreeableness may be related to higher scores on situational judgment tests. Again, this result must be interpreted with caution, as it is also likely that Agreeableness is perceived as a universally desirable personality trait among jobs that require considerable interpersonal interaction (Barrick & Mount, 2005). Further, the main effect for Agreeableness just barely achieved significance; this may be due to range restriction. There is a certain degree of confound between having an agreeable personality and asking participants to behave in a certain way. Asking participants to respond as though they are very interested in getting a job with a company described in a certain way may be easier for participants with a higher level of Agreeableness, regardless of the description.

The ability of highly agreeable people to better impression manage regardless of Leadership Culture could explain why there was such a small effect for Agreeableness, as well as why there was no interaction effect between Agreeableness and the manipulations of Leadership Culture when highly Agreeable participants were initially expected to have lower SJT scores in the Directive/Achieving manipulation, for example.

57

When the interaction between Experience and the manipulation of Leadership Culture was explored, significant influence on SJT response was observed. This is explained in part by the likelihood that job experience influences how well respondents "fit" themselves into different aspects of organizational culture; those with more experience are likely to do what they have done in the past, regardless of the leadership culture of the organization, because they expect it to work. Alternatively, those with less experience are more likely to see a need to fit in with implicit organizational policies to succeed.

Limitations of this Research

This research is a meaningful early step in construct explication, but it was not without its problems. One of the most obvious concerns is that access to the original management sample was lost after a small amount of pilot data was obtained. This loss was beyond the control of the experimenter and his colleagues at that organization, and is likely all too common when utilizing applied samples. It was fortunate that another Management sample became available, but it is clear that potential differences could exist between the organization used in the pilot research and the organization used in the experimental research, while the student sample was taken from the same undergraduate population both times. While this change did not necessarily limit the validity of this research, it resulted in a rethinking of how to score participant responses after the study had been planned.

Another concern was that the student experimental sample was small compared to the management experimental sample. This is unusual, as it is typically easier to recruit

58

students than professionals. However, having fewer students resulted in unequal cell sizes, and may have adversely affected power. This is especially important in light of the fact that several significant relationships barely met the [$p < .05$] convention. While the size of the student sample was partially under the control of this experimenter, a decision was made to limit student participation to a single semester to allow other researchers to take advantage of the university's undergraduate research pool.

Additionally, the management samples for both the pilot and experimental research were convenience samples. Participants in the managerial sample of the pilot study agreed to participate as a favor to this researcher and may therefore have been unfairly biased toward the research. Participation in the managerial sample of the experimental research was limited to those managers who deigned to respond to the voluntary research request. It is virtually impossible to report metrics for those managers who chose not to participate.

Much of the data tested showed a lack of homogeneity of variance (according to Levene's test for Homogeneity of Variance). This presents some statistical concerns when interpreting any significant relationships. However, where possible, results were reported with unequal variances assumed in an attempt to correct for this finding.

An obvious drawback to this research was the finding that numerous items on the SJT were answered consistently regardless of condition or experience. Psychometrically, this may suggest these items were too simple, or that particular sets of responses contained too many poor or transparent distracters. However, since many of these items were taken from actual situational judgment inventories that have been validated and

59

found to be reliable, this concern may be a detriment to the criterion validity of the instruments themselves as much as the research methods. The additional step of editing the instrument to remove or revise those items that were endorsed in the same way across all conditions/experience levels should be considered in replications of this research. Eliminating consistently endorsed items might show more significant effects within this research. Elimination or revision of these items could also be beneficial from an applied perspective; although it would seem that items that are consistent across conditions would be useful to test developers, it is also possible that those items are not particularly predictive of organizational fit, or even performance across different organizations.

Directions for Future Research

It is important that construct explication be continued to discover more about how and why situational judgment works in selection. This research could be advanced by performing content analysis of the items that were sensitive to the effects of Organizational Leadership Culture to determine if the items possess commonalities, and what distinguishes them from the items that were not significantly sensitive (per Appendices F, H and I). Likewise, a replication of this study would benefit from a higher degree of fidelity and realism if participants applying to actual organizations that clearly differ in their Organizational Leadership Cultures could be tested instead of relying on descriptions of fictitious organizations.

It would also be very interesting to perform similar research with different managerial positions and/or different industries. Many of the SJT items were originally written for staff management positions; that is, positions that have supervisory

60

responsibilities over administrative personnel rather than those connected expressly to the product or service of the organization. The managerial sample that participated in the pilot research consisted entirely of staff managers, while the managerial sample that participated in the experimental research consisted entirely of line managers. The implication is that managers who are involved in these two different functions may have very different perspectives about what to do in the same situation, based on distinct differences in achievement and support orientations in staff vs. line managers (Church & Waclawski, 2001). A straightforward step toward better understanding the constructs behind SJT response would be to pilot an SJT on similar job titles in different organizations (or different divisions or even teams within the same organization) to explore differential criterion-related validity. It is possible that observed differences in the utility of the same SJT could be explained by situational or personality variables that are characteristically different in different organizations or micro-organizations.

Additionally, different aspects of leadership or management could be explored, rather than limiting the instrument to addressing "dealing with subordinates." There continues to be a call for research on increased specificity in situational judgment tests; that is, focusing SJTs toward specific job tasks and abilities (*e.g.*, Ployhart & Weekley, 2006; McDaniel *et al.*, 2001; Weekley & Jones, 1999). Although dealing with subordinates was considered one of the most important job duties of managers in general, there are numerous other job duties that are likely equally important and equally common across all management jobs, for example conducting performance appraisals, handling escalations (*i.e.*, situations that are too demanding for subordinates to handle that are

61

consequently passed up to a manager), and dealing with internal or external customers. These and other job duties would likely show different effects in the context of Path-Goal Theory; for example, the domain of conducting performance appraisals would be expected to be profoundly influenced by the Participative or Directive nature of an organization, while handling escalations might be more heavily influenced by the Supportive or Achievement-oriented nature of an organization.

The use of Path-Goal Theory was oversimplified for the purpose of limiting the manipulated conditions in this research. The combination of Participative with Supportive, and Directive with Achievement-oriented characteristics were used to illustrate stereotypes of an employee-friendly vs. an authoritarian, profit driven organization. It was expected that these stereotypes would be simplest for participant interpretation and impression formation. However, other combinations of the four Path-Goal characteristics are entirely conceivable. Organizations might alternatively be considered Participative and Achievement-oriented, or Directive and Supportive. These characteristics could be combined in multiple ways in future research. It would be ideal to recruit participants from organizations that have decidedly different Leadership Cultures; this would render unnecessary the manipulations used in this study.

Likewise, future researchers should consider evaluating response choices on continua other than "effectiveness". For example, it would be useful and possibly meaningful to have pilot raters evaluate response choices for agreeableness, conscientiousness, or other personality variables to determine how important these

62

variables might be in determining what courses of action are effective. This follows from the implicit trait policy research of Motowidlo and colleagues (2006).

Other ways that SJTs should be examined in the future include additional psychometric research, such as a comparative study that can demonstrate the most reliable and valid way to develop and score SJTs (Weekley, Ployhart & Holtz, 2006). For example, a larger expert pilot sample could be obtained so that the Conditional Key method could be effectively compared to the Frequency Key method.

The future of situational judgment is bright because of the method's low cost and high validity, as described above. As the U.S. becomes increasingly a service-industry based nation, it is likely that SJTs will become even more useful in predicting performance because of their relationship with procedural knowledge (Ployhart & Weekley, 2006). That is, while work sample tests are highly predictive for jobs requiring workers who are skilled in specific physical tasks (Schmidt & Hunter, 1998), many service jobs require abilities less objective in nature (*e.g.*, diagnosis, troubleshooting, creativity, *etc.*) that are difficult to test. This research found small, but significant differences based on a minute aspect of cultural differences (leadership culture) within a larger homogeneous culture (U.S. organizations). It follows that SJT research should take a global perspective in the near future, comparing situational judgment across regional and international cultures that are undoubtedly more diverse, possibly leading to richer differences in how judgment is engaged.

This research demonstrates that the question of "why and how" situational judgment tests work likely requires a multifaceted answer. It is vital that future research

63

focus on additional personality and situational traits to learn more about what affects situational judgment. A better understanding of the constructs that influence SJT responses could ultimately lead to more effective tests that can better predict performance, turnover, job fit and other outcomes central to the field of employment testing.

64

References

Barrick, M.R. & Mount, M.K. (1991). The Big-5 personality dimensions and job

performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Barrick, M.R., & Mount, M.K. (1993). Autonomy as a moderator of the relationships

between the big five personality dimensions and job performance. *Journal of

Applied Psychology, 78*, 111-118.

Barrick, M.R. & Mount, M.K. (1993). Yes, personality matters: Moving on to more

important matters. *Human Performance, 18,* 349-372.

Barrick, M.R., Mount, M.K. & Judge, T.A. (2001). Personality and performance at the

beginning of the new millennium: What do we know and where do we go next?

International Journal of Selection and Assessment, 9, 9-30.

Bass, B.M. (1990). *Bass & Stogdill's Handbook of Leadership*. New York: Free Press.

Borman, W.C., Hanson, M.A. & Hedge, J.W. (1997). Personnel selection. *Annual

Review of Psychology, 48,* 299-337.

Callinan, M. & Robertson, I.T. (2000). Work sample testing. *International Journal of

Selection and Assessment, 8,* 248-160.

Church, A.H. & Waclawski, J. (2001). Hold the line: An examination of line vs. staff

differences. *Human Resource Management, 40, 21-34.*

Clause, C.S., Mullins, M.E., Nee, M.T., Pulakos, E., & Schmitt, N. (1998). Parallel test

form development: A procedure for alternate predictors and an example.

Personnel Psychology, 51, 193-208.

65

Clevenger, J, Pereira, G.M., Wiechmann, D., Schmitt, N., & Harvey, V.S. (2001).

Incremental validity of situational judgment tests. *Journal of Applied Psychology, 86,* 410-417.

Coleman, P.T. (2004). Implicit theories of organizational power and priming effects on managerial power-sharing decisions: An experimental study. *Journal of Applied Social Psychology, 34,* 297-321.

Costa, McCrae, R.R. & Dye, D.A. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12,* 887-898.

Cucina, J.M., Vasilopoulos, N.L., & Leaman, J.A. (2003). *The bandwith-fidelity dilemma and situational judgment test validity.* Poster presented at the 18th meeting of the Society of Industrial and Organizational Psychology, Orlando, Florida.

D'Alessio, A.T. (1994). Predicting insurance agent turnover using a video-based SJT. *Journal of Business & Psychology, 9,* 23-32.

Dalen, L.H., Stanton, N.A., & Roberts, A.D. (2001). Faking personality questionnaires in personnel selection. *Journal of Management Development, 20,* 729-742.

Daley, B.J. (1999). Novice to expert: An exploration of how professionals learn. *Adult Education Quarterly, 49,* 133-147.

Dreyfus, H. & Dreyfus, S. (1985). *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer.* New York: Free Press.

66

Evans, (1996). M.G. (1996). R.J. House's "A Path-Goal theory of leadership effectiveness." *Leadership Quarterly, 7,* 305-309.

Fleishman, E.A. (1995). Consideration and structure: Another look at their role in leadership research. In F. Dansereau & F.J. Yammarino (Eds.), *Leadership: The Multi-level Approaches* (pp. 51-60). Stamford, CT: JAI Press.

Genberg, V. (1992). Patterns and organizing perspectives: A view of expertise. *Teaching and Teacher Education, 8,* 485-495.

Greenberg, J. & Baron, R.A. (1997). *Behavior in Organizations, 6th Ed.* Upper Saddle River, NJ: Prentice-Hall.

Greguras, G. J. (2005). Managerial experience and the measurement equivalence of performance ratings. *Journal of Business and Psychology, 19,* 383-397.

Hanson, M.A. & Borman, W.C. (1992). Development and construct validation of the Situational Judgment Test (Institute Report #230). Submitted to the U.S. Army Research Institute.

Hanson, M.A. & Ramos, R.A. (1996). Situational judgment tests. In R. S. Barrett (Ed.) *Fair Employment Strategies in Human Resource Management*, Westport, CT: Quorum. 119-124.

Hogan, R. (1998). What is personality psychology? *Psychological Inquiry, 9,* 152-153.

Hooper, A.C., Cullen, M.J. & Sackett, P.R. (2006). Operational threats to the use of SJTs: Faking, coaching and retesting issues. In J.A. Weekley & R.E. Ployhart (Eds.), *Situational Judgment Tests: Theory, Measurement and Application* (pp. 205-232). Mahwah, NJ: Lawrence Erlbaum Associates.

67

House, R.J. (1971). A path-goal theory of leader effectiveness. *Administrative Science Quarterly, 16,* 321-328.

House, R.J. (1996). Path-Goal theory of leadership: Lessons, legacy, and a reformulated theory. *Leadership Quarterly, 7,* 323-352.

House, R.J. & Mitchell T.R. (1974). Path-Goal theory of leadership. *Journal of Contemporary Business, 3,* 81-97.

House, R.J., Shane, S.A. & Herold, D.M. (1996). Rumors of the death of dispositional research are vastly exaggerated. *Academy of Management Review, 21,* 203-224.

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/). Internet Web Site.

Judge, T.A., Piccolo, R.F. & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology, 89,* 36-51.

Kahai, S.S., Sosik, J.J. & Avolio, B.J. (1997). Effects of leadership style and problem structure on work group process and outcomes in an electronic meeting system environment. *Personnel Psychology, 50,* 121-146.

Kolz, A.R., McFarland, L.A., & Silverman, S.B. (1998). Cognitive ability and job experience as predictors of work performance. *Journal of Psychology: Interdisciplinary & Applied, 132,* 539-548.

68

Kristof-Brown, A., Barrick, M.R. & Franke, M. (2002). Applicant impression
management: Dispositional influences and consequences for recruiter perceptions
of fit and similarity. *Journal of Management, 28,* 27-46.

Lee, S., Choi, S.K., & Choe, I.-K. (2005). Issues in measuring situated cognition: Cases
of situational judgment tests. Paper presented at the 20th meeting of the Society
for Industrial and Organizational Psychology, Dallas, Texas.

McCrae, R.R. & Costa, P.T., Jr. (1985). Updating Norman's "adequate taxonomy":
Intelligence and personality dimensions in natural language and in questionnaires.
Journal of Personality and Social Psychology, 49, 710-721.

McCrae, R.R. & Costa, P.T., Jr. (2003). *Personality in Adulthood: A Five Factor Theory
Perspective, 2nd Ed.,* New York: Guilford Press.

McCrae, R.R., Costa, P.T., Jr. & Busch, C.M. (1986). Evaluating comprehensiveness in
personality systems: The California Q-Set and the Five-Factor Model. *Journal of
Personality, 54,* 430-446.

McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A., & Braverman, E.P.
(2001). Use of situational judgment tests to predict job performance: A
clarification of the literature. *Journal of Applied Psychology, 86*, 730-740.

McDaniel, M.A., Schmidt, F.L., & Hunter, J.E. (1988). Job experience correlates of job
performance. *Journal of Applied Psychology, 73,* 327-330.

McDaniel, M.A., & Nguyen, N.T. (2001). Situational judgment tests: A review of
practice and constructs assessed. *International Journal of Selection and
Assessment, 9*, 103-113.

69

McElreath, J., & Vasilopoulos, N.L. (2002). *Situational judgment: What is most or least likely to happen here?* Paper presented at the 17th meeting of the Society of Industrial and Organizational Psychology, Toronto, Ontario, Canada.

McFarland, L.A. & Ryan, A.M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85,* 812-821.

Motowidlo, S.J., Dunnette, M.D. & Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75,* 640-647.

Motowidlo, S.J., Hooper, A.C. & Jackson, H.L. (2006). A theoretical basis for situational judgment tests. In J.A. Weekley & R.E. Ployhart (Eds.), *Situational Judgment Tests: Theory, Measurement and Application* (pp. 57-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Motowidlo, S.J. & McDaniel, M.A. (2005). Situational judgment tests part 2: A theory based on implicit trait policies. Paper presented at the 20th meeting of the Society of Industrial and Organizational Psychology, Los Angeles, California.

Motowidlo, S.J. & Tippins, N. (1993). Further studies of the low-fidelity simulation in the form of a situational inventory. *Journal of Occupational & Organizational Psychology, 66,* 337-344.

Mowry, H.W. (1957). A measure of supervisory quality. *Journal of Applied Psychology, 41,* 405-408.

Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes:

Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66,* 574-583.

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric Theory, 3rd Ed.* New York:

McGraw-Hill.

Ones, D.S. & Vishwesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11,* 245-269.

Oshagbemi, T. (2004). Age influences on the leadership styles and behaviour of managers. *Employee Relations, 26,* 14-29.

Ployhart, R.E. & Weekley, J.A. (2006). Situational judgment: Some suggestions for future science and practice. In J.A. Weekley & R.E. Ployhart (Eds.), *Situational Judgment Tests: Theory, Measurement and Application* (pp. 345-350). Mahwah, NJ: Lawrence Erlbaum Associates.

Quinones, M.A., Ford, J.K. & Teachout, M.S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology, 48,* 887-910.

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

71

Schmidt, F.L. & Hunter, J.E. (1993). Tacit knowledge, practical intelligence, general mental ability and job knowledge. *Current Directions in Psychological Science, 2,* 8-9.

Schriescheim, C.A., Tepper, B.J. & Terault, L.A. (1994). Least preferred coworker score, situational control, and leadership effectiveness: A meta-analysis of contingency model performance predictions. *Journal of Applied Psychology, 79,* 561-573.

Silvester, J., Anderson-Gough, F.M., Anderson, N.R. & Mohamed, A.R. (2002). Locus of control, attributions and impression management in the selection interview. *Journal of Occupational and Organizational Psychology, 75,* 59-76.

Somech, A. (2003). Relationships of participative leadership with relational demography variables: A multi-level perspective. *Journal of Organizational Behavior, 24,* 1003-1018.

Sternberg, R.J. (1985). Implicit theories of intelligence, creativity and wisdom. *Journal of Personality and Social Psychology, 49,* 607-627.

Sternberg, R.J. & Wagner, R.K. (1993). The g-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science, 2,* 1-5.

Sternberg, R.J., Wagner, R.K., Williams, W.M., & Horvath, J.A. (1995). Testing common sense. *American Psychologist, 50,* 912-927.

Stogdill, R.M. & Coons, A.E. (1957). *Leader Behavior: Its Description and Measurement.* Columbus, OH: Ohio State University, Bureau of Business Research.

72

Sturman, M.C. (2003). Searching for the Inverted U-Shaped Relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management, 29,* 609-640.

Styhre, A. (2004). Rethinking knowledge: A Bergsonian critique of the notion of tacit knowledge. *British Journal of Management, 15,* 177-188.

Tett, R.P., Jackson, D.N. & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology, 44,* 703-743.

U.S. Department of Labor Bureau of Labor Statistics Standard Occupational Codes (*http://www.bls.gov/oes/current/oes_stru.htm*). Internet Web Site.

Vishwesvaran, , Ones, D.S. & Hough, L.M. (2001). Do impression management scales in personality inventories predict managerial job performance ratings? *International Journal of Selection and Assessment, 9, 277-289.*

Wagner, R.K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52,* 1236-1247.

Wagner, R.K. & Sternberg, R.J. (1991). *Tacit Knowledge Inventory for Managers*. San Antonio, TX: The Psychological Corporation.

Weekly, J.A. & Jones, C (1997). Video-based Situational Testing. *Personnel Psychology, 50,* 25-49.

Weekly, J.A., Ployhart, R.E. & Holtz, B.C. (2006). On the development of situational

judgment tests: Issues in item development, scaling and scoring. In J.A. Weekley

& R.E. Ployhart (Eds.), *Situational Judgment Tests: Theory, Measurement and*

Application (pp. 157-182). Mahwah, NJ: Lawrence Erlbaum Associates.

Yukl, G. (1994). *Leadership in Organizations* (3rd Ed.). Englewood Cliffs, NJ: Prentice-

Hall.

Appendices

<u>Instructions to Participants (*Participative/Supportive Culture* Condition):</u>

Please carefully read the following information about the company where you are applying and keep it in mind as you respond to this assessment:

ABC Company has a well-known reputation for being employee friendly and having strong family values. It is common knowledge that Joseph Meyers, the CEO of ABC Company, worked his way up from the mailroom to make ABC what it is today: a Fortune 500 company with one of the highest employee satisfaction ratings in the business. At ABC, the philosophy is, behind every good manager, there is a team of great people. The leadership style at this company is Participative and Supportive. That means that managers should allow their subordinates to contribute ideas and even question management, if necessary. It also means that providing personal and professional support to employees is very important. Managers at ABC are expected to be highly supportive of their employees and to provide a balance between the demands of work and family. Your employees are encouraged to contribute to the planning and execution of tasks. ABC places a lot of weight on feedback from subordinates when they evaluate performance and award raises.

You received the following confidential email from a friend of a friend (whom you've met a few times) who works at ABC Company:

Hey, excited you might be working with us at ABC! It is such a great company to work for. The environment is so friendly, and supervisors are always supportive. You really get the feeling that we're a team, from the very top to the newest employee. Everyone looks out for everyone else, and everyone's voice is heard and acknowledged. Let me know if I can put in a good word for you!

<u>Instructions to Participants (*Directive/Achieving Culture* Condition):</u>

Please carefully read the following information about the company where you are applying and keep it in mind as you respond to this assessment:

XYZ Company has a well-known reputation for being aggressive and competitive. At XYZ, the "bottom line" and profitability always come first. It is common knowledge that Joseph Meyers, the CEO of XYZ Company, doesn't like failure. He has had to lay off a lot of employees and step on a lot of toes to make XYZ Company what it is today: a Fortune 500 company with one of the highest profit margins in the business. At XYZ, the philosophy is, managers are accountable; successful managers are well rewarded, and

unsuccessful managers are gone. The leadership style at this company is Directive and Achievement-oriented. That means that getting results and meeting goals is very important. When your performance evaluations come around, the biggest question is "How much did you increase profitability this year?" Managers at XYZ are expected to delegate duties to their employees and make sure they are doing what they are supposed to from day to day, because at the end of the day, what gets done or doesn't get done is management's responsibility.

You received the following confidential email from a friend of a friend (whom you've met a few times) who works at XYZ Company:

> Hey, excited you might be working with us at XYZ! It is such a great company to work for. We're not one of those touchy-feely companies, but everyone knows what they need to do and we always get the job done. No one wastes a lot of time trying to get consensus. Management knows what they're doing and you can't argue with their results. Let me know if I can put in a good word for you!

Instructions to Participants (*Neutral/Control* Condition):

Please carefully read the following information about the company where you are applying and keep it in mind as you respond to this assessment:

NYT is a Fortune 500 company with over ten thousand employees nationwide. Business is growing and NYT will soon expand into international markets.

You received the following confidential email from a friend of a friend (whom you've met a few times) who works at NYT Company:

Hey, excited you might be working with us at NYT! It is such a great company to work for. Let me know if I can put in a good word for you!

Welcome!

This assessment involves the use of two measures: a "situational judgment test" or SJT, and an opinion survey.

An SJT is a kind of test that presents you with questions about how employees should react in realistic work situations.

An opinion survey asks you to provide your honest thoughts and experiences. In this case, the opinion survey will ask some questions about what you thought of the SJT as well as some questions about you.

What to Expect

Imagine that you are applying for a job in management with a real company. As part of your application, the company has asked you to take the following situational judgment test, the SJT-280.
As you take the SJT-280, please respond to the questions as though you really want this management job and it is very important to you that you get it. Remember, everything you will see for the first part of this assessment is based on a real company and a real employment test. So please try your best!

To help you do your best on this assessment, you will begin by reading some information about the company where you are applying.

During the SJT part of this assessment (26 questions), it will be up to you to read each question carefully and choose

One *Most Effective Response*

and

One *Least Effective Response*

based on what you know about this company.

Remember to try your best, just as you would if you were really applying for a job!

During the opinion survey part of this assessment (31 questions) you will be asked to answer some questions about yourself and about the SJT assessment that you just finished.

If you need to quit at any time, you can always exit and return to this assessment later. When you click on the link to this site, your computer will automatically return you to where you left off (you must be using the same computer).

Before we begin the assessment, on what date (day of the month) were you born?
1st-9th
10th-19th
20th-31st

(Note: Instructional Manipulation Information is presented here)

Choose one *Most Effective Response* and one *Least Effective Response*, based on what you know about this company.*

1) You ask an experienced employee to do a particular task. The employee responds curtly, "That's not my job." What should you do?

Ask the employee if something is bothering him or her.
Do the task yourself but discipline the employee later.
Explain why the task is important and ask the employee to reconsider.
Get someone else to do it and talk with the employee later.
Insist that it is part of the employee's job and see that he/she does it.

2) Of the following, which one method would good managers use most often for monitoring and controlling the work of employees?

Activity status reports by employees.
Feedback from others familiar with the employees' work.
Hands on inspection.
Impromptu telephone calls and meetings.
Time and action calendars.

---

* This instruction is repeated for the first 3 situational judgment items. It is presented in this Appendix only once for the sake of brevity.

3) At the end of the day, an employee's car won't start. She is in a hurry to pick up her children. You have not finished closing the office for the evening. What should you do?

Allow the employee to use the phone while you finish closing the office.
Lend the employee cab fare so she can get her children.
Take the employee in your car to get her children, then come back and close the office.

4) Of the following, which do you feel is the most effective way for a Manager to improve communications with Employees?

Adopt an "open door" policy.
Ask employees a lot of questions.
Be very visible and accessible.
Schedule regular group meetings.
Schedule regular, one-on-one meetings with employees.

5) One of the Managers reporting to you is reluctant to hold his employees accountable for results. He is too willing to accept reasons for why things can't get done according to standard. What should you do?

Counsel the Manager regarding his performance.
Hold a meeting with the Manager and his employees regarding the importance of meeting performance expectations.
Train the Manager on how to set performance standards and follow-up with employees.

6) You are a new Manager. Just before you took over, one of the supervisors working under you was promoted into greater responsibility. She is highly intelligent, but not very experienced. You start to get lots of complaints from the employees that she is inflexible and has a philosophy of "my way or the highway." Employee turnover in her area has been rising since she got there. What should you do?

Check it out with a few employees.
Consult with the prior Manager.
Do an employee survey and review the results with this supervisor.
Have a small meeting with this supervisor and some of the persons complaining.
Tell this supervisor what you've been hearing.

7) You are a Manager. An employee keeps showing up for work late. She is otherwise a good Employee. However, the other employees are noticing that she's coming in late, and it's setting a bad example. Since she started coming in late, you've been trying to find out why. She hasn't been willing to tell you until now: her husband has become a serious problem. She has already been referred to counseling. What else should you do?

Ask her what help she needs.
Be empathetic but explain the need to be on time.
Consult with your boss.
Give her more time to work things out.
Offer to change her schedule.

8) You are informing the employees in your work unit that they have just won an award for achieving the best quality record in the company. What should you say?

"Great job team! Now that we have quality taken care of, let's strive to achieve the same recognition for quantity."
"Great job team! See how far a little hard work and dedication can take you?"
"This is a proud moment for all of us and clear evidence that your hard work has been recognized in this company."
"This is a proud moment for all of us. At the same time I know we can achieve even higher quality standards. Let's show them what we can really do."

9) A serious problem has arisen with a project that your work team is currently working on. What should you say?

"For some reason we seem to be having a problem concerning this project. Here are my thoughts."
"Let's analyze the project step by step and determine what caused the problem."
"The good news is that we can solve the problem. Now here's what we all need to do."
"We seem to be having some problems with our current project. Does anyone have any suggestions?"

10) One of your newer employees is not pulling her weight in the sales department. For the second month in a row, she has not sold the required amount of goods. What should you do?

Ask one of her experienced coworkers to coach her and help her develop better sales skills.
Give her a month to improve; if she doesn't do better by then, talk to her about whether this company is the right fit for her.
Give her some "easy customers" that will make her sales numbers look better until she can get in the swing of the job.
Ignore the problem; people usually do better after they get some experience.

11) A top performer in your department has just asked you for next week off due to a death in the family. Your whole department will be needed next week to prepare an important report for your company's CEO, and usually no one is given time off during that week. What should you say?

81

"I'm sorry I can't spare you for a whole week. The best I can do is three days. We really need you for this report."

"Take the week off. But make sure you find someone to cover your portion of the report that is due next week."

"Take the week off. I wouldn't do this for just anyone, but you are one of my top performers and it is an emergency situation."

12) Of the following, what should a manager do when introducing a new policy that is likely to be unpopular?

Call a group meeting to introduce and discuss the policy.
Meet with each employee individually to discuss the policy.
Post the policy on a bulletin board and invite questions.
Send a memo to each employee explaining the policy.
Sound out employee opinions before announcing the policy.

13) In a staff meeting, you propose that a new project be handled in the usual way, but one of your employees (that you don't always get along with) interrupts to say he doesn't think that the "usual way" will work in this case. What should you say?

"Do you have a better idea? If so, you should have mentioned it to me before this meeting."

"Okay, let's hear your plan out, and if it sounds good, I'll expect you to take the lead on this project."

"Please don't interrupt. If you have a different idea, let's talk about it in my office."

"The usual way has worked for a long time, and I'm going to make an executive decision here to at least give it a try this time."

14) Jill is one of your hardest workers, but she sometimes has trouble getting along with others in the department. She recently came to you with a complaint about Tom, one of your newest employees. Jill reports that Tom is dragging down the rest of the department, taking frequent breaks and generally not getting much work done. You decide to confront Tom with this information. His immediate reply is, "I'll bet Jill was the one who came to you about this. She's been giving me a hard time ever since I started here." What should you say?

"I've spoken to several people in the department, and they are all giving me the same story: you're not getting your work done."

"It doesn't matter who it was. I'm concerned to hear this kind of information about you, Tom."

"Matter of fact, it was Jill. You'd do better to be more like her, and complain less about her, Tom."

"Tom, it seems like you haven't been happy here since you started. What can I do to help you get established with us?"

"Why don't you tell me your side of the story, Tom."

15) Betsy has three children, and it seems like they are always getting sick. The result is that Betsy has to take a lot of time off on short notice to be with her children. It seems her work is suffering because of it, too. Some of your other employees are getting upset about Betsy's absences, though no one has complained to you directly yet. What should you do?

Let Betsy know that her time off is almost used up, and that taking additional time off could result in a written warning, leading up to termination.

Let Betsy know that her work is suffering and that if her work doesn't improve, you will have no choice but to give her a written warning, leading up to termination.

Let Betsy know you have set a meeting to discuss some possible changes she could pursue to improve her commitment to work, such as using onsite child care, getting a babysitter, or getting her significant other to watch the kids when they are sick.

Let Betsy know you have set a meeting to discuss some possible scheduling options that will better fit her lifestyle, such as working part-time, working a 4 day week, or changing to evening shifts.

Wait until someone approaches you about the problem. If no one has made an official complaint, it probably isn't serious enough to worry about.

16) Mike is probably one of the most productive people in your department. He is nearly always punctual, organized and diligent, and the work he does is first rate. He has a reputation in the department for being sort of a loner and not being too talkative with other employees. The other day in the break room, you see Mike sitting by himself drinking coffee. What should you do?

Go over and strike up a conversation with Mike about an interesting movie you saw recently.

Go over and strike up a conversation with Mike about how things are going at work.

Go over and tell Mike, "I wish I had 10 other employees just like you."

Say, "Hi Mike," as you pass him on your way out.

Talk to one or two of your employees about helping Mike feel more included in the company.

17) Pat, one of your employees, just shouted at a very important client over the telephone, and everyone in the office heard it. You ask what's going on, and Pat says, "I'd rather not talk about it." What should you say?

"Okay, but I can't tolerate you talking like that in this office."

"Okay, but I can't tolerate you talking like that to one of our most important clients."

"Okay, but I'm going to have to ask you to call back the client and apologize right now."

"Okay, but I'm going to have to ask you to leave the area until you settle down."

"Okay, but if you tell me what's going on, maybe I can help."

18) The home office informed you that you must have a workplace seminar next week to comply with company regulations. They will send a speaker on one of the following topics, all of which have been well received by employees in the past. You must make a decision right away so they can schedule the speaker. What is your choice?

A Stitch in Time: Maximizing Workplace Efficiency
Balancing the Scales: How to Find a Happy Medium Between Work and Family
Brainstorming: Getting the Most Out of Everyone's Ideas
Closing the Deal: Selling Yourself and Your Company

19) An employee who is supervised by one of your subordinates has asked to talk with you. He says your subordinate is guilty of some possible violations of company policy. This is the first you have heard of any problem with the subordinate in question. In the company, employees are expected to take all concerns to their immediate supervisor before going to anyone else. The employee who wants to talk with you has not discussed this matter with his supervisor (your subordinate) because of its sensitive nature. What should you do?

Refuse to meet with this employee until he has first discussed this matter with his supervisor (your subordinate).
Meet with this employee, but only with your subordinate present.
Meet with this employee to discuss the matter, and then decide whether you need to meet with your subordinate.
Meet with your subordinate to discuss the matter, and then decide whether you need to meet with this employee.

20) You are a manager, and you have an outstanding work team. Lately, you have been getting complaints from your team. They say they seem to get assigned every project that comes along. You feel that this is probably true. What should you do?

Talk with your director and ask him if he thinks your team is getting more than their fair share of new assignments.
Talk with your director's supervisor and tell him that your team is getting more assignments than any other team in the organization.
Talk with other managers in your department. Explain that new projects should be divided up evenly.
Tell your team that because they are the best in the department, sometimes they have to pick up the slack for other work teams.

21) You are a newly assigned manager. Your department director has given you guidelines on how to run your work team. What input should you ask from your team?

Ask your team what they think is the best way to handle projects, but keep in mind that it is your decision to make.
Tell your team exactly how you want them to handle projects. Do not ask for their input unless a problem comes up.
Assuming the team has been performing well let them decide how they would like to be run. Try to merge what they say with your director's guidelines.
Determine how the team has been performing, combine that information with your director's guidelines, then run the team in the manner that you believe is best.

22) One of your employees, Bob, has always performed his work in an excellent manner, and seemed to be happy working for you. Today he came to you and said he wanted to transfer to a different work team and if possible to a different department. What should you do?

Talk with Bob. If you find that he really wants to transfer, then help him to do so.
Agree to begin the transfer if that's what Bob really wants. Do not process the transfer too quickly, and in the meantime, try to help Bob with whatever is leading him to request the transfer.
Let Bob know very clearly how pleased you have been with his performance and how much you'd like him to remain on your team. Then ask him if there are any problems you might be able to help with.
Tell Bob you would like him to hold off on the transfer for 1 month to see if together you can solve whatever problems are behind his request for transfer.

23) As a manager, you have noticed that one of your employees, Joan, has been taking a lot of her own time to help out another employee. What should you do?

Commend Joan and recommend her for special recognition.
Tell Joan that you appreciate her help, but also mention that you expect her not to neglect her own duties.
Talk with Joan to find out if the other employee has some special problem that you should know about. Let Joan know you appreciate her help.
Tell Joan that you appreciate her help, but that the other employee needs to learn to do his/her own work.

24) One of your employees is performing an assigned task exactly the way it is supposed to be done, but he is taking entirely too much time to get it done. Your work team has several more tasks to do today, and you told the employee to work faster so that the entire group will not have to stay late. An hour later, when you came back to see what he had accomplished, you found that he had not don anything since the last time you talked with him. Because of this, you and your whole team will have to work late tonight to meet a deadline. What should you do?

85

Talk to the employee and find out if anything is bothering him.
Assign another employee to assist him until the job is completed. Meet with him later to find out why he was so slow.
Put this employee on a different task and have one of your better employees finish up this task.
Tell the whole team that everyone is staying late because of one employee. Hopefully, this will create peer pressure and speed up his performance.

25) You are a manager. Yesterday, your work team finished up a difficult and exhausting 2-week project. This morning, your department director overlooked several other teams and assigned an important "rush" project to your team. Your employees have given you a lot of negative feedback about this assignment. What do you do?

Advise your director that your employees have just finished a difficult project and you do not feel they are capable of giving their best work on a new "rush" project at this time. Persuade your director to select another team for the new project.
Determine why your team was selected (because of superior performance or just necessity), then inform your team of the reason for their selection and the importance of going forward with the new project.
Ask your director to consider giving your team a different project since they just finished a difficult project and you do not feel they are capable of giving their best work on this new "rush" project. If your director won't reconsider, take the matter to your director's supervisor.

26) You have a new director who has been the head of your department for about 3 months. He just told you that he has recommended one of your employees for a promotion. You told him that you have worked with this employee for over a year and that you don't think the employee is ready for a promotion. Your director says he has already made up his mind, but you are sure that the employee is not ready. What should you do?

Present your new director with all the documentation you have to support your views of the employee in question.
Ask other managers in your department what they have done in similar circumstances.
Present your documentation in an informational email to your director's supervisor.

Appendix C: Secondary Inventory (SPE-30) and Instructions

Thank you for taking the time to complete the SJT-280. The next part of this research will ask you some questions about the instructions you read, the situational judgment test (SJT-280) you just completed, and some additional questions about you and your personality.

At this point, you DO NOT need to imagine you are trying to get a job. Answer the following questionnaire about *yourself* and *your true opinions* as honestly as you can.

Choose the response that best describes how much you agree or disagree with each statement about how you responded to the SJT.

[The following anchors are presented above each set of items:
1 = I Very Much Disagree
2 = I Somewhat Disagree
3 = I am Neutral
4 = I Somewhat Agree
5 = I Very Much Agree]

When answering this inventory, I found myself focusing on situations I have been in before, even though they did not happen at work.	1	2	3	4	5
When answering this inventory, I chose the answers that I did because I have a pretty good idea of what will happen as a result of each choice.	1	2	3	4	5
When answering this inventory, I chose the answers that I did because they seemed like the "right thing to do."	1	2	3	4	5
When answering this inventory, I did not answer the way I normally would, because I was asked to answer as though I really wanted the job.	1	2	3	4	5

Choose the response that best describes how much you agree or disagree with each statement about *your everyday life*.

I base most of my decisions in everyday life on how I have made decisions in the past.	1	2	3	4	5
I think I have a good idea of what a manager's job involves.	1	2	3	4	5
I think I am the kind of person who would enjoy a full-time job in management.	1	2	3	4	5
Most of my work experience has been with a company that is very similar to the company in the description that I read.	1	2	3	4	5
I expect that the values at most companies are similar to the values at the company in the description that I read.	1	2	3	4	5
The description of the imaginary company provided some clues about how I should answer the situational judgment items in order to get the job.	1	2	3	4	5
The email from the imaginary coworker provided some clues about how I should answer the situational judgment items in order to get the job.	1	2	3	4	5
The company that I work for now is very similar to the company in the description that I read.	1	2	3	4	5
In my experience, most managers would fit right in at the company in the description that I read.	1	2	3	4	5
Most of my experience as a manager has been in a company very similar to the description that I read. (if you have no management experience, please answer this question as an employee).	1	2	3	4	5

Choose the response that best describes how much you agree or disagree with each statement about *your everyday life*.

I feel little concern for others.	1	2	3	4	5
I am interested in people.	1	2	3	4	5
I never insult people.	1	2	3	4	5
I sympathize with others' feelings.	1	2	3	4	5
I have a soft heart.	1	2	3	4	5
I am not really interested in others.	1	2	3	4	5
I take time out for others.	1	2	3	4	5
I feel others' emotions.	1	2	3	4	5
I make people feel at ease.	1	2	3	4	5
I am not interested in other people's problems.	1	2	3	4	5
I tend to trust other people, for the most part.	1	2	3	4	5
I will listen to others if I believe they can make a contribution.	1	2	3	4	5
I am the kind of person whom others can trust.	1	2	3	4	5
I seldom have all the answers.	1	2	3	4	5

Choose the answer that best describes your work experience:

I have worked part-time or full-time in a management job (with any company) for:

0 years (never)
less than 1 year total
1 to 3 years total
3 to 5 years total
more than 5 years total

Choose the answer that best describes your experience with your current company:

I have been with my current company (in any job) for:

0 years (I am not currently employed)
less than 1 year
1 to 3 years
3 to 5 years
more than 5 years

If you are currently employed, please select the industry that best describes your current job, or select "Other" and describe your current job in the box provided.

If you are not currently employed, please choose NOT EMPLOYED.

NOT EMPLOYED
Communications
Energy
Finance
Government
Healthcare
Manufacturing
Retail
Social Services
Science/Technology
Transportation
Other (Please specify Below):

Appendix D: Manipulation Pilot Inventory

For each statement below, indicate whether you
1 = Strongly Disagree
2 = Moderately Disagree
3 = Neither Disagree nor Agree
4 = Moderately Agree
5 = Strongly Agree

	SD	MD	N	MA	SA
After reading the description of this company, it sounds like a place where employees are told what is expected of them.	1	2	3	4	5
After reading the description of this company, it sounds like a place where employees are part of the decision making process.	1	2	3	4	5
After reading the description of this company, it sounds like a place where employees are told how to perform their jobs.	1	2	3	4	5
After reading the description of this company, it sounds like a place where employees are asked for suggestions.	1	2	3	4	5
After reading the description of this company, it sounds like a place where leadership is friendly and approachable.	1	2	3	4	5
After reading the description of this company, it sounds like a place where leaders show concern for their employees' well-being.	1	2	3	4	5
After reading the description of this company, it sounds like a place where employees are expected to perform at their highest level.	1	2	3	4	5
After reading the description of this company, it sounds like a place where leaders set challenging goals for their employees.	1	2	3	4	5
After reading the description of this company, it sounds like the kind of place I would like to work.	1	2	3	4	5

Appendix E: Situational Judgment Inventory Pilot Rating Instructions

Used to Create Conditional Response Keys and Measure Reliability

Imagine that you work as a manager at a company that is creating a test to help select new managers. Your company has asked you to help create a key for this new test. A *key* is a set of correct or best responses to a test against which a candidate's responses can be compared.

First, some information will be presented to help you understand how things work at your company. This should help you understand how to create a key that will be specific to your company.

Then, it will be up to you to read each question on the test and **rate** the response choices for each question using the following scale:

> **A Very Effective Response**
> **A Somewhat Effective Response**
> **A Somewhat Ineffective Response**
> **A Very Ineffective Response**

You will find that the questions on the test have 3, 4, or 5 response choices for you to rate. Therefore, you might not use every rating for each question, or you might use some ratings more than once for each question.

Please try to rate every response choice for every question. Try to find at least one effective and one ineffective response for each question. Finally, do not over-think any response choice or question; just go with your first impression using your best judgment. Thank you for your participation!

Appendix F: Significant SJT Item-level Chi-Square and Phi Values

	Most Effective			**Item**		**Least Effective**		
^{2}	p	*Phi*	R^2		^{2}	p	*Phi*	R^2
39.17	0.000	0.315	0.10	**1**	28.5	0.000	0.268	0.07
32.63	0.000	0.387	0.15	**2**				
				3	11.44	0.022	0.17	0.03
				4				
				5				
				6				
				7	16.44	0.036	0.204	0.04
				8				
28.7	0.000	0.269	0.07	**9**	13.67	0.034	0.186	0.03
				10	15.55	0.016	0.198	0.04
				11				
				12				
				13				
22.49	0.004	0.238	0.06	**14**				
16.75	0.033	0.206	0.04	**15**				
18.78	0.016	0.218	0.05	**16**				
35.27	0.000	0.298	0.09	**17**	16.45	0.036	0.204	0.04
22.51	0.001	0.238	0.06	**18**	28.15	0.000	0.267	0.07
				19	15.42	0.017	0.197	0.04
27.91	0.000	0.265	0.07	**20**	15.44	0.017	0.197	0.04
14.43	0.025	0.191	0.04	**21**				
				22				
				23				
17.55	0.007	0.211	0.04	**24**	18.9	0.004	0.218	0.05
10.47	0.033	0.163	0.03	**25**				
				26				
	12			**Number Significant**		**10**		
	46.15%			**Percent Significant**		**38.46%**		

92

Appendix G: Item-Level Frequency Tables for Managers and Students

MANAGERS (N = 258)

Response #	Most Effective			Least Effective		
	Participative	**Directive**	**Control**	**Participative**	**Directive**	**Control**
	N=74	*N=88*	*N=96*	*N=74*	*N=88*	*N=96*
1a	25.7%	9.1%	14.6%	6.8%	17.0%	10.4%
b	0.0%	0.0%	0.0%	43.2%	52.3%	49.0%
c	66.2%	61.4%	72.9%	1.4%	0.0%	1.0%
d	4.1%	4.5%	1.0%	18.9%	14.8%	21.9%
e	4.1%	25.0%	11.5%	29.7%	15.9%	17.7%
2a	18.9%	17.0%	10.4%	6.8%	12.5%	8.3%
b	5.4%	0.0%	1.0%	41.9%	38.6%	53.1%
c	74.3%	73.9%	82.3%	4.1%	2.3%	0.0%
d	0.0%	0.0%	0.0%	35.1%	33.0%	31.3%
e	1.4%	9.1%	6.3%	12.2%	13.6%	7.3%
3a	77.0%	90.9%	92.7%	14.9%	4.5%	2.1%
b	9.5%	2.3%	5.2%	14.9%	34.1%	21.9%
c	13.5%	6.8%	2.1%	70.3%	61.4%	76.0%
4a	10.8%	13.6%	11.5%	10.8%	13.6%	7.3%
b	0.0%	1.1%	0.0%	64.9%	63.6%	62.5%
c	50.0%	40.9%	35.4%	4.1%	1.1%	2.1%
d	1.4%	2.3%	1.0%	20.3%	20.5%	28.1%
e	37.8%	42.0%	52.1%	0.0%	1.1%	0.0%
5a	9.5%	14.8%	4.2%	32.4%	33.0%	37.5%
b	0.0%	6.8%	7.3%	64.9%	60.2%	59.4%
c	90.5%	78.4%	88.5%	2.7%	6.8%	3.1%
6a	5.4%	4.5%	2.1%	29.7%	19.3%	22.9%
b	14.9%	11.4%	7.3%	14.9%	19.3%	20.8%
c	44.6%	34.1%	34.4%	9.5%	11.4%	10.4%
d	17.6%	22.7%	19.8%	28.4%	27.3%	28.1%
e	17.6%	27.3%	36.5%	17.6%	22.7%	17.7%
7a	29.7%	18.2%	18.8%	2.7%	2.3%	7.3%
b	50.0%	73.9%	72.9%	4.1%	2.3%	4.2%
c	1.4%	2.3%	0.0%	21.6%	18.2%	11.5%
d	0.0%	0.0%	0.0%	56.8%	62.5%	68.8%
e	20.3%	5.7%	8.3%	14.9%	14.8%	8.3%
8a	9.5%	17.0%	9.4%	43.2%	33.0%	39.6%
b	2.7%	4.5%	7.3%	33.8%	36.4%	39.6%
c	62.2%	44.3%	47.9%	6.8%	6.8%	7.3%
d	25.7%	34.1%	35.4%	16.2%	23.9%	13.5%
9a	1.4%	9.1%	0.0%	40.5%	44.3%	45.8%
b	27.0%	48.9%	38.5%	16.2%	19.3%	10.4%
c	14.9%	18.2%	19.8%	32.4%	14.8%	31.3%
d	56.8%	23.9%	41.7%	10.8%	21.6%	12.5%

93

Appendix G: Item-Level Frequency Tables (Continued)

10a	86.5%	79.5%	89.6%	5.4%	2.3%	0.0%
b	10.8%	18.2%	9.4%	12.2%	2.3%	4.2%
c	2.7%	2.3%	1.0%	12.2%	17.0%	9.4%
d	0.0%	0.0%	0.0%	70.3%	78.4%	86.5%
11a	28.4%	38.6%	31.3%	29.7%	29.5%	38.5%
b	55.4%	45.5%	57.3%	13.5%	12.5%	9.4%
c	16.2%	15.9%	11.5%	56.8%	58.0%	52.1%
12a	71.6%	61.4%	60.4%	1.4%	3.4%	1.0%
b	25.7%	33.0%	36.5%	4.1%	2.3%	2.1%
c	0.0%	0.0%	0.0%	58.1%	42.0%	47.9%
d	0.0%	4.5%	2.1%	10.8%	23.9%	15.6%
e	2.7%	1.1%	1.0%	25.7%	28.4%	33.3%
13a	1.4%	4.5%	3.1%	41.9%	40.9%	53.1%
b	79.7%	71.6%	80.2%	4.1%	4.5%	0.0%
c	5.4%	11.4%	9.4%	37.8%	38.6%	31.3%
d	13.5%	12.5%	7.3%	16.2%	15.9%	15.6%
14a	2.7%	1.1%	0.0%	10.8%	6.8%	3.1%
b	14.9%	21.6%	26.0%	6.8%	3.4%	1.0%
c	0.0%	4.5%	0.0%	81.1%	84.1%	87.5%
d	40.5%	20.5%	24.0%	1.4%	4.5%	6.3%
e	41.9%	52.3%	50.0%	0.0%	1.1%	2.1%
15a	10.8%	12.5%	9.4%	9.5%	6.8%	3.1%
b	2.7%	10.2%	10.4%	6.8%	2.3%	0.0%
c	33.8%	50.0%	41.7%	1.4%	1.1%	3.1%
d	52.7%	26.1%	38.5%	1.4%	4.5%	0.0%
e	0.0%	1.1%	0.0%	81.1%	85.2%	93.8%
16a	20.3%	5.7%	11.5%	9.5%	11.4%	7.3%
b	51.4%	70.5%	69.8%	0.0%	1.1%	0.0%
c	4.1%	8.0%	1.0%	24.3%	21.6%	16.7%
d	5.4%	8.0%	5.2%	41.9%	35.2%	54.2%
e	18.9%	8.0%	12.5%	24.3%	30.7%	21.9%
17a	6.8%	10.2%	5.2%	23.0%	18.2%	18.8%
b	5.4%	12.5%	1.0%	6.8%	12.5%	21.9%
c	4.1%	8.0%	3.1%	39.2%	35.2%	34.4%
d	8.1%	5.7%	10.4%	28.4%	21.6%	21.9%
e	75.7%	63.6%	80.2%	2.7%	12.5%	3.1%
18a	21.6%	38.6%	27.1%	32.4%	12.5%	16.7%
b	25.7%	13.6%	11.5%	23.0%	56.8%	41.7%
c	27.0%	9.1%	20.8%	21.6%	19.3%	24.0%
d	27.0%	38.6%	40.6%	23.0%	11.4%	17.7%

94

Appendix G: Item-Level Frequency Tables (Continued)

19a	5.4%	3.4%	2.1%	63.5%	76.1%	79.2%
b	6.8%	9.1%	6.3%	27.0%	9.1%	10.4%
c	81.1%	73.9%	80.2%	2.7%	5.7%	2.1%
d	6.8%	13.6%	11.5%	6.8%	9.1%	8.3%
20a	74.3%	48.9%	55.2%	2.7%	11.4%	4.2%
b	1.4%	0.0%	0.0%	43.2%	44.3%	54.2%
c	12.2%	27.3%	26.0%	12.2%	9.1%	7.3%
d	12.2%	23.9%	18.8%	41.9%	35.2%	34.4%
21a	21.6%	13.6%	13.5%	5.4%	9.1%	6.3%
b	1.4%	8.0%	1.0%	86.5%	78.4%	86.5%
c	31.1%	12.5%	21.9%	6.8%	10.2%	6.3%
d	45.9%	65.9%	63.5%	1.4%	2.3%	1.0%
22a	20.3%	18.2%	24.0%	13.5%	13.6%	8.3%
b	1.4%	2.3%	1.0%	50.0%	48.9%	54.2%
c	77.0%	76.1%	72.9%	0.0%	2.3%	0.0%
d	1.4%	3.4%	2.1%	36.5%	35.2%	37.5%
23a	14.9%	8.0%	9.4%	4.1%	12.5%	10.4%
b	6.8%	20.5%	16.7%	36.5%	25.0%	25.0%
c	74.3%	62.5%	68.8%	8.1%	11.4%	6.3%
d	4.1%	9.1%	5.2%	51.4%	51.1%	58.3%
24a	33.8%	28.4%	28.1%	1.4%	5.7%	0.0%
b	60.8%	62.5%	70.8%	1.4%	1.1%	0.0%
c	4.1%	3.4%	1.0%	4.1%	3.4%	3.1%
d	1.4%	5.7%	0.0%	93.2%	89.8%	96.9%
25a	24.3%	12.5%	9.4%	9.5%	22.7%	16.7%
b	74.3%	87.5%	89.6%	8.1%	2.3%	1.0%
c	1.4%	0.0%	1.0%	82.4%	75.0%	82.3%
26a	97.3%	94.3%	95.8%	0.0%	3.4%	1.0%
b	1.4%	2.3%	3.1%	27.0%	28.4%	21.9%
c	1.4%	3.4%	1.0%	73.0%	68.2%	77.1%

[ ] = Summed Six Keyed Responses

STUDENTS (N = 138)

	Most Effective			**Least Effective**		
Response #	**Participative**	**Directive**	**Control**	**Participative**	**Directive**	**Control**
	N=40	*N=53*	*N=45*	*N=40*	*N=53*	*N=45*
1a	15.0%	9.4%	25.0%	15.0%	43.4%	18.2%
b	0.0%	1.9%	0.0%	40.0%	43.4%	38.6%
c	65.0%	37.7%	47.7%	2.5%	1.9%	0.0%
d	10.0%	3.8%	6.8%	10.0%	3.8%	15.9%
e	10.0%	47.2%	22.7%	32.5%	7.5%	29.5%

95

Appendix G: Item-Level Frequency Tables (Continued)

2a	17.5%	26.4%	13.6%	15.0%	17.0%	25.0%
b	20.0%	1.9%	6.8%	35.0%	39.6%	36.4%
c	52.5%	50.9%	59.1%	7.5%	0.0%	2.3%
d	5.0%	1.9%	13.6%	30.0%	20.8%	25.0%
e	5.0%	18.9%	9.1%	12.5%	22.6%	13.6%
3a	52.5%	54.7%	52.3%	15.0%	11.3%	18.2%
b	30.0%	30.2%	29.5%	12.5%	17.0%	18.2%
c	17.5%	15.1%	20.5%	72.5%	71.7%	65.9%
4a	27.5%	20.8%	18.2%	12.5%	9.4%	9.1%
b	0.0%	1.9%	2.3%	72.5%	64.2%	56.8%
c	32.5%	34.0%	22.7%	0.0%	5.7%	18.2%
d	10.0%	17.0%	13.6%	10.0%	17.0%	13.6%
e	30.0%	26.4%	45.5%	5.0%	3.8%	4.5%
5a	7.5%	9.4%	20.5%	62.5%	64.2%	50.0%
b	32.5%	35.8%	36.4%	27.5%	22.6%	36.4%
c	60.0%	54.7%	45.5%	10.0%	13.2%	15.9%
6a	10.0%	1.9%	4.5%	35.0%	28.3%	22.7%
b	7.5%	5.7%	15.9%	20.0%	17.0%	25.0%
c	42.5%	45.3%	43.2%	2.5%	3.8%	6.8%
d	20.0%	32.1%	27.3%	12.5%	18.9%	18.2%
e	20.0%	15.1%	11.4%	30.0%	32.1%	29.5%
7a	12.5%	9.4%	18.2%	10.0%	9.4%	18.2%
b	42.5%	50.9%	34.1%	15.0%	15.1%	13.6%
c	5.0%	1.9%	6.8%	37.5%	26.4%	15.9%
d	2.5%	5.7%	2.3%	30.0%	37.7%	50.0%
e	37.5%	32.1%	40.9%	7.5%	11.3%	4.5%
8a	7.5%	0.0%	11.4%	30.0%	41.5%	38.6%
b	7.5%	13.2%	2.3%	32.5%	34.0%	31.8%
c	57.5%	49.1%	56.8%	7.5%	9.4%	9.1%
d	27.5%	37.7%	31.8%	30.0%	15.1%	22.7%
9a	7.5%	3.8%	9.1%	42.5%	35.8%	52.3%
b	32.5%	54.7%	34.1%	12.5%	5.7%	13.6%
c	10.0%	18.9%	13.6%	32.5%	34.0%	31.8%
d	50.0%	22.6%	45.5%	12.5%	26.4%	4.5%
10a	87.5%	69.8%	75.0%	0.0%	0.0%	0.0%
b	12.5%	28.3%	20.5%	7.5%	1.9%	6.8%
c	0.0%	1.9%	6.8%	10.0%	18.9%	22.7%
d	0.0%	0.0%	0.0%	82.5%	79.2%	72.7%
11a	25.0%	30.2%	34.1%	47.5%	32.1%	38.6%
b	45.0%	54.7%	56.8%	12.5%	15.1%	4.5%
c	30.0%	15.1%	11.4%	40.0%	52.8%	59.1%

96

12a	67.5%	69.8%	70.5%	2.5%	1.9%	4.5%
b	17.5%	18.9%	18.2%	10.0%	5.7%	9.1%
c	7.5%	1.9%	2.3%	40.0%	34.0%	38.6%
d	2.5%	3.8%	4.5%	32.5%	22.6%	34.1%
e	5.0%	5.7%	6.8%	15.0%	35.8%	15.9%
13a	2.5%	5.7%	11.4%	47.5%	47.2%	59.1%
b	67.5%	52.8%	65.9%	5.0%	18.9%	6.8%
c	17.5%	24.5%	20.5%	32.5%	22.6%	20.5%
d	12.5%	17.0%	4.5%	15.0%	11.3%	15.9%
14a	2.5%	7.5%	4.5%	7.5%	3.8%	11.4%
b	20.0%	35.8%	29.5%	2.5%	1.9%	4.5%
c	0.0%	1.9%	2.3%	82.5%	83.0%	77.3%
d	42.5%	22.6%	22.7%	5.0%	3.8%	6.8%
e	35.0%	32.1%	43.2%	2.5%	7.5%	2.3%
15a	5.0%	7.5%	9.1%	10.0%	9.4%	15.9%
b	5.0%	26.4%	15.9%	2.5%	9.4%	2.3%
c	50.0%	32.1%	40.9%	0.0%	0.0%	4.5%
d	37.5%	34.0%	36.4%	0.0%	3.8%	6.8%
e	2.5%	0.0%	0.0%	87.5%	77.4%	72.7%
16a	30.0%	22.6%	27.3%	2.5%	5.7%	4.5%
b	40.0%	47.2%	43.2%	5.0%	1.9%	6.8%
c	7.5%	13.2%	15.9%	20.0%	18.9%	13.6%
d	7.5%	9.4%	2.3%	47.5%	37.7%	50.0%
e	17.5%	7.5%	13.6%	25.0%	35.8%	27.3%
17a	2.5%	15.1%	9.1%	22.5%	20.8%	20.5%
b	2.5%	15.1%	2.3%	12.5%	9.4%	6.8%
c	17.5%	32.1%	22.7%	27.5%	24.5%	31.8%
d	0.0%	5.7%	6.8%	32.5%	26.4%	29.5%
e	77.5%	32.1%	61.4%	5.0%	18.9%	13.6%
18a	15.0%	34.0%	36.4%	12.5%	11.3%	18.2%
b	22.5%	15.1%	27.3%	27.5%	50.9%	27.3%
c	55.0%	32.1%	31.8%	5.0%	5.7%	11.4%
d	7.5%	18.9%	6.8%	55.0%	32.1%	45.5%
19a	12.5%	7.5%	4.5%	62.5%	56.6%	61.4%
b	20.0%	20.8%	27.3%	20.0%	11.3%	25.0%
c	45.0%	52.8%	56.8%	10.0%	18.9%	9.1%
d	22.5%	18.9%	13.6%	7.5%	13.2%	6.8%
20a	50.0%	43.4%	27.3%	2.5%	17.0%	11.4%
b	22.5%	7.5%	9.1%	20.0%	18.9%	25.0%
c	25.0%	28.3%	56.8%	7.5%	13.2%	6.8%
d	2.5%	20.8%	9.1%	70.0%	50.9%	59.1%

97

Appendix G: Item-Level Frequency Tables (Continued)

21a	25.0%	26.4%	22.7%	2.5%	7.5%	9.1%
b	5.0%	7.5%	4.5%	87.5%	71.7%	72.7%
c	17.5%	17.0%	20.5%	7.5%	17.0%	11.4%
d	52.5%	49.1%	54.5%	2.5%	3.8%	9.1%
22a	7.5%	15.1%	9.1%	12.5%	20.8%	29.5%
b	22.5%	9.4%	9.1%	42.5%	37.7%	31.8%
c	57.5%	66.0%	79.5%	7.5%	5.7%	0.0%
d	12.5%	9.4%	4.5%	37.5%	35.8%	40.9%
23a	12.5%	1.9%	13.6%	45.0%	32.1%	22.7%
b	25.0%	34.0%	18.2%	10.0%	18.9%	20.5%
c	52.5%	54.7%	56.8%	7.5%	5.7%	6.8%
d	10.0%	9.4%	13.6%	37.5%	43.4%	52.3%
24a	27.5%	9.4%	22.7%	5.0%	20.8%	6.8%
b	57.5%	64.2%	68.2%	0.0%	5.7%	2.3%
c	10.0%	15.1%	11.4%	10.0%	7.5%	2.3%
d	5.0%	11.3%	0.0%	85.0%	66.0%	90.9%
25a	40.0%	43.4%	25.0%	17.5%	18.9%	38.6%
b	45.0%	45.3%	54.5%	27.5%	30.2%	31.8%
c	15.0%	11.3%	22.7%	55.0%	50.9%	31.8%
26a	72.5%	77.4%	77.3%	7.5%	3.8%	4.5%
b	17.5%	9.4%	13.6%	32.5%	49.1%	50.0%
c	10.0%	13.2%	11.4%	60.0%	47.2%	47.7%

= Summed Six Keyed Responses

98

Appendix H: Student SJT Item Level Chi-Square and Phi Values

Student Most Effective				**Item**	**Student Least Effective**			
χ^2	*p*	*Phi*	R^2		χ^2	*p*	*Phi*	R^2
21.98	**0.005**	0.399	0.159	**1**	21.815	**0.005**	0.398	0.158
20.686	**0.008**	0.387	0.150	**2**	8.501	0.386	0.248	0.062
0.414	0.981	0.055	0.003	**3**	1.392	0.846	0.1	0.010
6.448	0.597	0.216	0.047	**4**	11.14	0.194	0.284	0.081
4.375	0.358	0.178	0.032	**5**	3.055	0.549	0.149	0.022
8.166	0.417	0.243	0.059	**6**	3.494	0.9	0.159	0.025
5.849	0.664	0.206	0.042	**7**	9.029	0.34	0.256	0.066
10.293	0.113	0.273	0.075	**8**	3.328	0.767	0.155	0.024
11.448	0.075	0.288	0.083	**9**	10.036	0.123	0.27	0.073
7.296	0.121	0.23	0.053	**10**	4.064	0.397	0.172	0.030
5.665	0.226	0.203	0.041	**11**	5.491	0.241	0.199	0.040
2.72	0.951	0.14	0.020	**12**	8.577	0.379	0.249	0.062
7.359	0.289	0.231	0.053	**13**	7.635	0.266	0.235	0.055
8.765	0.362	0.252	0.064	**14**	5.023	0.755	0.191	0.036
11.143	0.194	0.284	0.081	**15**	11.612	0.169	0.29	0.084
6.916	0.546	0.224	0.050	**16**	4.443	0.815	0.179	0.032
24.851	**0.002**	0.429	0.184	**17**	4.871	0.771	0.188	0.035
13.416	**0.037**	0.312	0.097	**18**	10.184	0.117	0.272	0.074
3.774	0.707	0.165	0.027	**19**	6.145	0.407	0.211	0.045
21.862	**0.001**	0.398	0.158	**20**	7.535	0.274	0.234	0.055
0.875	0.99	0.08	0.006	**21**	6.162	0.405	0.211	0.045
8.268	0.219	0.245	0.060	**22**	6.659	0.354	0.22	0.048
7.492	0.278	0.233	0.054	**23**	5.918	0.432	0.207	0.043
10.655	0.1	0.278	0.077	**24**	12.874	**0.045**	0.305	0.093
4.903	0.297	0.188	0.035	**25**	7.971	0.093	0.24	0.058
1.429	0.839	0.102	0.010	**26**	3.378	0.497	0.156	0.024

5	Number Significant **(bold)**	**2**
19.2%	Percent Significant	**7.7%**

Appendix I: Manager SJT Item Level Chi-Square and Phi Values

	Mgr Most			Item		Mgr Least		
x^2	p	Phi	R^2		x^2	p	Phi	R^2
23.49	**0.001**	0.302	0.091	**1**	11.21	0.19	0.208	0.043
13.742	**0.033**	0.231	0.053	**2**	9.62	0.293	0.193	0.037
13.01	**0.011**	0.225	0.051	**3**	19.124	**0.001**	0.272	0.074
6.96	0.541	0.164	0.027	**4**	6.816	0.557	0.163	0.027
11.642	**0.02**	0.212	0.045	**5**	2.693	0.61	0.102	0.010
11.011	0.201	0.207	0.043	**6**	3.646	0.888	0.119	0.014
16.88	**0.01**	0.256	0.066	**7**	9.65	0.29	0.193	0.037
8.776	0.187	0.184	0.034	**8**	4.474	0.613	0.132	0.017
28.772	**0.0001**	0.334	0.112	**9**	13.017	**0.043**	0.225	0.051
4.335	0.363	0.13	0.017	**10**	16.264	**0.012**	0.251	0.063
3.669	0.453	0.119	0.014	**11**	2.475	0.649	0.098	0.010
6.997	0.321	0.165	0.027	**12**	9.437	0.307	0.191	0.036
5.317	0.504	0.144	0.021	**13**	6.867	0.333	0.167	0.028
20.374	**0.009**	0.281	0.079	**14**	11.816	0.16	0.214	0.046
16.024	**0.042**	0.249	0.062	**15**	16.942	**0.031**	0.256	0.066
19.641	**0.012**	0.276	0.076	**16**	9.324	0.316	0.19	0.036
16.843	**0.032**	0.256	0.066	**17**	17.357	**0.027**	0.259	0.067
18.774	**0.005**	0.27	0.073	**18**	23.598	**0.001**	0.302	0.091
4.023	0.674	0.125	0.016	**19**	14.377	**0.026**	0.236	0.056
15.132	**0.019**	0.242	0.059	**20**	9.05	0.171	0.187	0.035
19.49	**0.003**	0.275	0.076	**21**	2.893	0.822	0.106	0.011
2.109	0.909	0.09	0.008	**22**	5.636	0.465	0.148	0.022
9.958	0.126	0.196	0.038	**23**	7.741	0.258	0.173	0.030
9.759	0.135	0.194	0.038	**24**	8.36	0.213	0.18	0.032
9.172	0.057	0.189	0.036	**25**	11.186	**0.025**	0.208	0.043
2.116	0.714	0.091	0.008	**26**	4.619	0.329	0.134	0.018

| | **13** | | | **Number Significant** | | **8** | | |
| | **50.0%** | | | **Percent Significant** | | **30.8%** | | |

100

About the Author

Jonathan Adam Shoemaker has held an interest in general psychology since childhood. He completed his B.A. with Honors in Psychology at The College of William & Mary in 1995. He completed his M.S. in Applied Psychology at Georgia College & State University in 1997.   He has taught numerous courses as a graduate student and provided selection, training and development consulting services for various organizations. He continues to be interested in selection, training, organizational culture, and occupational health psychology.