THE IMPACT OF SITUATIONAL STRENGTH ON THE VALIDITY OF SITUATIONAL JUDGMENT ITEMS

By

Juliya Golubovich

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2013

ABSTRACT

THE IMPACT OF SITUATIONAL STRENGTH ON THE VALIDITY OF SITUATIONAL JUDGMENT ITEMS

By

Juliya Golubovich

Situational judgment tests (SJTs) require test takers to respond to hypothetical work-related scenarios. These tests can be used in college admissions to supplement traditional cognitive predictors of student success, predictors that tend to disadvantage minority applicants and may not predict students' performance in important areas like leadership, integrity, and social responsibility. However, what SJTs measure is often unclear, their criterion-related validities can vary widely, and few validity moderators have been identified. This study aimed to extend existing knowledge and the practical value of SJTs by considering whether certain situational characteristics of the scenarios comprising an SJT have implications for the test's validity. The situational strength (i.e., the extent to which a situation constrains individuals' behavior) of scenarios comprising an SJT was manipulated. Participants completed either the *weak* or *strong* version of the SJT, along with measures of biographical data, personality, self-rated college performance, and citizenship behaviors. Participants' high school GPA, ACT/SAT scores, and first-year college GPA were collected from the university. In partial support of the expectation that strong situations would constrain the expression of individual differences in behavior, the weak SJT was a better measure of agreeableness and openness than the strong SJT. However, relations of SJT scores with scores on other personality traits, biographical data measures, and criteria (first-year college GPA, self-rated college performance, and citizenship behaviors) did not vary as a function of situational strength. Implications of the findings, study limitations, and future research directions are discussed.

ACKNOWLEDGMENTS

A lot of people deserve thanks for helping to make this project possible. First of all, I would like to thank my advisor, Dr. Neal Schmitt, for his willingness to take on this role, as well as for his timely and helpful feedback and encouragement throughout the entire process. I would also like to thank my other committee members, Drs. Ann Marie Ryan, John Schaubroeck, and Frederick Morgeson for their critical feedback and helpful insights. I am also very grateful to the College Board for offering the financial support that made this study possible.

A special thanks to the sixteen graduate students who served as judges of my situational judgment items, my two research assistants, Abraham Kim and Alex Seeley, who helped me with data collection, and Leslie Baldwin, who could always be counted on for advice and support with any and all issues related to the psychology participant pool.

Finally, I want to offer heartfelt thanks to my mom, older brother, and fiancé for their continuous moral support, willingness to listen, and earnest attempts to understand what I was trying to do.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Practitioners invest considerable effort into designing organizational selection systems that reflect the goals of the organization and can effectively identify the applicants who are most likely to be successful in that organization. In doing so, they rely on various tools to assess individuals' abilities, knowledge, personality, and interests. Companies and higher education institutions are two types of organizations that heavily depend on selection systems. Given the similarity between the employment and academic contexts (e.g., schoolwork is the student's job; Munson & Rubenstein, 1992), hiring managers and admissions officers assess individuals on some of the same characteristics and using similar tools, and tend to struggle with similar challenges (e.g., maximizing selection system validity while maintaining fairness).

Selection systems in higher education in particular, aim to quickly and effectively screen a large number of applicants to fill available spots with individuals who are more likely to perform well and persevere until graduation (e.g., Buyse, 2011). College admissions officers typically evaluate high school students based on their academic achievement (i.e., GPA) and math and verbal standardized test scores (on the SAT or ACT) (Schmitt et al., 2009), as these measures are known to predict success in college (Hezlett et al., 2001; Kobrin et al., 2008; Kuncel, Hezlett, & Ones, 2004; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). These predictors may be referred to as more "cognitive" in nature to distinguish them from more "non-cognitive" predictors (e.g., personality, interpersonal skills) and measurement methods (e.g., biographical data, situational judgment tests, interviews, assessment centers). High school GPA and standardized test scores are not cognitive measures in the sense that they are assessments of the general factor of intelligence (psychometric conceptualization; Jensen, 1998); rather, they are

measures of accumulated knowledge and achievement, which intelligence facilitates, and can therefore be used as proxies for cognitive ability (Ployhart & Holtz, 2008; Schmitt et al., 2010).

Although these more cognitive measures predict success, there are drawbacks to relying on them in admissions/selection decisions: 1) They may not effectively predict students' performance in areas such as leadership, integrity, and social responsibility, in which schools are also interested (Schmitt et al., 2009; Shultz & Zedeck, 2011), 2) They may predict initial academic performance better than performance in later years (Buyse, 2011), and 3) Standardized test scores, in particular, show subgroup differences such that minority test takers tend to score 0.7 to 1.0 standard deviations lower than White, and may be subsequently disadvantaged when these scores are used to make admissions decisions (Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Sackett, Schmitt, Ellingson, & Kabin, 2001).

In light of the outlined issues, there has been increased interest in and push for using non-cognitive predictors to supplement the variables traditionally used in college and professional school admissions (Schmitt et al., 2009). To this end, several large-scale efforts have been undertaken both in the USA and abroad (see Buyse, 2011, Santelices, Ugarte, Flotts, Radovic, & Kyllonen, 2011, Schmitt et al., 2010, Shultz & Zedeck, 2011, and Sternberg, Bonney, Gabora, & Merrifield, 2012 for descriptions of this work). This work shows that non-cognitive measures can demonstrate incremental validity above cognitive measures for predicting performance (both academic and subsequently on-the-job; Lievens & Sackett, 2011), facilitate the expansion of the criterion space beyond grades to other criteria schools are interested in and which these non-cognitive measures can predict, and show reduced subgroup differences relative to cognitive measures. Because non-cognitive measures evidence reduced subgroup score differences, using them in conjunction with cognitive measures can, in some cases (see Ryan, Ployhart, & Friedel,

1998 for a discussion), result in a more demographically diverse student body (Sinha, Oswald, Imus, & Schmitt, 2011) with only a small decrement or even improvement in average student performance (Schmitt et al., 2009; Sternberg et al., 2012).

The current paper will focus on one non-cognitive predictor of college student potential, situational judgment test scores. Situational judgment tests have demonstrated potential for predicting academic and non-academic criteria both on their own and incrementally above high school grades and standardized test scores (Schmitt et al., 2009). In addition to predicting important criteria, situational judgment tests can enhance fairness and increase selection rates for minority applicants (Schmitt et al., 2009). These properties can make situational judgment tests a useful selection tool for academic institutions negotiating the dual concerns of having predictors that are *both* valid and evidence minimal adverse impact against minority applicants (diversity-validity dilemma; Ployhart & Holtz, 2008). However, the validity of situational judgment tests can vary widely (McDaniel, Hartman, Whetzel, & Grubb, 2007), and relatively little is known about the variables that moderate relations with criteria. As such, further research attention should be directed at identifying moderators of situational judgment tests' validity and enhancing these tests' utility as a selection tool in college admissions.

The current study considers whether the validity of situational judgment test scores for predicting students' performance (broadly defined) may be enhanced by manipulating the strength of the situations that comprise the items in the test. In the sections that follow, existing research on the multi-dimensionality of college student performance and situational judgment tests will be reviewed first. An overview of the research tradition of interactionism and attempts to derive taxonomies of situations will follow. The concept of situational strength as an overarching dimension characterizing situations will be introduced and its application for the

design of more valid situational judgment tests will be considered. Thereafter, the hypotheses and methodology for the current research will be reviewed. Finally, results will be presented and discussed in terms of their implications for theory and practice. Notably, although the current research focuses on the validity of situational judgment tests in *college admissions*, implications for selection more broadly (including selection in both higher education institutions and other types of organizations) will be considered.

College Student Success: Criteria

GPA is the most widely used measure of college performance (Robbins et al., 2004). First year college GPA in particular is typically used to examine the criterion-related validity of assessments used in admissions (Stemler, 2012). Another important consideration has been length of enrollment or retention. Attrition can be costly for academic institutions (Daugherty & Lane, 1999) and institutional eligibility for federal financial aid for students may be dependent on an institution's graduation rate (Titus, 2004).

Although academic institutions use easily quantifiable aspects of students' performance for purposes of selection and evaluation, they actually define student success much more broadly. Institutions' objectives include not only traditional academic criteria (e.g., grades, test scores, graduation) but also consideration of personal qualities like leadership and strength of character (Taber & Hackman, 1976). A great deal of research has aimed at advancing our understanding of college students' performance and success. But a couple of studies stand out in their attempt to identify and operationalize a set of critical dimensions of college student performance (e.g., Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004; Taber & Hackman, 1976). Taber and Hackman (1976) derived five academic and eight non-academic performance factors from ratings collected on a behaviorally-anchored rating instrument. Their academic dimensions

included cognitive proficiency, academic effort and achievement, communication proficiency, career goals, and artistic performance. The non-academic dimensions were optimistic and emotionally stable behavior, ethical behavior, interpersonal responsiveness, participation in organizations, athletic performance, congruence with the college, discrimination issues, and interpersonal sociability. Yet, given that Taber and Hackman's effort to identify dimensions of college student performance was limited to subject matter experts and sampled from one [private and highly selective] academic institution, the generalizability of their dimensions was a potential concern.

In a more recent and extensive effort, Oswald and colleagues (2004) derived a set of college performance dimensions by drawing on the mission statements and objectives of twenty three colleges and universities that varied in their characteristics (e.g., public versus private, large versus small). Their performance dimensions can be classified as intellectual behaviors, interpersonal behaviors, and intrapersonal behaviors. Intellectual behaviors include knowledge, continuous learning, and artistic appreciation. Interpersonal behaviors include multicultural appreciation, leadership, interpersonal skills, and citizenship. Intrapersonal behaviors include health, career orientation, adaptability, perseverance, and ethics. This multidimensional view of college student performance points to additional relevant criteria beyond GPA and retention.

Behaviorally anchored rating scale (BARS). Oswald et al. (2004) identified positive and negative behaviors for each of their performance dimensions and used them to create behaviorally anchored rating scales (BARS). Each of the twelve items (one for each performance dimension) included the dimension name and definition and two examples of college-related critical incidents. Each incident was accompanied by behavioral anchors representing three levels of performance on a seven-point scale. Individuals used this information as a basis for

rating their level on the dimension. Scores on the BARS represent students' performance in college in the intellectual, interpersonal, and intrapersonal domains.

Organizational citizenship behaviors (OCBs). The construct of organizational citizenship behaviors (OCBs) comes out of the employment setting. Although OCBs are discretionary behaviors and are not explicitly required of individuals, their performance is desirable as these behaviors promote the effective functioning of the organization. In the college context specifically, OCBs are behaviors that are not required for degree attainment, but which help the welfare of the institution (Organ, 1997). Examples would include helping other students academically and socially, taking part in community service efforts, and participating in clubs that seek to improve the school (Schmitt et al., 2007).

BARS and OCBs capture the intrapersonal and interpersonal dimensions of the performance domain that colleges claim to value. Thus, BARS and OCB measures could be used to supplement college GPA and retention when assessing students' success in a broader—both cognitive and non-cognitive—sense. The criteria used to evaluate success determine the nature of the predictors used to select those who have a greater likelihood of being successful. Accordingly, predictors of college students' success are discussed next.

Traditional Predictors of College Student Success: Uses and Shortcomings

Colleges and universities have traditionally used high school GPA, class rank, and standardized test scores (SAT, ACT) to select students for admission (Breland, Maxey, Gernand, Cumming, & Trapani, 2002). These more cognitive measures are highly valid predictors of college GPA; in combination, they explain about 25% of the variance in first year college GPA (Robbins et al., 2004). High school GPA and standardized test scores can also be valid predictors of retention or persistence in college (Robbins et al., 2004; Zwick, 2007a). Although they do

predict the outcomes colleges closely monitor, as predictors, high school grades and standardized scores have a number of shortcomings. Each of these will be reviewed in the sections that follow.

Predicting college student success (broadly defined). Although high school grades and standardized test scores are predictive of traditional academic criteria, these measures do not capture the full range of qualities that make students successful (Schmitt et al., 2009). Evaluating the utility of traditional predictors for identifying students who are likely to succeed in a broader sense requires a good understanding of the college performance domain.

Taber and Hackman (1976) and Oswald et al. (2004) both examined the relations of various predictor measures with performance on their dimensions of student success. Taber and Hackman found that SAT scores had relatively stronger relations with ratings on their academic dimensions than ratings on the non-academic dimensions. In fact, SAT scores were either not related to, or had negative relations with ratings on the non-academic dimensions. A similar pattern of relations were observed between ratings and the traditional performance criterion of college GPA. These results indicate that non-academic dimensions of college student performance may not be effectively predicted with SAT scores nor effectively reflected in college GPA.

Schmitt et al. (2009) discuss the results of research that used the performance dimensions reported on by Oswald et al. (2004). This longitudinal, multi-institutional effort, involved 2,771 students from ten academic institutions. Students provided self-ratings on the twelve performance dimensions (using the BARS measure discussed earlier) as well as data on a number of other cognitive and non-cognitive measures. Schmitt and colleagues' findings were similar to those of Taber and Hackman (1976). Traditional predictors of college success, which in this case included standardized test scores (SAT/ACT) and high school GPA, had stronger

relations with the more traditional academic outcomes (college GPA, graduation status) than with the less traditional outcomes (performance measured via BARS, OCBs). OCBs showed weak relations with the traditional academic measures. However, the relations of the BARS with college GPA and graduation status were somewhat higher than expected in light of Taber and Hackman's (1976) findings with regard to their less traditional, non-academic dimensions of performance. This may have been due in part to the fact that students' ratings on the 12 dimensions of performance (three of which were "intellectual" and akin to those Taber and Hackman referred to as academic) in Schmitt et al.'s study were combined into a single BARS measure.

Taken together, Taber and Hackman's (1976) and Schmitt et al.'s (2009) findings show that: 1) Measures traditionally used in college admissions (e.g., standardized test scores, high school GPA) are better suited for predicting student success as narrowly defined by academic performance and graduation status and 2) Typically measured performance outcomes (e.g., college GPA, graduation status) provide little information about the non-academic dimensions of success. Many have argued for better alignment of academic institutions' stated objectives and selection practices (e.g., Schmitt, 2012; Stemler, 2012). Taber and Hackman's and Schmitt et al.'s work suggests that this goal will require the two pronged approach of supplementing traditional cognitive predictors with those of a more non-cognitive nature, and evaluating their validity against theoretically relevant non-cognitive criteria. Performance measured via BARS and OCBs can be effectively applied for the latter purpose.

Predicting college student performance over time. It has been suggested that ability is more important for performance in the early stages of a new job, as those with higher ability are able to master new material more quickly (Murphy, 1989). In the meantime, expression of

individual differences in motivation is minimized because of the demands of novel and challenging tasks (Helmreich, Sawin, & Carsrud, 1986). Later on, motivational processes should become more important for performance. Consistent with this logic, cognitive ability tests have been shown to predict academic performance best in early semesters or years and to evidence declining validities thereafter (Buyse, 2011; Humphreys, 1968; Humphreys & Taber, 1973; Lin & Humphreys, 1977). Lievens, Ones, and Dilchert (2009) found that personality traits (extraversion, openness, and conscientiousness) became more valid predictors of medical students' GPAs over time as demands shifted focus from knowledge acquisition to practical application. Thus, academic institutions could more effectively predict students' long-term performance by supplementing traditional cognitive predictors with more non-cognitive ones.

Demonstrating fairness for test takers and enhancing diversity. A significant concern with the standardized tests that are typically used for selection in academic contexts (e.g., SAT/ACT, GRE) are large race-based mean differences in scores, as minority test takers tend to score approximately 0.7 to 1.0 standard deviations lower than White test takers on these tests (Roth, Bevier, Bobko, Switzer, & Tyler, 2001). When groups score differently on selection tests, adverse impact (i.e., differential hiring rates for groups) can result. Thus, lower test scores decrease the likelihood that individuals from minority groups will receive positive outcomes in high stakes decisions in educational and organizational settings (Sackett, Schmitt, Ellingson, & Kabin, 2001). Further, selection systems that disadvantage minority applicants can result in institutions having poor demographic diversity. Notably, student body diversity can be an important goal for colleges and universities as it can benefit current students' learning, engagement, and satisfaction among other things (Stemler, 2012 provides a review). Relative to cognitive tests, non-cognitive predictors evidence lower adverse impact against minority test

takers (Ployhart & Holtz, 2008). Research suggests that overall adverse impact can sometimes be reduced by supplementing cognitive ability measures with predictor methods that have lower adverse impact (the effectiveness of this strategy can depend on factors such as the number of predictors, the correlations between predictors, subgroup score differences associated with each predictor, and applicant pool characteristics; Ryan et al., 1998; Sackett & Ellingson, 1997; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997) and by assessing a broader range of constructs relevant for success (Ployhart & Holtz, 2008).

Summary. Although traditionally used cognitive measures are, in combination, valid predictors of college GPA and retention, they may not effectively predict all facets of student success or even academic performance over the entire college career. Minority applicants may also score substantially worse on the SAT and ACT, making these groups less likely recipients of favorable selection decisions. Research (e.g., Oswald et al., 2004) has been done to better explicate the domain of college student performance and to identify additional predictors of college student success that may be used to supplement traditional cognitive measures.

Importantly, some institutions may already request and take into consideration additional information about applicants' qualifications (e.g., achievements, interpersonal skills) (Oswald et al., 2004). However, gleaning this additional information from methods like personal statements, letters of recommendation, and interviews may be less practical for large universities with a larger volume of applications to review than for small schools (Zwick, 2007b). In fact, larger institutions tend to value standardized test scores more than smaller institutions do because of their utility for quickly screening a large number of candidates (Hawkins & Lautz, 2005, as cited by Zwick, 2007b). A standardized method of collecting and scoring additional non-cognitive information about applicants could be a welcomed addition to large institutions' selection

batteries (Oswald et al., 2004). A method that meets this need as well as helps to address the discussed shortcomings of traditionally used cognitive measures is discussed next.

Overview of Situational Judgment Tests

One of the measures (or more appropriately, measurement methods) that has shown promise as a supplement to traditional predictors of student potential is the situational judgment test (SJT). SJTs present test takers with hypothetical situations (i.e., item stems) and ask them to indicate what they would (or should) do in each of these situations (i.e., the response options). The situations on an SJT may be representative of typical situations that occur on the job that the test was designed to reflect, or situations that are designed to be psychologically analogous to actual job situations (Chan & Schmitt, 1997). There is value in asking individuals to indicate their behavioral intentions in response to hypothetical situations because intentions are known to predict subsequent behavior (Latham, Saari, Pursell, & Campion, 1980; see Armitage & Conner, 2001 and Sheppard, Hartwick, & Warshaw, 1988 for meta-analytic findings).

Whereas SJTs have a long history of use in the area of employee selection (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001), attempts to apply them in academic contexts are fairly recent. Given the similarity between the academic and employment contexts (Munson & Rubenstein, 1992), evidence for the practical benefits of SJTs will be drawn from research in both domains. Evidence of SJTs' versatility as a measurement method, criterion-related validity, and incremental validity will be presented next. Evidence pertaining to fairness, applicant reactions, fakability, and susceptibility to coaching interventions will be reviewed as well.

Construct-related Validity of SJTs

Although some have argued that SJTs measure some sort of unique construct like judgment or practical intelligence (e.g., Brooks & Highhouse, 2006; Schmitt & Chan, 2006; Stemler & Sternberg, 2006), these tests are generally recognized to be multidimensional "measurement methods that may be used to assess a variety of constructs" (McDaniel et al., 2001). SJTs are typically designed to represent aspects of a particular work or academic domain, and are likely to tap several different constructs that are relevant for performance in that domain (Christian et al., 2010). These constructs can include cognitive ability (further discussed below), personality, job knowledge, and job experience (Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001; McDaniel et al., 2001; McDaniel & Nguyen, 2001; Weekley & Jones, 1997, 1999). However, there may be limits to the range of constructs that can be assessed with an SJT (Schmitt & Chan, 2006) and certain constructs tend to be assessed with greater frequency than others. A literature review conducted by Christian and colleagues (2010) focused on the higher-order construct domains measured by SJTs and showed that these tests most typically assess leadership and interpersonal skills (Christian et al., 2010).

As mentioned above, SJTs can measure cognitive ability; McDaniel et al.'s (2001) meta-analysis estimated the SJT-cognitive ability correlation at .46. However, the size of the correlation in any specific case depends a great deal on the way the SJT is designed (e.g., response instructions, response format, delivery medium) (Chan & Schmitt, 1997; Glaze, Jarrett, Schurig, Arthur, & Taylor, 2011; McDaniel et al., 2007) and the context it is used in (e.g., with applicants versus with incumbents; MacKenzie, Ployhart, Weekley, & Ehlers, 2010). The correlations of SJT scores with cognitive ability are lower when respondents are asked to indicate how they "would" respond (as opposed to how they "should" respond), respondents

provide ratings for an item's response options (as opposed to indicating their most and least likely response), the test is video-based (as opposed to paper-and-pencil-based), and the test is used with applicants (as opposed to incumbents). As further discussed below (see section on subgroup differences), these considerations also have implications for the magnitude of subgroup score differences in SJT scores.

Because of their versatility as a measurement method, SJTs can be designed to assess a variety of knowledge, skills, abilities, and other characteristics students require for success in an academic context (though not necessarily to the exclusion of unintended constructs; McDaniel et al., 2006). Oswald and colleagues (2004) designed an SJT to tap their twelve dimensions of college performance, including knowledge, continuous learning, artistic appreciation, multicultural appreciation, leadership, interpersonal skills, citizenship, health, career orientation, adaptability, perseverance, and ethics. The ability of their SJT (which will be adapted for use in the current research), and SJTs more broadly, to predict criteria relevant for student or employee success is considered next.

Criterion-related Validity of SJTs

Situational judgment tests can be valid predictors in both employment and academic contexts. SJTs have been used to predict a variety of important criteria in the employment context (where, again, SJTs have a longer history), including job performance (e.g., Weekley & Jones, 1997, 1999; Stevens & Campion, 1999), training performance (Krokos, 1999), and turnover (Dalessio, 1994).

In the academic context, SJTs have been used to predict students' academic (e.g., GPA) and non-academic performance (e.g., citizenship behavior, deviant behavior), absenteeism, graduation status, internship performance, and job performance after graduation (Buyse, 2011;

Lievens & Sackett, 2011; Oswald et al., 2004; Schmitt et al., 2009; Schmitt et al., 2010). Importantly, SJTs may effectively predict academic performance in later semesters of school, as the validity of cognitive predictors starts to shrink (Buyse, 2011). SJTs have also been used to predict attitudinal outcomes (e.g., intentions to quit, satisfaction) that may serve as precursors to various performance-related outcomes for students (Schmitt et al., 2007).

Meta-analyses have reported overall criterion-related validities for SJTs of ρ = .34 (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001) and ρ = .26 (McDaniel et al., 2007). The latter meta-analysis was more comprehensive, including 118 coefficients and a pooled sample size of 24,756 (as compared to the earlier meta-analysis's 102 coefficients and sample size of 10,640). SJTs measuring leadership skills and interpersonal skills (construct domains most frequently assessed), in particular, have average validities of .28 and .25, respectively (Christian et al., 2010). Notably, reported criterion-related validities for SJTs show considerable variability across studies, indicating the presence of moderating variables. For instance, SJTs that are based on a job analysis tend to show higher validities (McDaniel et al., 2001). There is also some evidence to suggest that "would do" or behavioral tendency (as opposed to "should do" or knowledge-based) response instructions and a video-based (as opposed to a paper-and-pencil) format may result in SJTs with higher criterion-related validity (with SJT content held constant) (Lievens & Sackett, 2006; Ployhart & Ehrhart, 2003). It is interesting to note that these findings are apparently at odds with those reviewed earlier showing that behavioral tendency instructions and video-based delivery minimize SJT-cognitive ability correlations. Given that cognitive ability tends to be the best predictor of performance (Schmidt & Hunter, 1998), it would stand to reason that the stated design elements would result in lower, not higher, criterion-related validity. Ployhart and Ehrhart (2003) suggest that the criterion under

consideration could determine the observed relations. More specifically, the match between the predictor and criterion can influence validity such that well-matched predictor constructs and criterion facets could result in the SJT having higher validity (Christian et al., 2010). The issue of validity moderators is central to the current study and will be considered further (see section on additional moderators of SJTs' criterion-related validity) after a brief review of other benefits of SJTs.

Incremental Validity of SJTs

SJTs have been shown to add incrementally to the prediction of various criteria. Reported analyses vary both with regard to the criteria that are predicted and the set of predictors examined. McDaniel et al.'s (2007) meta-analysis estimates that an SJT can have incremental validity over a measure of cognitive ability in the 3% to 5% (percent of variance accounted for) range, over a measure of personality in the 6% to 7% range, and over a combination of cognitive ability and personality in the 1% to 2% range.

Primary studies in the academic context find evidence of SJTs' incremental validity over cognitive ability and/or personality for predicting a range of criteria. Lievens and colleagues (Lievens, Buyse, & Sackett, 2005; Lievens & Sackett, 2011) demonstrated incremental validity for an interpersonal skills SJT. More variability in students' grades in interpersonally oriented classes, internship performance, and later job performance was explained by a combination of cognitive factors (including knowledge test, reading comprehension, and cognitive ability test scores) and SJT scores than by the cognitive factors alone. In another study, Mumford, Van Iddekinge, Morgeson, and Campion (2008) found a team role SJT to have incremental validity over a combination of cognitive ability and personality for predicting students' performance in academic project teams.

Primary studies in the organizational context offer additional evidence of SJTs' incremental validity. These tests have been found to predict job performance above and beyond: 1) Cognitive ability and experience combined (Weekley & Jones, 1997, 1999) and 2) Cognitive ability, job experience or job knowledge, and personality combined (Chan & Schmitt, 2002; Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001). In academic contexts, analyses of incremental validity provided by SJT scores often incorporate past performance (i.e., high school GPA) as a predictor, and consider a broader range of criteria. Schmitt and colleagues (Schmitt et al., 2009; Schmitt et al., 2010) show incremental validity for an SJT above combinations of standardized test scores (SAT/ACT), high school GPA, and biographical data (measures of relevant experiences). Their criteria in these analyses included college GPA, absenteeism, turnover intentions, deviant behavior, and drinking problems. In related work from Oswald et al. (2004), situational judgment test scores explained additional variability in college GPA, ratings on twelve academic and non-academic dimensions, and absenteeism above and beyond a composite of standardized test scores, personality, and biographical data.

Subgroup Differences on SJTs

A major benefit of SJTs is that they evidence lower subgroup score differences than do tests of cognitive ability. Meta-analytic findings indicate White-Black, White-Hispanic, White-Asian, and male-female SJT score differences of .38, .24, .29, and -.11 standard deviations, respectively (Whetzel, McDaniel, & Nguyen, 2008). Thus, although minority test takers may score lower than Whites on SJTs, these differences are of considerably smaller magnitude than those seen with standardized tests like the SAT/ACT. Women outperform men on SJTs but the average difference is small.

The magnitude of mean race differences that might be expected on SJTs depends a lot on how the test is designed. A multimedia (as opposed to paper-and-pencil) format, behavioral tendency (as opposed to knowledge) response instructions, and a response format that requires ratings of all the response options (as opposed to choosing a best and worst or most likely and least likely response) are SJT design elements that minimize score differences between racial subgroups (Chan & Schmitt, 1997; Glaze et al., 2011; McDaniel et al., 2007; Whetzel et al., 2008). These are elements that reduce SJT-cognitive ability correlations. Gender-based score differences are not affected by the cognitive loading of SJT scores, but do increase (in favor of women) to the extent that the test measures the personality constructs of agreeableness and conscientiousness (Whetzel et al., 2008).

Oswald and colleagues' (2004) SJT favored females (d = .70). But more importantly for adverse impact and diversity concerns (since traditional cognitive predictors disadvantage racial minorities), the test showed substantially reduced mean score differences between minority (Hispanic, Black) and White samples relative to differences in high school GPA and standardized test scores. On the SJT, mean differences between minority (Hispanic, Black) and White samples were four to six times smaller than on high school GPA and standardized test scores (Schmitt et al., 2009). The selection rates for minority subgroups (Blacks, Hispanics) could, under some circumstances (see Ryan et al., 1998), increase when more non-cognitive measures like an SJT are used to supplement cognitive measures in college admissions (Schmitt et al., 2009). The greater student body diversity may come at only small decrements (which might be expected because cognitive measures are more strongly related to college GPA than non-cognitive measures) or even improvement in average academic performance among the student body (Schmitt et al., 2009; Sternberg et al., 2012).

Applicant Reactions to SJTs

Applicant reactions are frequently an important concern for selection procedures as individuals' motivation and test performance, attraction to the organization, intentions to accept an offer, willingness to recommend the organization to others, and inclination to undertake legal challenges have all been tied to applicant reactions (Chan & Schmitt, 1997; Hausknecht, Day, & Thomas, 2004). Individuals react more positively to selection procedures that they perceive to be job-related (Bauer & Truxillo, 2006). As SJTs ask for responses to concrete situations that are typical in the jobs (or other contexts) for which they are designed, these tests should appear more job-relevant than tests of a more abstract nature (e.g., personality) (Bauer & Truxillo, 2006). Some consider SJTs to be a type of work sample (e.g., Gessner & Klimoski, 2006; Schmitt, Clause, & Pulakos, 1996), and multiple studies show that work samples tend to be rated more favorably (e.g., more job-related, more fair) than personality and cognitive ability tests (e.g., Lievens, De Corte, & Brysse, 2003; Steiner & Gilliland, 1996).

One study, that of Schmitt, Oswald, Kim, Gillespie, and Ramsay (2004), examined students' reactions to an SJT relative to other selection methods. The situational judgment test was seen as more relevant (for the job of being a college student) but less fair than standardized test scores. That the SJT would be seen as less fair is surprising in light of its greater perceived relevance in Schmitt et al.'s study and findings of favorable reactions to work samples more generally. Notably though, more commonly used selection procedures may be seen as more fair by virtue of individuals having more experience with them (Ryan & Huth, 2008); use of standardized tests in college admissions is common, whereas use of SJTs is not (Schmitt et al., 2004). Importantly, whereas minority (Black, Hispanic) respondents viewed standardized test scores as less fair than did White respondents, these groups had similar views on the SJT's

fairness. This is a desirable finding in light of the possibility that subgroup differences in test scores may stem in part from minorities having less favorable perceptions of the test instrument and lower motivation to perform (Edwards & Arthur, 2007; Ryan, 2001); if that is the case, improving reactions might lead to some reduction in adverse impact (Ployhart & Ehrhart, 2002; Ployhart & Holtz, 2008).

Research has examined reactions to SJTs as a function of delivery medium and indicates that reactions can be further enhanced by administering SJTs in a multimedia as opposed to a paper-and-pencil based format (Chan & Schmitt, 1997; Richman-Hirsch, Olson-Buchanan, & Drasgow, 2000). Chan and Schmitt (1997) found that students perceived a video-based SJT of work habits and interpersonal skills to have higher face validity with respect to a specific job they were asked to consider than a paper-and-pencil version of the test. In Richman-Hirsch et al.'s (2000) study, managers perceived a video-based conflict resolution SJT to have higher content and predictive validity relative to paper-and-pencil and computerized page-turner test versions. Managers also found the multimedia version more enjoyable and were more satisfied with the overall assessment process when completing that version of the test.

SJTs' Susceptibility to Faking and Coaching

A concern associated with use of SJTs for selection is that applicants may be motivated to distort their responses or "fake good" to appear better than they really are. It is therefore important to consider these tests' susceptibility to faking. Interestingly, it is typically harder to fake on SJTs than on traditional personality tests (Hooper, Cullen, & Sackett, 2006 provide a review), even when the two tests are designed to measure the same set of constructs (Hooper, Jackson, & Motowidlo, 2004). Personality items tend to be unidimensional; higher agreement with more positively valenced items and lower agreement with more negatively valenced items

tends to yield better scores. If respondents can figure out which end of the scale for an item is more desirable, they can fake good (Hooper et al., 2006). SJT items, on the other hand, are typically multidimensional, making it more difficult for respondents to identify the constructs they are measuring and to choose a response strategy that would yield higher scores (Hooper et al., 2006).

Hooper, et al.'s (2006) review of existing research indicated that the degree of faking on an SJT can vary widely (from .08 to .89 standard deviations). The extent of a particular SJT's fakability will depend on how the test was designed. Hooper et al. (2006) suggest that SJTs are harder to fake when they 1) Tap less fakable domains like cognitive ability (as opposed to more fakable domains like personality), 2) Have less transparent response options, and 3) Have knowledge-based (as opposed to behavioral tendency) instructions. The first and third points are closely related as knowledge-based instructions result in SJT scores with a higher cognitive loading (McDaniel et al., 2007). When individuals are provided with external motivation to achieve high scores on the SJT designed by Oswald et al. (2004), they are able to raise their scores by only a slight amount (d = .13) (Ramsay, Schmitt, Oswald, Kim, & Gillespie, 2006).

Closely related to the issue of faking is that of coaching, as both can lead to incorrect inferences about individuals' true standing on the constructs assessed by the SJT and have implications for process fairness (Hooper et al., 2006). A number of test preparation activities can be included under the "coaching" umbrella: increasing familiarity with the test, increasing knowledge of the content of the test domain, and improving test wiseness (i.e., knowledge of strategies that can be used to raise scores) (Hooper et al., 2006). It is likely that some of the same test design elements that make an SJT susceptible to faking (e.g., transparency; reviewed above) would also make it susceptible to coaching. It is not surprising then that different SJTs appear to

be differentially resistant to attempts to coach respondents (Cullen, Sackett, & Lievens, 2006). Importantly, the SJT developed by Oswald et al. (2004) appears relatively more resistant to the effects of coaching as compared to another SJT designed for use in academic contexts (Common Sense Questionnaire (CSQ); Sternberg & the Rainbow Project Collaborators, 2006). Test taking strategies for Oswald et al.'s (2004) SJT are more item-specific than the ones applicable for the CSQ, and thus may be more difficult and require higher cognitive ability to learn and apply (Cullen, Sackett, & Lievens, 2006). But it is important to note that when individuals are both motivated and receive coaching to achieve higher scores on an SJT, they may be able to raise their scores by a considerable amount (Ramsay et al., 2006). Although the validity of the test might remain unaffected (Cullen, Sackett, & Lievens, 2006), issues of fairness (e.g., differential access to coaching for majority and minority test takers) would still need to be considered (Ramsay et al., 2006).

Benefits of SJTs: Summary

Thus, there is mounting evidence to suggest utility for the use of SJTs in admissions settings. Importantly, the SJT developed by Oswald et al. (2004), which was adapted for use in the current research, has been demonstrated to predict a variety of academic and non-academic criteria relevant for the college context (Schmitt et al., 2009; Schmitt et al., 2010), and to have desirable properties such as reduced subgroup differences in performance (Schmitt et al., 2009), favorable reactions with regard to relevance for the college context (Schmitt et al., 2004), and relatively better resistance to a coaching intervention to raise scores as compared to another SJT developed for use with students (Cullen, Sackett, & Lievens, 2006).

Limitations of SJTs: Summary

SJTs do, however, have certain limitations, leaving room for further research on and improvement of these tests. First, fairly little is known about the construct validity of SJTs. Although it is generally accepted that SJTs are a method of measurement, debate continues about whether these tests measure some sort of unique construct as well (e.g., Chan & Schmitt, 2006). Most SJTs are developed via non-theoretical approaches and little attention is paid to the constructs they assess (Christian et al., 2010). By sampling common situations from the job domain, a developer can expect that the SJT will capture some relevant knowledge, skills, and abilities. But the developer typically cannot specify which constructs are assessed or maintain that these constructs are measured to the exclusion of other ones. SJTs are inherently multidimensional, mirroring the complex nature of human behavior in real situations. SJTs also tend to lack an interpretable factor structure. In the rare cases that researchers try to take a construct-centered approach to SJT development (e.g., Ployhart, 1999; Oswald et al., 2004), they have a challenging time producing SJTs with reliable and construct-valid subscales. Oswald et al. (2004), for example, reported low reliabilities in the .32-.55 range and poor discriminant validities for their SJT scales. In subsequent efforts, these researchers had to rely on overall SJT scores to have a measure with adequate reliability (e.g., Schmitt et al., 2009). A lack of understanding of the constructs a given SJT measures limits practitioners' ability to make good practical decisions about the use of an SJT. For example, it can be difficult to judge whether it will benefit incremental validity to add an SJT that measures an unknown set of constructs to a test battery that already assesses certain constructs (Gessner & Klimoski, 2006).

Second, understanding of SJTs' criterion-related validity is limited as well. There is a considerable amount of variability in SJT validity across applications but few moderators, aside

from response instructions (behavioral tendency versus knowledge-based) have been identified (McDaniel et al., 2001; McDaniel et al., 2007). The identification of additional moderators of SJTs' criterion-related validities can increase the validity of SJTs for predicting important outcomes in academic and employment settings. It is typical for researchers to examine the implications of overarching test design elements (e.g., response instructions, response format, delivery medium) for validity. But researchers seldom look to the individual items comprising the test (Reynolds, Sydell, Scott, & Winter, 2000). One heretofore unexamined moderator may be the "strength" of the scenarios making up the test, as situational strength has implications for how individual differences become manifest in behavior. But before defining situational strength and considering its relevance for the validity of SJTs, it is instructive to discuss the value of studying how individual difference and situational variables interact. Thus, the interactional approach to the study of behavior is reviewed next.

Interactional Approach to the Study of Behavior

Scholars in the area of individual differences have long called for research that takes into consideration how individual differences interact with situational elements to predict behavior (e.g., Endler & Magnusson, 1976; Hattrup & Jackson, 1996; Pervin, 1985). Whereas earlier approaches seek to explain variations between individuals' behavior with individual difference (e.g., traits, abilities, values, motives, affect) *or* situational variables, the interactional approach acknowledges that behavior will always be a function of *both* the person and the situation (Hattrup & Jackson, 1996). As aptly summarized by Hattrup and Jackson (1996), the interactional approach looks to "identify how different types of individuals respond to different situations in predictably unique ways".

Research findings attest to the relative merits of using person-situation interactions relative to individual differences and situational attributes on their own to predict behavior (Bowers, 1973 and Endler & Magnusson, 1976 provide reviews). However, where situational determinants of behavior are concerned, research has been held back by a tendency to study situational features in a non-systematic way. Interactional approaches seek to address this issue by trying to identify relevant situational dimensions ("taxonomies") that characterize situations in general and may be expected to influence how individual differences become manifest in behavior. A brief review of a number of attempts to categorize situations follows and transitions into a discussion of situational strength.

Situational Taxonomies

Several examples of attempts to categorize situations include, but are not limited to, those of Magnusson (1971), Forgas (1976), Wish, Deutsch, and Kaplan (1976), Eckes (1995), Ten Berge and De Raad (2002), and Wagerman and Funder (2006). Magnusson (1971) factor analyzed similarity judgments for thirty six college situations (e.g., undergo a written examination, carry out a joint group task together with fellow students) and derived five situational dimensions: positive, negative, passive, active, and social. Forgas (1976) likewise examined college situations and came up with three dimensions to characterize them. These were intimacy/involvement, self confidence/know how to behave, and pleasantness. Wish, Deutsch, and Kaplan (1976) analyzed ratings on 45 interpersonal relations (e.g., between close friends, between supervisor and employee) and derived four underlying dimensions, including intense versus superficial, cooperative/friendly versus competitive/hostile, socioemotional/informal versus task oriented/formal, and equal versus unequal. Eckes's (1995) cluster analysis of ratings of 30 situations that college students typically find themselves in produced nine situational

dimensions: non-intimate, emotionally uninvolving, emotionally involving, informal, relaxed, social, familiar social, frightening, and competitive. Ten Berge and De Raad (2002) derived four dimensions—pleasure, individual adversity, interpersonal conflict, and social demand—based on a factor analysis of individuals' ratings on their ability to deal with a given situation (237 situations were included—e.g., having an exam, having to speak in public). To date, the most comprehensive effort to categorize situations is that of Wagerman and Funder (Riverside Situational Q-sort; 2006). The instrument, which is still undergoing development, currently consists of 81 items (e.g., someone is trying to impress someone or convince someone of something, situation may cause feelings of hostility) (Wagerman & Funder, 2009). The intent is for raters to be able to describe any given situation by sorting the 81 items into 9 categories ranging from "highly uncharacteristic" to "highly characteristic" of the situation.

The extensive literature on this topic and the apparent lack of consistency across taxonomic efforts attests to a continuing lack of agreement on a taxonomy of situations (Johns, 2006). Two challenges in this research are that scholars generally target just one of many parts of the situational domain when proposing taxonomies, and that different situational dimensions may emerge depending on the characteristics of the participants who provide the data and the method used for classifying situations (Magnusson, 1971; Ten Berge & De Raad, 2002). But given that behavior results from individuals' personalized interpretations of objective features of situations (Lewin, 1951, as cited by Hattrup & Jackson, 1996), there *is* substantial agreement on the need to conceive of situations in terms of their psychological meaning as opposed to concrete manifestations (e.g., Campion & Ployhart, 2012; Endler & Magnusson, 1976; Funder & Colvin, 1991).

Situational Strength

The practical utility of using any given situational taxonomy arguably lies in the purpose for which it will be applied. In light of the current goal to predict success in college with certain individual differences that can be captured by an SJT, of concern are the dimensions of situations that could have implications for the test's criterion-related validity. One situational dimension that seems particularly relevant is situational strength, or the constraints a situation imposes on individuals' behavior (Mischel, 1977). Mischel (1977) describes strong situations as ones that **"lead everyone to construe the particular events the same way, induce uniform expectancies regarding the most appropriate response pattern, provide adequate incentives for the performance of that response pattern and require skills that everyone has to the same extent"** (p. 347). In sum, in strong situations, the "appropriate course of action" is clear (Meyer et al., in press).

Situational strength provides a strong motivation for dissimilar individuals to exhibit a common set of behaviors in a situation (**i.e., it constrains the effects of individual differences on behavior**). For instance, during a class lecture (a strong situation), students will tend to exhibit a common set of behaviors: listening to the professor, refraining from talking, and taking notes. During class discussion (a weaker situation), these students will exhibit different degrees of participation behavior: conscientious individuals who read ahead for class and those who feel comfortable talking in front of others will likely participate more than less conscientious and more socially anxious ones. As the notion of situational strength comes out of the personality literature, the concern has typically been with the way it may constrain personality effects (Cooper & Withey, 2009 provide a review).

Researchers suggest several attributes that can characterize situations as weak or strong (Hattrup & Jackson, 1996; Meyer et al., 2009; Meyer, Dalal, & Hermida, 2010; Schneider & Hough, 1995). Hattrup and Jackson (1996) delineate four broad attributes of situations: information attributes, task attributes, physical attributes, and social attributes. **Information attributes** refer to the "availability, ambiguity, and degree of consensus (versus conflict) among information cues" (p. 528). Stronger situations provide more, less ambiguous, and less conflicting information cues than weaker situations. **Task attributes** refer to the amount of autonomy and structure provided by tasks. Stronger situations are characterized by less autonomy and greater structure than weaker situations. **Physical attributes** of situations refer to the amount of privacy and danger present. Stronger situations involve less privacy (i.e., they tend to be more public) and more danger relative to weaker situations. Finally, **social attributes** pertain to what others in the situation want or expect. Stronger situations involve greater demands from others than weaker situations.

Meyer et al. (2010) also discuss four dimensions of situational strength: clarity, consistency, constraints, and consequences. **Clarity** is the degree to which cues regarding one's responsibilities are available and comprehensible. **Consistency** is the degree to which cues regarding one's responsibilities are consistent with each other. **Constraints** represent the degree to which one's decision-making and actions are constrained by external forces one cannot control. **Consequences** represent the degree to which one's decisions or actions have considerable consequences (positive, negative) for someone/something of relevance. Each of these four dimensions is positively associated with situational strength (e.g., situations with high clarity tend to be stronger). Meyer et al.'s dimensions show some conceptual overlap with those of Hattrup and Jackson (1996). Hattrup and Jackson's information attributes dimension is

represented by Meyer et al.'s clarity and consistency dimensions. Hattrup and Jackson's task attributes and social attributes dimensions are fairly similar to Meyer et al.'s constraints and consequences dimensions, respectively. Hattrup and Jackson's physical attributes dimension is the only one that does not seem to be captured by Meyer et al. All in all, the dimensions proposed by Meyer et al. (2010) seem to offer greater conceptual clarity (e.g., they treat information availability and consistency as distinct) than those of Hattrup and Jackson (1996) and may be more practical to use in studying situational strength as Meyer and colleagues (in press), in follow up work, designed a scale to measure their four dimensions described above. Next, research findings that speak to the validity of the situational strength hypothesis are reviewed.

Empirical Research on Situational Strength

Research findings provide some support for the hypothesis that strong situations, by providing strong cues for how individuals should behave, allow less latitude for individual differences to impact behavior. Consistent with the prerequisite that strong situations should be interpreted similarly by different individuals (Mischel, 1977), Meyer et al. (in press) found that raters better agree in their perceptions of strong situations than of weak situations. Further, research findings indicate that situational strength can indeed constrain the effects of individual differences (e.g., personality variables, attitudes) on behavioral intentions and behavior (e.g., citizenship, lateness, expended effort, talking) (Beaty, Cleveland, & Murphy, 2001; Elicker, Foust, O'Malley, & Levy, 2008; Meyer et al., in press; Monson, Hesley, & Chernick, 1982; Withey, Gellatly, & Annett, 2005). Meta-analytic findings for the personality variable of conscientiousness are available. Consistent with the situational strength hypothesis, the criterion-related validity of conscientiousness is moderated by situational strength such that validity is

higher for "weak" occupations than for "strong" ones (Meyer, Dalal, & Bonaccio, 2009). Notably, a couple of research studies have also considered the moderating role of situational strength in relations between personality and counterproductive work behavior, but findings are mixed. Smithikrai's (2008) findings were consistent with the situational strength hypothesis—correlations of agreeableness and conscientiousness with counterproductive work behaviors were smaller under strong situations than under weak, whereas Meyer and colleagues' (in press) findings were not (the authors unexpectedly found relations to be *stronger* under strong situations).

Cooper and Withey (2009) point out certain limitations in existing research on the situational strength hypothesis. Surprisingly, few studies have explicitly manipulated or measured situational strength (as conceptualized by Mischel, 1977) (Cooper & Withey, 2009). Some studies have only indirectly examined the role of situational strength by using other constructs as proxies for it (Cooper & Withey, 2009). For example, researchers (e.g., Barrick & Mount, 1993; Fuller, Hester, & Cox, 2010; Gellatly & Irving, 2001) have used job autonomy in lieu of situational strength, equating low job autonomy with a relatively strong situation and high job autonomy with a relatively weak situation. Group norms have likewise been used as a proxy for situational strength (e.g., Smithikrai, 2008). Relatedly, strict and lenient climates in organizations have been used as examples of strong and weak situations, respectively (e.g., Elicker et al., 2008). Cooper and Withey (2009) observe that researchers will often use situational strength as a way to interpret findings post hoc (e.g., Elicker et al., 2008 note that a reviewer drew the authors' attention to the fact that their study speaks to the situational strength hypothesis).

Taking into consideration the limitations of existing studies, Cooper and Withey (2009) take the position that "there are no well-designed studies that test the [situational strength] hypothesis" (p. 67). Recent research has started to address some of the limitations noted above. As noted above, Meyer and colleagues (2010, in press) conceptualized the sub-dimensions of situational strength and designed a scale that may be used to assess them. This work should facilitate research that better measures and/or manipulates situational strength. The current study will draw on this recent research to apply situational strength to the study of SJTs' validity. Existing research that can speak (at least indirectly) to this topic is reviewed next.

Situational Strength and Situational Judgment Tests

Situational strength is a meta-dimension of sorts as situations with a variety of characteristics can be located on a common continuum of situational strength (Hattrup & Jackson, 1996). This makes strength a particularly useful way of thinking about a wide range of possible situations comprising an SJT. The current research considers the role of situational strength as a moderator of the validity of SJTs. The implications of situational strength for the psychometric properties of SJTs have not been considered in previous research in any explicit sense. However, several potential correlates of situational strength (as it relates to SJT items) have been examined or at least discussed as a topic worthy of research.

Elaboration, complexity, detail, specificity, and length are all terms that come up in these discussions and refer to the amount of information SJT items contain. Item length is clearly not synonymous with these other terms, but in practice is related to them, as more complex situations require more words to describe (McDaniel & Nguyen, 2001). McDaniel et al. (2001, 2006) reason that more specific SJTs will tend to be longer, and as a result, have higher reading requirements and stronger correlations with cognitive ability. Given that cognitive ability is the

best predictor of performance (Schmidt & Hunter, 1998), the more SJTs tap test takers' cognitive ability, the higher their criterion-related validity could be (McDaniel et al., 2006). Consistent with this reasoning, two studies found better validity for more specific items (Reynolds, Sydell, Scott, & Winter, 2000; Reynolds, Winter, & Scott, 1999). However, McDaniel et al.'s (2001) meta-analysis, contrary to their expectations, found less detailed SJTs to have higher criterion-related validity. To the extent that amount of detail is a proxy for situational strength, this handful of studies may speak to the implications of situational strength for SJTs' validity. Brooks and Highhouse (2006), for example, implicitly equate level of detail in an item with its situational strength when they suggest that too little information may lead different test takers to construe and respond to a situation differently. Clearly though, research that explicitly considers situational strength is needed to better understand the implications of situational strength for the validity of SJTs.

Although existing investigations of SJT item specificity may not be directly relevant to the issue of situational strength, they do point to the importance of keeping item length constant (so as to keep constant the cognitive loading of SJT scores) when manipulating situational strength. In other words, it is critical to not confound item strength with item length, as longer items will probably have higher criterion-related validity because of stronger correlations with cognitive ability. Next, the hypotheses for the current research are reviewed.

Current Study

The primary goal of the current study was to examine whether situational strength can moderate the criterion-related validity of an SJT. Given that situational strength can constrain the effects of individual differences on behavior (e.g., Beaty et al., 2001; Meyer et al., in press; Withey et al., 2005), there should be more (less) variability in individuals' interpretations of, and

responses to weak (strong) test situations. Greater (lower) variability in responses on the test should be associated with greater (lower) validity when using test scores to predict academic performance and other criteria. A representation of this prediction is included in Figure 1.

H1a: There will be an interaction of SJT dimension scores[1] *and SJT strength (weak, strong) on criterion-related validities such that validities will be lower for a strong SJT.*

H1b: There will be an interaction of the overall SJT score and SJT strength (weak, strong) on criterion-related validities such that validities will be lower for a strong SJT.

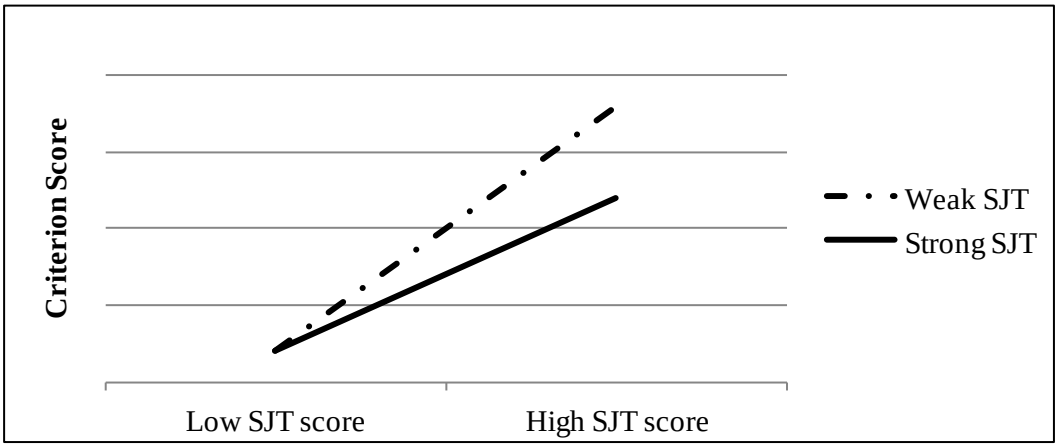

Figure 1. **Illustration of Hypothesis 1.**

There is evidence that the construct validity of SJTs can be altered by changing elements of the test's design (e.g., delivery medium, response instructions, response format, item specificity) (e.g., Chan & Schmitt, 1997; Glaze et al., 2011; McDaniel et al., 2001; Ployhart & Ehrhart, 2003). The current study will also examine whether situational strength is another

[1] Analyses to test hypotheses at the level of SJT dimensions are contingent on finding internally consistent SJT dimensions, which is a challenge as attested to by prior work (e.g., Oswald et al., 2004). As described below, an effort was made to add items to Oswald et al.'s SJT dimensions so as to improve dimensions' internal consistency.

element that can affect the construct validity of an SJT. Construct validity is typically assessed by examining measures' convergent and discriminant validities (or correlations) with other measures designed to assess the same and different individual differences, respectively (Campbell & Fiske, 1959). Correlations between different measures of the same phenomenon should be high (indicating convergent validity for the target measure); correlations between different measures of different phenomena should be low (indicating discriminant validity for the target measure).

To the extent that altering situational strength of SJT items can change how well the test measures the individual differences it is designed to assess[2], the convergent validities of the SJT items (or more appropriately, dimensions consisting of multiple items each) with other measures designed to tap the same individual differences should be affected as well. In other words, the situational strength of the SJT should moderate convergent validities. Given the expectation that strong SJT items will constrain the effect of individual differences on test responses, associations between responses on the SJT dimensions and the dimensions of another measure of the same individual differences should be weaker (stronger) for the strong (weak) SJT. The same effect should be observed at the level of the overall measure (i.e., different dimensions combined). A representation of this prediction is included in Figure 2.

H2a: Convergent validities of SJT dimension scores with dimension scores on a measure designed to index the same performance dimensions as the SJT will be moderated by SJT strength such that validities will be lower (i.e., less favorable) for a strong SJT.

[2] Note though the challenge that SJTs generally do not have very good construct validity (McDaniel et al., 2006).

H2b: Convergent validity of the overall SJT score with the overall score on a measure designed to index the same performance dimensions as the SJT will be moderated by SJT strength such that validity will be lower (i.e., less favorable) for a strong SJT.

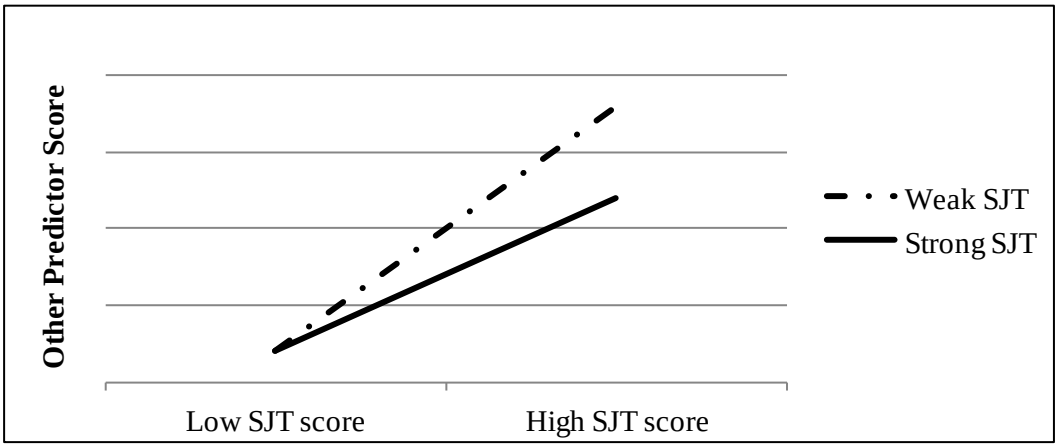

Figure 2. **Illustration of Hypothesis 2.**

Situational strength may also moderate the discriminant validities of the SJT's dimensions. Given the expectation that weak SJT items will capture the individual differences that they are designed to assess more effectively than strong SJT items, associations between dimensions designed to assess different individual differences should be weaker (stronger) for the weak (strong) SJT. This effect should be observed for intercorrelations between different dimensions of the SJT as well as intercorrelations of SJT dimensions with the dimensions of another measure designed to predict students' performance. A representation of this prediction is included in Figure 3.

H3: Discriminant validities of SJT dimension scores with dimension scores designed to index different performance dimensions will be moderated by SJT strength such that discriminant validities will be lower (i.e., more favorable) for a weak SJT.

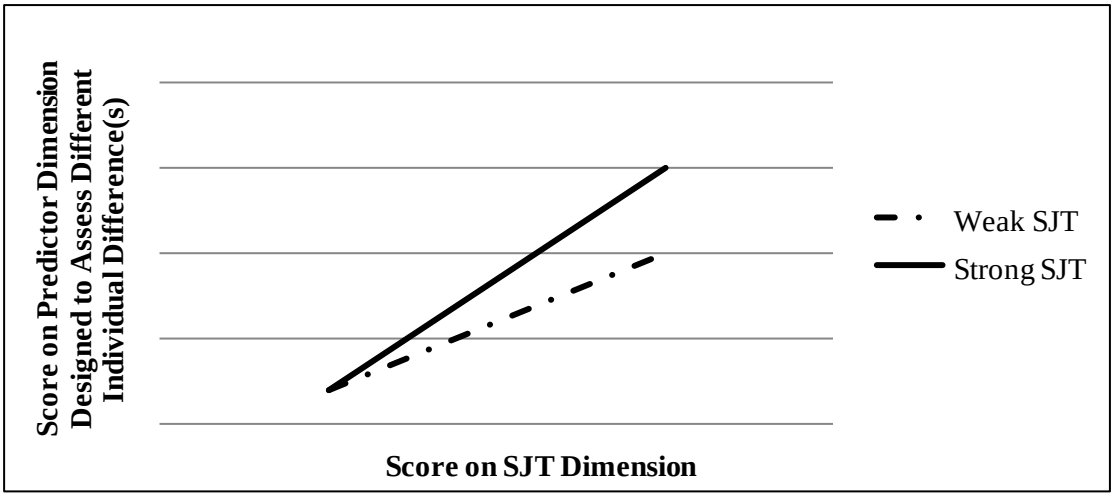

Figure 3. **Illustration of Hypothesis 3.**

Given that SJT items may capture multiple individual differences (e.g., conscientiousness, agreeableness) pertinent to performance in a domain (e.g., social responsibility), additional explanatory value for any situational strength effects may be provided by examining the interrelations of SJT dimensions with personality scales. While there is evidence that the effects of conscientiousness may be stronger in weak situations (Meyer et al., 2009), there is less or less consistent evidence available for other personality dimensions. Thus, this issue will be posed in the form of a research question.

RQ1: Will situational strength moderate the interrelations of the SJT dimension scores with openness, conscientiousness, extraversion, agreeableness, and emotional stability scores?

Given that the appropriate course of action is clearer in strong situations than in weak (Meyer et al., in press), a stronger SJT could prove to be more susceptible to faking. One approach to examining this issue without experimentally manipulating faking, is to administer a measure of socially desirable responding and check to see whether the association between

scores on this measure and scores on the SJT is stronger for the stronger SJT. This issue of test fakability will take the form of a research question as well.

RQ2: Will situational strength moderate the relations between scores on the SJT and scores on a measure of socially desirable responding?

Other questions addressed by the current study are gender- and race-based score differences on the SJT relative to score differences on the more cognitive predictors of student success (i.e., standardized test scores and high school GPA), as well as incremental validity of the SJT above the more cognitive measures for academic and non-academic criteria. The method employed to test this study's hypotheses is described next.

METHOD

Sample

Pilot. A total of 16 graduate students in organizational psychology provided ratings for the SJT stems (six raters provided ratings for the weak SJT stems, six raters for the strong stems, and four raters for the stem pairs). The sample was 56.3% female and 68.8% White (the remainder were Asian—18.8%, Hispanic—6.3%, and Black—6.3%). The sample was 12.5% 1st years, 31.3% 2nd years, 12.5% 3rd years, 12.5% 4th years, 25% 5th years, and 6.3% recent graduates.

Scoring key development. The sample for the scoring key development was drawn from the psychology participant pool at Michigan State. A total of 157 juniors and seniors (experts) participated in exchange for partial credit in their psychology courses. After excluding 44 individuals (i.e., 28% of total sample) who were not juniors/seniors (4 people) and/or who failed one or more of the four attention checks (42 people), the final sample size was 113 individuals (59 for the weak SJT items and 54 for the strong SJT items). Notably, approximately equal percentages of the data had to be excluded for the weak (29% excluded) and strong (27% excluded) SJT items. The final sample was 79.6% female, 76.1% White (the remainder were Asian—7.1%, Black or African American—7.1%, Hispanic—4.4%, multiracial—4.4%, and "other"—0.9%), and 56.6% juniors. These students were 21.0 years old on average (*SD* = 1.46). The majority (86.7%) reported a cumulative college GPA between 2.60 and 3.79.

Lab study. The sample for the lab study was also drawn from the psychology participant pool at Michigan State. Second semester freshmen were invited to participate in the study in exchange for partial credit in their psychology courses and entrance into a lottery to win one of fifteen $30 Amazon.com gift cards. A total of 557 students participated in the study. After

excluding 149 individuals (i.e., 27% of total sample) who indicated they were not freshmen (43 people) and/or who failed one or more of the five attention checks (111 people), the final sample size was 408 individuals (208 for the weak SJT condition and 200 for the strong SJT condition). Notably, equal percentages of the data had to be excluded for the weak (27% excluded) and strong (27% excluded) SJT conditions. The final sample was 78.5% female and 80.2% White (the remainder were Black or African American—6.9%, Asian—6.7%, Hispanic—2%, multiracial—3%, American Indian or Alaska Native—0.2%, and "other"—1%). These students were 18.3 years old on average (*SD* = .49).

Procedure

Creation of two SJT versions. A set of items from the SJT designed by Oswald et al. (2004) to tap twelve dimensions of college student performance were adapted in the current study. This study focused on four of their performance dimensions: citizenship, adaptability, perseverance, and ethics. It was expected that the items comprising these dimensions would be most amenable to manipulation of situational strength.

An issue of concern with the SJT adapted for this study was that the dimensions the test was designed to tap had poor psychometric properties. Oswald et al. (2004) reported low reliabilities in the .32-.55 range and poor discriminant validities for the individual 3-6 item SJT scales. This test's broad dimensions (i.e., intellectual, interpersonal, and intrapersonal) evidenced low reliabilities as well (.44-.65 range; based on data from Schmitt et al., 2009). As a result, only total SJT scores have been used in the past (e.g., Schmitt et al., 2009). Nonetheless, the current study attempted to examine the test at the dimension level with the expectation that increasing scale length and rewriting the scenarios to be "weaker" could improve the psychometric

properties (e.g., internal consistencies, convergent validities, discriminant validities) of these dimensions.

Drzakowski et al. (2004) selected 3 SJT items per dimension from the Oswald et al.'s initial bank of 153 items that maximized the expected reliability and validity of the test; the current effort started with this particular set of items. Available data were used to estimate the internal consistency reliability of these scales. Reliabilities were .102, .266, .327, and .420 for citizenship, adaptability, perseverance, and ethics, respectively. An effort was made to improve scale reliabilities by adding items.

Six additional items were selected per dimension (with the goal of having more reliable, nine-item dimensions). Items with favorable correlations with college GPA and item-total correlations were preferred. In cases where the item bank did not provide enough items with favorable properties to choose from, or when items could not be adapted to vary in situational strength, additional items were drawn from available research. Appendix A provides unaltered items organized by dimension.

Information in the SJT item stems was varied along the dimensions suggested by Meyer et al. (2010) to create low strength (e.g., low clarity, low consistency, low constraints, or low consequences) and high strength (e.g., high clarity, high consistency, high constraints, or high consequences) versions of the test. Several guidelines were followed when developing the two versions of each item: (1) Keep the basic situation the same across test versions; (2) Avoid completely invalidating any of the response options; (3) Avoid introducing additional necessary response options in cases when the item already has six response options; and (4) Ensure an approximately equal number of words to keep test length constant (to avoid confounding test strength with test length) and try to keep readability statistics at similar levels. Given the

limitations this imposed on item development, it was necessary to interpret the situational strength dimensions somewhat more flexibly than intended by Meyer et al. (in press). For example, *constraints* included not only constraints imposed on a person by external forces, but in some cases constraints created by the individual's own needs or abilities (see the first adaptability item for an example).

Across the entire set of 41 items, the following aspects of situations were varied most often (at least ten times) when creating a weak and strong version of a particular situation:

1. Clarity of someone's (e.g., professor's, friend's, parents') expectation/preference/need/plan or knowledge of relevant policy regarding what should be done (17 times)
2. Consequences of some behavior (e.g., cheating) or work product (e.g., project, test) for someone's grade or job prospects (15 times)
3. Amount of time remaining for getting something accomplished (10 times)

The following aspects of situations were varied less often (fewer than ten times) when creating weak and strong versions of situations:

1. Closeness of relationship to individual(s) involved (e.g., acquaintance vs. close friend) (8 times)
2. Availability of another helper(s) (e.g., friend in class, professor) (7 times)
3. Likelihood of getting into trouble or experiencing embarrassment (6 times)
4. Individual's level of relevant ability, knowledge, or preparation (6 times)
5. Workload/number of other commitments; level of time availability (5 times)
6. Number of opportunities remaining to raise one's grade (e.g., first exam vs. final exam) (5 times)

7. Severity of the issue at hand (e.g., friend had a few drinks vs. is intoxicated; someone was mistreated vs. physically hurt) (5 times)
8. Previous/to date performance in the class(es) (e.g., receiving Bs and Cs vs. Ds and Fs) (5 times)
9. Needed/desired grade (e.g., in the class, on a test) (5 times)
10. Consistency of someone's behavior (e.g., over time) (4 times)
11. Consistency of relevant individuals' advice/opinions/preferences (4 times)
12. Availability of relevant information or advice (4 times)
13. Likelihood that someone will get hurt (3 times)
14. Effort or time already invested (3 times)
15. Amount of time expected to commit (3 times)
16. Level of interest in the topic (2 times)
17. Class size (2 times)
18. Amount of pressure from close others (2 times)
19. Physical proximity to person needing help (1 time)
20. Have a personal need versus not (1 time)

While item stems varied slightly across test versions, the response options did not. Notably, in cases where an item had fewer than 5 response options, one to two additional options were added for greater flexibility during scoring key development. Appendix B offers two versions of each item included in Pilot study (see "Pilot" below). One to two extra items per dimension are shown, for a total of 10 to 11 items per SJT scale. These additional items were included as backup options in case SMEs judged certain items to be poor (see "Pilot" below).

Pilot. Subject matter experts (SMEs)—graduate students in organizational psychology—were asked to examine the initial pool of 41 weak and strong item stems and to rate them on the four dimensions of situational strength and separately, on overall/global strength. Definitions of the facets of situational strength and of overall strength were provided for this task along with the item stems to be evaluated.

Two versions of the SJT item evaluation were used. The first version asked for ratings of either the weak *or* the strong SJT stems on the four dimensions of strength and on overall/global strength. The 7-point response scale ranged from *Extremely Low* to *Extremely High*. Raters were randomly assigned to the weak or strong SJT stems. The drawback to evaluating the effectiveness of the strength manipulation by asking raters to rate one set of stems without seeing the other is that raters would lack a common understanding of the strength continuum, and thus may show little agreement in their ratings of item strength (dimensions of, overall). On the other hand, individuals completing the SJT as a test also responded to either the weak or strong SJT items, not both.

The second version of the SJT item evaluation presented each pair of SJT stems (weak and strong) side by side and asked raters to indicate the extent to which the stems were different on each dimension of strength and overall strength and the direction of that difference. The 7-point response scale ranged from *A Extremely Higher* to *B Extremely Higher*. The order of the item pair was randomized (e.g., weak first, strong first) so that raters would not expect the stronger stem to always be in the same position. Instructions that were provided to SMEs for their rating task and a sample item stem or item stem set that was rated are included in Appendices C (instructions to those who rated the weak OR strong stems) and D (instructions to those who compared the weak and strong stems side by side).

Scoring key development. Given that varying the SJT item stems on situational strength could change the interpretation of these items, and subsequently the responses to these items, two separate scoring keys were developed. The scoring keys were developed using the approach described by Friede et al. (2003). The two test versions were administered to two different samples of advanced college students during a supervised data collection. Students were randomly assigned to respond to one of the two test versions. They were asked to rate each SJT response option on effectiveness (ranging from 1-*highly ineffective* to 5-*highly effective*) as well as to indicate the "best" and "worst" response for each item. The order of the SJT items was counterbalanced to alleviate the effects of fatigue on later items in the rating form. Four attention checks were inserted at approximately equal intervals of the survey in order to flag random responding (e.g., For quality assurance purposes, please select 'Highly Ineffective' for this item).

Asking advanced students to serve as experts assumes that they have been successful in college (e.g., they have persisted). Advanced students may indeed know the appropriate way to behave in various situations; earlier work has shown that advanced students and resident hall advisor experts produce similar scoring keys (Friede et al., 2003). The scoring key combined mean effectiveness ratings for a given alternative with the percentage of SMEs who indicated it as the best or worst alternative. The item keying procedure is detailed in Appendix E.

Because these data were being collected at the same time as SME ratings of the item stems, and therefore the final set of items (based on SME ratings) was not yet determined, students were asked to evaluate the entire pool of forty one SJT items. When data were collected from an additional set of students during the subsequent semester to supplement initial sample sizes, results of the "Pilot" were already available, so the survey was abbreviated to include only the eighteen items with an effective strength manipulation (see results of "Pilot" below). The

trimming of the survey was expected to improve the quality of the collected data as participants should have been less affected by fatigue and tedium.

Lab study. Research participants attended a supervised data collection session. They were asked to provide their permission for the researcher to collect their HS GPA, SAT/ACT, and first year college GPA from the university. Participants were then randomly assigned to complete an SJT consisting of either high or low strength SJT items. Everyone also completed the biodata, personality, BARS, organizational citizenship, impression management, and demographic items. Participants were also randomly assigned to one of two survey orders. Version 1 ordered the measures as follows: personality, SJT, manipulation check, biodata, BARS, organizational citizenship, impression management, demographics. Version 2 ordered the measures as follows: organizational citizenship, biodata, SJT, manipulation check, personality, BARS, impression management, demographics). Thus, there were four versions of the survey (2 SJT versions x 2 order versions). The manipulation check always followed directly after the SJT to allow for maximum recall of the set of SJT items referred to by the manipulation check questions. BARS were always completed after the SJT because some of the behavioral examples and anchors for performance that is "Unsatisfactory", "Fulfills Expectations" and "Exceptional" were similar to the SJT items and suggested which responses may be more socially desirable.

Measures

Predictor measures in this study included an SJT (two versions), measures of biographical data and personality, high school GPA, and standardized test scores. Consistent with earlier research on non-cognitive predictors of college student success (e.g., Schmitt et al., 2010), outcome measures included first year college GPA, behaviorally anchored rating scales, and organizational citizenship behaviors. An impression management scale was used to examine

the susceptibility of the two SJT versions to faking. Two measures were used for quality assurance purposes and to screen out individuals not paying attention or providing random responses (see Quality Check and Manipulation Check sections below). Basic demographic information was collected as well.

SJT. The SJT asked individuals to respond to hypothetical situations with appropriate courses of action. Consistent with earlier research using this test in the academic context (e.g., Oswald et al., 2004; Schmitt et al., 2009), respondents picked their *Most* and *Least Likely* responses. Additionally, they were asked to rate each response option according to their likelihood of responding to the given situation in that way.[3] The 5-point response scale ranged from *Very Unlikely* to *Very Likely*. Behavioral tendency instructions were an appropriate choice in light of the current goals to contribute to research on the non-cognitive predictors of college student success; behavioral tendency instructions make SJTs better measures of personality (versus cognitive ability) (Whetzel et al., 2008). Relatedly, behavioral tendency instructions can produce lower mean score differences between White and minority test takers, and alleviate issues of adverse impact (McDaniel et al., 2007). Likelihood ratings were collected in addition to most and least likely responses as they tend to produce higher reliability (Ployhart & Ehrhart, 2003). This provided for two methods of scoring individuals' responses on the SJT.

Scores on each item based on Most and Least Likely responses ranged from -2 (respondent's most and least likely responses to the item corresponded with alternatives experts judged to be worst and best, respectively) to +2 (respondent's Most and Least Likely responses corresponded with alternatives experts judged to be best and worst, respectively) (Friede et al.,

[3] The decision to also ask for likelihood ratings of the SJT response options was made following work that was done with the initial pool of 41 items to 1) Ensure that the strength manipulations were working (SME ratings of SJT item stems) and 2) To develop the scoring keys (expert ratings of SJT response options). See discussion of that work below.

2003). Scores on each item based on likelihood ratings were computed by summing the ratings for the correct option(s) and reverse-scored ratings for the incorrect option(s) (Pereira & Harvey, 1999, as cited by Ployhart & Ehrhart, 2003). With this approach, scores on an item with one correct and one incorrect response ranged from 2 (respondent rated the experts' best response as Very Unlikely and the worst response as Very Likely) to 10 (respondent rated the expert's best response as Very Likely and the worst response as Very Unlikely). If an item had more than one best (worst) response as judged by the experts, a respondent's ratings of the multiple best (worst) responses were averaged together.

Item scores based on applying a given scoring approach were summed into overall test scores. As discussed above, two different keys were applied to score the weak and strong versions of the SJT. SJT scores based on Most Likely/Least Likely responses had relatively low internal consistency ($\alpha = .55$); the internal consistency of SJT scores based on Likert ratings of responses was higher ($\alpha = .72$).

Biodata. Biographical data ("biodata") items developed by Oswald et al. (2004) to measure the same performance dimensions as the SJT were used to examine the convergent and discriminant validities of the low and high strength versions of the SJT. Biodata measures traditionally assess individuals' past behavior with the expectation that past behavior is the best predictor of future behavior, but have evolved to include interests, preferences, and self appraisals as well (Schmitt & Golubovich, 2013).

The biodata scales used in this study contained about 10 items per performance dimension (again, citizenship, adaptability, perseverance, and ethics in this study), for a total of 40 items. Each item is scored on a four- or five-point multiple-choice scale. A sample citizenship item is "How many times in the past year have you volunteered in social service or charity

organizations?" Appendix F provides the biodata items organized by dimension. The four biodata dimensions have marginal to adequate reliability (citizenship $\alpha = .76$, adaptability $\alpha = .65$, perseverance $\alpha = .70$, and ethics $\alpha = .64$).

Personality. A personality measure was administered with the purpose of examining the interrelations between the SJT dimensions and personality scales. A 50-item measure of the Big Five personality traits (agreeableness, conscientiousness, openness to experience, extraversion, and emotional stability) from the International Personality Item Pool (Goldberg, 2011) was used. A sample conscientiousness item is "Like order". The 5-point response scale ranged from *Very Inaccurate* to *Very Accurate*. Items are presented in Appendix G. The scales comprising the measure have good internal consistency (openness $\alpha = .81$, extraversion $\alpha = .89$, conscientiousness $\alpha = .83$, agreeableness $\alpha = .80$, and emotional stability $\alpha = .84$).

HS GPA and standardized test scores. Participants were asked for permission to collect their HS GPA and ACT/SAT scores from the university's Office of the Registrar. These data were collected to examine the incremental validity of the SJT over cognitive measures for predicting the outcomes of interest as well as to compare subgroup score differences.

First year college GPA. Participants were also asked for permission to collect their first year college GPA from the university's Office of the Registrar. GPA was collected to examine the validity of different versions of the SJT for predicting academic performance.

Behaviorally Anchored Rating Scales (BARS). The BARS were designed to measure students' self-reported college performance along twelve dimensions (knowledge, continuous learning, artistic appreciation, multicultural appreciation, leadership, interpersonal skills, citizenship, health, career orientation, adaptability, perseverance, and ethics; Oswald et al., 2004) and were used to examine the criterion-related validity of the two versions of the SJT. Each of

the twelve items includes a dimension name and definition, and two examples of college-related critical incidents. Each critical incident has behavioral anchors representing three levels of performance on a seven-point scale. Test takers used this information as a basis for rating their level on a given dimension. Ratings were made on a 7-point scale ranging from *Unsatisfactory* to *Exceptional*. Items are shown in Appendix H. The 12 item-item measure has adequate internal consistency (α = .77).

Organizational Citizenship Behaviors (OCBs). OCBs are discretionary behaviors that are not directly rewarded but that promote the functioning and welfare of an institution (Organ, 1997). A measure of OCBs was another criterion to examine the validity of the two SJT versions against. The current study used a ten-item OCB measure, tapping interpersonal helping and loyalty, which has been used in earlier research with college students (Sinha et al., 2011). A sample item is "Gone out of your way to help other students from your school with social problems". Responses were made on a 5-point scale ranging from *Very Infrequently/Never* to *Very Frequently/Always.* Items are shown in Appendix I. The items have adequate internal consistency (α = .77).

Impression management. The impression management scale from the balanced inventory of desirable responding (BIDR; Paulhus, 1991), a widely used measure of socially desirable responding (Li & Bagger, 2006), was administered to examine whether the stronger version of the SJT may be more susceptible to faking than the weaker version. The impression management scale consists of 20 items and is used to evaluate a person's tendency toward presenting him or herself in a social desirable manner. A sample item is "I always obey laws, even if I'm unlikely to get caught." One of the items from the scale—"I never read sexy books or magazines"—was excluded as individuals could conceivably find it particularly intrusive and

inappropriate. Items were rated on a 5-point scale ranging from *Not True* to *Very True*. Items are shown in Appendix J. The items have adequate internal consistency (α = .76).

Manipulation check. To verify that the situational strength manipulation worked, test takers were asked to respond to a set of questions following the SJT. Twenty seven items (seven each for clarity, consistency, and constraints, and six for consequences) were adapted from a measure created by Meyer et al. (in press) to assess job-related situational strength and rewritten to refer to the SJT. A sample clarity item is "The situations in the questions you responded to above provided straightforward information about what someone in the situation needs to do to succeed". Individuals responded to these items on a 5-point scale ranging from *Strongly Disagree* to *Strongly Agree*. Appendix K presents the list of items. The four scales have good internal consistency (clarity α = .82, consistency α = .88, constraints α = .88, and consequences α = .87).

Quality check. Five items were scattered throughout administered self-report measures—one quality check item per 35 self-report items (175 self-report items, not counting demographics)—to flag random responding. Such items stated: "This item is for data processing purposes only. Please mark '__'." Items varied according to the response option test takers were instructed to select.

Demographics. Name and university personal identification number were collected for linking self-report data to objective data collected from the university. Ethnicity/race, gender, age, year in school, and intended major were collected. This information was used to characterize the sample as well as to examine mean SJT score differences by gender and ethnicity/race. Participants were also asked to self-report their HS GPA and ACT/SAT scores.

Demographic questions are included in Appendix L. Results of the various stages on this research are reviewed next.

RESULTS

Pilot

Table 1 presents the ratings SMEs provided for the weak and strong stems independently. Mean differences were calculated between the two independent sets of ratings for each SJT item so as to examine whether the strong version of the stem was perceived to be stronger than the weak version of that stem. As expected, rater agreement was poor for this rating task (*M* r_{wg} = .11, *SD* = .54 for weak stems; *M* r_{wg} = .11, *SD* = .59 for strong stems)[4] as SMEs apparently differed in their understanding of the strength continuum. Another contributing factor to the low agreement values may have been the fact that item stems were presented to SMEs without the corresponding response options. Some SMEs who completed this task indicated that they had come up with their own versions of possible response options, which could have influenced their situational strength ratings. This issue was remedied for the second version of the SME rating task (where SMEs rated the item stems side by side); the response options for each SJT item were provided to these individuals.

Table 2 presents the ratings that SMEs provided while considering the two versions of each SJT item side by side. SMEs' ratings were recoded for analyses to range from -3 ("A Extremely Higher") to +3 ("B Extremely Higher") such that a rating of "A and B Equal" corresponded to a value of 0. As expected, rater agreement was considerably better for this rating

[4] Agreement was slightly better for ratings of consequences (*M* r_{wg} = .37, *SD* = .54 for weak stems; *M* r_{wg} = .24, *SD* = .48 for strong stems) and consistency (*M* r_{wg} = .20, *SD* = .59 for weak stems; *M* r_{wg} = .26, *SD* = .38 for strong stems) relative to ratings of clarity (*M* r_{wg} = -.07, *SD* = .53 for weak stems; *M* r_{wg} = -.01, *SD* = .63 for strong stems), constraints (*M* r_{wg} = -.05, *SD* = .66 for weak stems; *M* r_{wg} = -.07, *SD* = .55 for strong stems), and global strength (*M* r_{wg} = .05, *SD* = .53 for weak stems; *M* r_{wg} = .13, *SD* = .59 for strong stems).

task (*M* r_{wg} = .57, *SD* = .33)[5]; these SMEs were more likely to have a common understanding of the strength continuum because 1) Both the weak and strong version of an SJT stem was available to them, and 2) They had each SJT item's response options to refer to for additional contextual information.

Both Tables 1 and 2 present ratings on clarity, consequences, consistency, constraints, and global strength, for each of the 41 SJT stems. It is noted whether a given strength dimension was manipulated for a particular stem (fewer than four dimensions of strength were manipulated for most of the SJT items). Based on whether a strength dimension was manipulated and the ratings SMEs provided for that dimension, the difference between the two versions of the stem perceived by the SMEs is noted to be either *expected* or *unexpected*. The difference was labeled as *expected* if 1) The stronger version of a particular stem was seen as higher on the manipulated dimension than the weaker version of that stem or 2) No difference was seen between the stronger version of a particular stem and the weaker version along a dimension that was not manipulated. The nature of the difference was labeled as *unexpected* if 1) The stronger version of a particular stem was not seen as higher on the manipulated dimension that the weaker version of that stem or 2) There was a difference seen between the stronger version of a particular stem and the weaker version along a dimension that was not manipulated. In cases where the difference was unexpected due to a difference between stem versions along a non-manipulated dimension, the difference was further noted as being either *desirable* (i.e., stronger version of the stem was perceived as higher on the non-manipulated dimension) or *undesirable* (i.e., stronger version of

[5]Agreement was slightly better for ratings of consequences (*M* r_{wg} = .64, *SD* = .25) and constraints (*M* r_{wg} = .63, *SD* = .32) relative to ratings of clarity (*M* r_{wg} = .56, *SD* = .35), consistency (*M* r_{wg} = .49, *SD* = .38), and global strength (*M* r_{wg} = .52, *SD* = .34).

the stem was perceived as lower on the non-manipulated dimension) for the purposes of the study.

The two sets of SME ratings (Tables 1 and 2) were examined to identify SJT items for which the strength manipulation appeared to have worked (i.e., the stronger version of the SJT item was rated as being higher on the various strength dimensions). Particular attention was paid to whether or not the global strength difference between the weak and strong versions of a stem was in the expected direction for both sets of SME ratings. The difference in the manipulated facets of strength needed to be in the expected direction as well, particularly for SME ratings that were provided on the stem pairs viewed side by side. For six SJT items, SMEs who rated the weak and strong stems independently did not rate all the manipulated strength dimensions higher for the stronger version of the stem. For all these items, however, ratings were as expected for SMEs who compared the two versions of a stem side by side. More weight was given to the latter ratings in light of the better agreement between raters. For the vast majority of items that met the aforementioned criteria, the raters unexpectedly perceived differences in non-manipulated dimensions between the weak and strong versions of a stem. However, these differences were in a desirable direction (the stronger stem was seen as higher on a particular facet of strength).

Table 3 presents the eighteen SJT items that were kept based on the criteria described above. Five were citizenship, four were adaptability, six were perseverance, and three were ethics items. Average rater agreement by strength dimension (see Table 4) tended to be the same or slightly higher for this subset of items relative to average rater agreement values for the initial set of 41 items.

Rather than revising additional items that did not meet criteria to end up with thirty six usable items in total, the decision was made to use an eighteen item measure but to require test takers to provide likelihood ratings of the response options (in addition to most likely and least likely responses). Getting ratings of the response options was expected to help in addressing the issue of low scale reliabilities (Ployhart & Ehrhart, 2003). Further, the reduced test length of 18 items instead of the planned 36 lowered the burden on respondents (who had a number of other measures to complete during the experiment), thereby also maximizing data quality. Notably, initial data collection to develop the scoring keys resulted in a large number of participants having to be excluded due to inattention. Fatigue and tedium as a result of having to respond to a large number of SJT items were likely contributing factors. This data collection is described below.

Table 1. **Independent ratings of weak and strong SJT stems**

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	C 1	4.67	1.97	-.81	6	5.67	1.86	-.62	6	NM	-1.00	.51	UE	UD			
conseq	C 1	6.40	.55	.86	5	2.50	1.97	-.82	6	M	3.90	.01	E				
consist	C 1	2.33	1.03	.50	6	4.17	1.94	-.76	6	NM	-1.83	.09	UE	UD			
constr	C 1	6.50	.84	.67	6	6.83	.41	.92	6	NM	-.33	.46	UE	UD			
strength	**C 1**	4.33	1.97	-.81	6	5.67	1.51	-.06	6	**M**	**-1.33**	.16	**UE**		Y	N	No, undesirable
clarity	C 2	3.83	2.04	-.95	6	3.67	1.51	-.06	6	NM	.17	.87	UE	D			
conseq	C 2	5.67	1.51	-.06	6	5.00	1.26	.25	6	M	.67	.27	E				
consist	C 2	2.83	.75	.74	6	3.00	.63	.81	6	NM	-.17	.65	UE	UD			
constr	C 2	3.17	1.47	-.01	6	3.83	1.60	-.20	6	M	-.67	.36	UE				
strength	**C 2**	4.00	1.79	-.50	6	4.00	1.10	.44	6	**M**	**.00**	.80	**UE**		P	N	No; 1 desirable, 1 undesirable
clarity	C 3	4.83	1.94	-.76	6	3.33	1.37	.13	6	M	1.50	.21	E				
conseq	C 3	5.00	.63	.81	6	3.17	1.47	-.01	6	M	1.83	.02	E				
consist	C 3	5.17	1.60	-.20	6	3.83	.98	.55	6	NM	1.33	.07	UE	D			
constr	C 3	2.33	1.03	.50	6	2.00	1.10	.44	6	NM	.33	.62	UE	D			
strength	**C 3**	4.67	1.86	-.62	6	2.00	.89	.63	6	**M**	**2.67**	.02	**E**		**Y**	**Y**	**No, but desirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item[2]**	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)[3]**	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.[6]**	**Nature of Non-Manip. Diff.[7]**	**Manip. Dimensions in Expected Direction?[8]**	**Strength Diff. in Expected Direction?[8]**	**Non-manip. Dimensions in Expected Direction?**
clarity	C 4	5.00	2.10	-1.06	6	5.17	2.32	-1.51	6	M	-.17	.87	UE				
conseq	C 4	5.00	1.10	.44	6	5.83	1.17	.36	6	NM	-.83	.21	UE	UD			
consist	C 4	5.83	.75	.74	6	4.83	.75	.74	6	NM	1.00	.05	UE	D			
constr	C 4	4.00	1.90	-.68	6	1.83	.41	.92	6	M	2.17	.02	E				
strength	**C 4**	4.83	1.94	-.76	6	4.83	.75	.74	6	**M**	**.00**	.68	**UE**		P	N	No; 1 desirable, 1 undesirable
clarity	C 5	5.67	1.51	-.06	6	1.67	.82	.69	6	M	4.00	.00	E				
conseq	C 5	5.67	1.03	.50	6	2.50	1.22	.30	6	M	3.17	.00	E				
consist	C 5	5.83	1.17	.36	6	3.00	.89	.63	6	NM	2.83	.01	UE	D			
constr	C 5	2.17	1.47	-.01	6	1.83	.75	.74	6	NM	.33	.93	UE	D			
strength	**C 5**	5.50	1.22	.30	6	1.67	.52	.88	6	**M**	**3.83**	.00	**E**		**Y**	**Y**	**No, but desirable**
clarity	C 6	5.00	2.37	-1.62	6	5.83	.75	.74	6	NM	-.83	1.00	UE	UD			
conseq	C 6	6.50	.84	.67	6	6.17	.75	.74	6	M	.33	.38	E				
consist	C 6	4.33	1.21	.31	6	3.67	1.21	.31	6	NM	.67	.36	UE	D			
constr	C 6	4.67	1.37	.13	6	3.50	.84	.67	6	NM	1.17	.12	UE	D			
strength	**C 6**	5.00	1.67	-.31	6	5.33	.52	.88	6	**M**	**-.33**	1.00	**UE**		Y	N	No, but mostly desirable

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	C 7	4.67	1.63	-.25	6	3.83	1.33	.17	6	M	.83	.41	E				
conseq	C 7	4.17	1.33	.17	6	4.00	.89	.63	6	M	.17	.60	E				
consist	C 7	3.17	.75	.74	6	3.67	1.03	.50	6	NM	-.50	.31	UE	UD			
constr	C 7	3.50	2.07	-1.01	6	3.33	1.86	-.62	6	M	.17	1.00	E				
strength	**C 7**	4.33	1.03	.50	6	3.67	1.03	.50	6	**M**	**.67**	.24	**E**		**Y**	**Y**	**No, undesirable**
clarity	C 8	5.17	1.17	.36	6	3.33	1.86	-.62	6	M	1.83	.04	E				
conseq	C 8	2.00	1.26	.25	6	3.67	1.21	.31	6	NM	-1.67	.05	UE	UD			
consist	C 8	4.00	.89	.63	6	4.17	.75	.74	6	NM	-.17	.73	UE	UD			
constr	C 8	4.83	1.47	-.01	6	3.00	1.41	.07	6	M	1.83	.06	E				
strength	**C 8**	5.17	.41	.92	6	3.33	1.63	-.25	6	**M**	**1.83**	.02	**E**		**Y**	**Y**	**No, undesirable**
clarity	C 9	5.83	.75	.74	6	3.33	1.21	.31	6	NM	2.50	.01	UE	D			
conseq	C 9	2.83	1.6	-.20	6	2.33	1.03	.50	6	NM	.50	.57	UE	D			
consist	C 9	4.33	1.86	-.62	6	4.67	1.03	.50	6	NM	-.33	.62	UE	UD			
constr	C 9	5.33	1.37	.13	6	2.83	1.17	.36	6	M	2.50	.01	E				
strength	**C 9**	5.33	1.51	-.06	6	3.50	1.38	.11	6	**M**	**1.83**	.05	**E**		**Y**	**Y**	**No, but mostly desirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item[2]**	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)[3]**	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.[6]**	**Nature of Non-Manip. Diff.[7]**	**Manip. Dimensions in Expected Direction?[8]**	**Strength Diff. in Expected Direction?[8]**	**Non-manip. Dimensions in Expected Direction?**
clarity	C 10	3.67	1.21	.31	6	4.00	1.55	-.12	6	M	-.33	.68	UE				
conseq	C 10	2.67	1.21	.31	6	3.17	1.33	.17	6	NM	-.50	.51	UE	UD			
consist	C 10	4.50	1.52	-.07	6	4.33	1.21	.31	6	NM	.17	.93	UE	D			
constr	C 10	3.33	1.21	.31	6	3.00	1.41	.07	6	M	.33	.68	E				
strength	**C 10**	2.33	.52	.88	6	3.83	1.33	.17	6	**M**	**-1.50**	.04	**UE**		P	N	No; 1 desirable, 1 undesirable
clarity	A 1	4.00	2.10	-1.06	6	2.33	1.86	-.62	6	M	1.67	.08	E				
conseq	A 1	4.50	1.38	.11	6	5.33	.82	.69	6	NM	-.83	.27	UE	UD			
consist	A 1	3.67	1.63	-.25	6	3.17	1.47	-.01	6	NM	.50	.62	UE	D			
constr	A 1	4.67	1.86	-.62	6	4.00	2.45	-1.8	6	M	.67	.68	E				
strength	**A 1**	4.50	1.52	-.07	6	2.50	1.64	-.26	6	**M**	**2.00**	.06	**E**		**Y**	**Y**	**No; 1 desirable, 1 undesirable**
clarity	A 2	4.33	1.86	-.62	6	3.17	1.94	-.76	6	M	1.17	.21	E				
conseq	A 2	5.00	1.10	.44	6	5.17	.41	.92	6	M	-.17	.92	UE				
consist	A 2	3.50	1.76	-.45	6	3.50	1.05	.49	6	NM	.00	.87	E				
constr	A 2	6.17	.75	.74	6	3.83	2.14	-1.13	6	M	.33	.04	E				
strength	**A 2**	3.83	2.04	-.95	6	3.67	1.21	.31	6	**M**	**.17**	.93	**E**		**P**	**Y**	**Yes**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	A 3	5.17	.98	.55	6	2.50	1.38	.11	6	M	2.67	.01	E				
conseq	A 3	5.50	.84	.67	6	5.67	.82	.69	6	M	-.17	.93	UE				
consist	A 3	4.33	1.21	.31	6	2.83	1.17	.36	6	NM	1.50	.07	UE	D			
constr	A 3	5.33	1.37	.13	6	5.17	1.33	.17	6	M	.17	.74	E				
strength	**A 3**	4.83	1.33	.17	6	3.67	1.86	-.62	6	**M**	**1.17**	.22	**E**		**P**	**Y**	**No, but desirable**
clarity	A 4	4.17	.98	.55	6	3.17	1.47	-.01	6	M	1.00	.18	E				
conseq	A 4	4.17	1.60	-.20	6	4.33	.52	.88	6	M[4]	-.17	.40	UE				
consist	A 4	4.33	.82	.69	6	3.83	.75	.74	6	M	.50	.27	E				
constr	A 4	4.00	2.19	-1.24	6	4.17	1.83	-.57	6	M	-.17	.87	UE				
strength	**A 4**	4.17	.98	.55	6	3.50	1.38	.11	6	**M**	**.67**	.45	**E**		**P**	**Y**	**--**
clarity	A 5	6.17	.75	.74	6	4.33	.82	.69	6	M	1.83	.01	E				
conseq	A 5	5.83	1.17	.36	6	3.50	1.05	.49	6	M	2.33	.01	E				
consist	A 5	4.33	.82	.69	6	2.50	1.05	.49	6	NM	1.83	.01	UE	D			
constr	A 5	4.17	1.83	-.57	6	3.50	1.38	.11	6	M	.67	.51	E				
strength	**A 5**	5.83	1.17	.36	6	3.33	1.03	.50	6	**M**	**2.50**	.01	**E**		**Y**	**Y**	**No, but desirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	A 6	4.00	1.26	.25	6	4.33	.82	.69	6	M	-.33	.50	UE				
conseq	A 6	4.50	1.87	-.64	6	3.50	1.87	-.64	6	NM	1.00	.33	UE	D			
consist	A 6	3.83	1.17	.36	6	3.50	.84	.67	6	NM	.33	.59	UE	D			
constr	A 6	3.33	2.16	-1.18	6	3.17	2.23	-1.32	6	M	.17	.72	E				
strength	**A 6**	3.83	1.47	-.01	6	4.17	1.33	.17	6	**M**	**-.33**	.68	**UE**		P	N	No, but desirable
clarity	A 7	6.33	.82	.69	6	4.17	2.14	-1.13	6	NM	2.17	.03	UE	D			
conseq	A 7	6.50	.84	.67	6	5.17	1.33	.17	6	M	1.33	.06	E				
consist	A 7	5.67	1.37	.13	6	4.00	2.28	-1.43	6	NM	1.67	.19	UE	D			
constr	A 7	3.33	2.25	-1.37	6	3.50	1.76	-.45	6	M	-.17	.68	UE				
strength	**A 7**	6.67	.52	.88	6	3.67	1.86	-.62	6	**M**	**3.00**	.00	**E**		**P**	**Y**	**No, but desirable**
clarity	A 8	4.00	1.67	-.31	6	5.33	1.03	.50	6	NM	-1.33	.17	UE	UD			
conseq	A 8	5.67	.52	.88	6	5.17	.41	.92	6	NM	.50	.09	E	D			
consist	A 8	2.83	1.47	-.01	6	3.83	.98	.55	6	NM	-1.00	.19	UE	UD			
constr	A 8	4.00	1.55	-.12	6	5.33	1.03	.50	6	M	-1.33	.16	UE				
strength	**A 8**	4.33	1.86	-.62	6	5.17	.75	.74	6	**M**	**-.83**	.61	**UE**		N	N	No, mostly undesirable

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item[2]**	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)[3]**	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.[6]**	**Nature of Non-Manip. Diff.[7]**	**Manip. Dimensions in Expected Direction?[8]**	**Strength Diff. in Expected Direction?[8]**	**Non-manip. Dimensions in Expected Direction?**
clarity	A 9	3.67	2.07	-.99	6	3.17	1.60	-.20	6	NM	.50	.63	UE	D			
conseq	A 9	6.17	.41	.92	6	5.67	.82	.69	6	M	.50	.18	E				
consist	A 9	2.33	1.37	.13	6	2.83	1.47	-.01	6	NM	-.50	.51	UE	UD			
constr	A 9	5.00	1.26	.25	6	4.17	1.72	-.39	6	M	.83	.32	E				
strength	**A 9**	3.17	1.17	.36	6	3.67	.52	.88	6	**M**	**-.50**	.31	**UE**		Y	N	No; 1 desirable, 1 undesirable
clarity	A 10	4.50	1.64	-.26	6	3.33	1.86	-.62	6	NM	1.17	.18	UE	D			
conseq	A 10	6.00	.63	.81	6	5.17	.41	.92	6	M	.83	.03	E				
consist	A 10	4.17	1.47	-.01	6	2.67	1.86	-.62	6	M	1.50	.14	E				
constr	A 10	4.17	.98	.55	6	4.50	.84	.67	6	NM	-.33	.53	UE	UD			
strength	**A 10**	4.67	1.03	.50	6	4.33	1.86	-.62	6	**M**	**.33**	.93	**E**		**Y**	**Y**	**No; 1 desirable, 1 undesirable**
clarity	A 11	3.67	1.63	-.25	6	3.33	1.63	-.25	6	M	.33	.74	E				
conseq	A 11	2.83	1.33	.17	6	2.40	.55	.86	5	NM	.43	.49	UE	D			
consist	A 11	3.83	.75	.74	6	3.33	.52	.88	6	NM	.50	.21	UE	D			
constr	A 11	4.33	1.21	.31	6	2.17	1.47	-.01	6	M	2.17	.02	E				
strength	**A 11**	3.17	.75	.74	6	3.00	1.10	.44	6	**M**	**.17**	.55	**E**		**Y**	**Y**	**No, but desirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	P 1	6.00	.63	.81	6	2.67	1.86	-.62	6	M	3.33	.01	E				
conseq	P 1	3.33	1.86	-.62	6	3.50	1.38	.11	6	M	-.17	.87	UE				
consist	P 1	5.33	1.37	.13	6	3.83	.75	.74	6	NM	1.50	.05	UE	D			
constr	P 1	4.50	1.38	.11	6	2.67	1.21	.31	6	NM	1.83	.05	UE	D			
strength	**P 1**	5.67	.82	.69	6	3.00	1.67	-.31	6	**M**	**2.67**	.01	**E**		**P**	**Y**	**No, but desirable**
clarity	P 2	4.00	1.10	.44	6	4.67	1.03	.50	6	NM	-.67	.30	UE	UD			
conseq	P 2	5.17	1.33	.17	6	5.33	.52	.88	6	M	-.17	.86	UE				
consist	P 2	2.83	1.33	.17	6	4.00	2.00	-.87	6	M	-1.17	.28	UE				
constr	P 2	5.50	.84	.67	6	4.50	1.38	.11	6	M	1.00	.13	E				
strength	**P 2**	3.83	1.17	.36	6	5.33	1.03	.50	6	**M**	**-1.50**	.04	**UE**		P	N	No, undesirable
clarity	P 3	5.00	1.26	.25	6	4.17	1.17	.36	6	NM	.83	.21	UE	D			
conseq	P 3	3.83	1.47	-.01	6	2.50	.84	.67	6	M	1.33	.08	E				
consist	P 3	4.67	1.37	.13	6	5.17	1.17	.36	6	NM	-.50	.62	UE	UD			
constr	P 3	4.00	1.26	.25	6	1.83	.75	.74	6	M	2.17	.01	E				
strength	**P 3**	4.50	1.22	.30	6	3.67	1.51	-.06	6	**M**	**.83**	.26	**E**		**Y**	**Y**	**No; 1 desirable, 1 undesirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	P 4	3.83	1.33	.17	6	3.67	1.97	-.81	6	M[4]	.17	.80	E				
conseq	P 4	5.17	.75	.74	6	3.83	.98	.55	6	M[5]	1.33	.04	E				
consist	P 4	3.83	1.17	.36	6	3.83	1.47	-.01	6	NM	.00	.93	E				
constr	P 4	4.33	1.51	-.06	6	2.33	1.51	-.06	6	M	2.00	.05	E				
strength	**P 4**	4.50	.55	.86	6	3.33	1.21	.31	6	**M**	**1.17**	.07	**E**		**Y**	**Y**	**Yes**
clarity	P 5	5.67	1.21	.31	6	4.50	.84	.67	6	NM	1.17	.10	UE	D			
conseq	P 5	5.17	1.72	-.39	6	3.33	.82	.69	6	M	1.83	.03	E				
consist	P 5	5.50	1.05	.49	6	4.00	1.26	.25	6	NM	1.50	.06	UE	D			
constr	P 5	5.50	1.38	.11	6	1.83	.98	.55	6	M	3.67	.01	E				
strength	**P 5**	5.83	1.60	-.20	6	3.00	1.41	.07	6	**M**	**2.83**	.02	**E**		**Y**	**Y**	**No, but desirable**
clarity	P 6	5.33	.82	.69	6	4.00	1.79	-.50	6	M[4]	1.33	.13	E				
conseq	P 6	4.50	1.52	-.07	6	5.50	.84	.67	6	M	-1.00	.20	UE				
consist	P 6	5.00	.89	.63	6	4.17	2.14	-1.13	6	M[5]	.83	.51	E				
constr	P 6	4.00	1.67	-.31	6	4.00	1.67	-.31	6	NM	.00	1.00	E				
strength	**P 6**	5.17	.75	.74	6	3.83	2.14	-1.13	6	**M**	**1.33**	.36	**E**		**M**	**Y**	**Yes**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Nop-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	P 7	5.33	1.03	.50	6	4.67	1.37	.13	6	NM	.67	.50	UE	D			
conseq	P 7	5.67	1.03	.50	6	5.33	.52	.88	6	NM	.33	.39	UE	D			
consist	P 7	4.17	1.33	.17	6	3.17	.75	.74	6	NM	1.00	.13	UE	D			
constr	P 7	5.00	1.55	-.12	6	4.83	1.17	.36	6	M	.17	.61	E				
strength	**P 7**	4.83	.98	.55	6	4.33	1.21	.31	6	**M**	**.50**	.34	**E**		**Y**	**Y**	**No, but desirable**
clarity	P 8	4.33	1.51	-.06	6	3.33	1.63	-.25	6	M	1.00	.28	E				
conseq	P 8	3.17	1.60	-.20	6	3.67	1.03	.50	6	NM	-.50	.36	UE	UD			
consist	P 8	3.83	.75	.74	6	4.33	1.21	.31	6	NM	-.50	.45	UE	UD			
constr	P 8	3.17	.98	.55	6	3.17	1.60	-.20	6	M	.00	.93	UE				
strength	**P 8**	4.17	1.17	.36	6	3.83	1.94	-.76	6	**M**	**.33**	.74	**E**		**P**	**Y**	**No, undesirable**
clarity	P 9	2.67	.82	.69	6	4.67	1.51	-.06	6	NM	-2.00	.03	UE	U			
conseq	P 9	5.33	1.75	-.43	6	6.17	.75	.74	6	M	-.83	.39	UE				
consist	P 9	4.42	.92	.61	6	4.33	1.97	-.81	6	NM	.08	.68	UE	D			
constr	P 9	5.67	1.86	-.62	6	4.50	1.87	-.64	6	M	1.17	.25	E				
strength	**P 9**	3.33	1.21	.31	6	4.83	1.72	-.39	6	**M**	**-1.50**	.10	**UE**		P	N	No; 1 desirable, 1 undesirable

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	P 10	4.33	1.21	.31	6	4.83	1.17	.36	6	NM	-.50	.46	UE	UD			
conseq	P 10	5.33	1.21	.31	6	5.67	1.21	.31	6	NM	-.33	.74	UE	UD			
consist	P 10	4.50	1.38	.11	6	5.17	1.17	.36	6	NM	-.67	.56	UE	UD			
constr	P 10	4.50	1.38	.11	6	3.00	1.90	-.68	6	M	1.50	.13	E				
strength	**P 10**	4.17	1.17	.36	6	5.17	1.33	.17	6	**M**	**-1.00**	.19	**UE**		Y	N	No, undesirable
clarity	E 1	4.83	1.94	-.76	6	3.33	1.37	.13	6	M	1.50	.14	E				
conseq	E 1	5.00	1.10	.44	6	4.83	.98	.55	6	NM	.17	.65	UE	D			
consist	E 1	3.50	1.64	-.26	6	3.17	1.94	-.76	6	M	.33	.62	E				
constr	E 1	3.00	1.79	-.50	6	3.67	1.63	-.25	6	NM	-.67	.46	UE	UD			
strength	**E 1**	4.50	1.38	.11	6	4.00	1.41	.07	6	**M**	**.50**	.57	**E**		**Y**	**Y**	**No; 1 desirable, 1 undesirable**
clarity	E 2	4.50	1.64	-.26	6	5.00	1.41	.07	6	NM	-.50	.68	UE	UD			
conseq	E 2	5.17	1.47	-.01	6	4.33	2.25	-1.37	6	M	.83	.56	E				
consist	E 2	3.50	1.76	-.45	6	4.83	1.17	.36	6	M	-1.33	.16	UE				
constr	E 2	3.83	1.60	-.20	6	1.83	.75	.74	6	M	2.00	.04	E				
strength	**E 2**	3.33	1.86	-.62	6	4.00	1.55	-.12	6	**M**	**-.67**	.51	**UE**		M	N	No, undesirable

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	E 3	4.50	1.05	.49	6	4.50	1.22	.30	6	M	.00	.93	UE				
conseq	E 3	5.50	1.52	-.07	6	5.50	1.05	.49	6	M	.00	.87	UE				
consist	E 3	4.33	1.37	.13	6	4.33	1.03	.50	6	M	.00	.87	UE				
constr	E 3	3.50	1.38	.11	6	3.00	1.26	.25	6	M	.50	.50	E				
strength	**E 3**	4.17	1.17	.36	6	4.00	.89	.63	6	**M**	**.17**	.87	**E**		MN	Y	--
clarity	E 4	5.33	.82	.69	6	4.17	1.47	-.01	6	NM	1.17	.13	UE	D			
conseq	E 4	5.33	1.03	.50	6	5.17	1.33	.17	6	M	.17	.93	E				
consist	E 4	4.17	1.17	.36	6	3.17	1.72	-.39	6	M	1.00	.33	E				
constr	E 4	2.83	1.33	.17	6	2.17	1.94	-.76	6	M	.67	.17	E				
strength	**E 4**	4.50	1.05	.49	6	4.33	1.63	-.25	6	**M**	**.17**	.93	**E**		**Y**	**Y**	**No, but desirable**
clarity	E 5	3.50	1.22	.30	6	4.50	1.64	-.26	6	NM	-1.00	.24	UE	UD			
conseq	E 5	2.67	1.86	-.62	6	4.33	.82	.69	6	NM	-1.67	.07	UE	UD			
consist	E 5	4.00	.63	.81	6	3.83	1.47	-.01	6	NM	.17	.68	UE	D			
constr	E 5	1.67	.52	.88	6	1.33	.82	.69	6	M	.33	.20	E				
strength	**E 5**	3.00	.63	.81	6	4.00	1.67	-.31	6	**M**	**-1.00**	.28	**UE**		Y	N	No, mostly undesirable

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item**[2]	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)**[3]	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.**[6]	**Nature of Non-Manip. Diff.**[7]	**Manip. Dimensions in Expected Direction?**[8]	**Strength Diff. in Expected Direction?**[8]	**Non-manip. Dimensions in Expected Direction?**
clarity	E 6	5.67	.52	.88	6	4.67	1.37	.13	6	NM	1.00	.16	UE	D			
conseq	E 6	6.50	.55	.86	6	5.83	.41	.92	6	M	.67	.04	E				
consist	E 6	4.33	1.37	.13	6	3.33	1.37	.13	6	NM	1.00	.19	UE	D			
constr	E 6	4.17	1.72	-.39	6	3.33	1.75	-.43	6	M	.83	.50	E				
strength	**E 6**	5.67	.52	.88	6	4.17	1.60	-.20	6	**M**	**1.50**	.07	**E**		**Y**	**Y**	**No, but desirable**
clarity	E 7	4.83	1.47	-.01	6	3.83	1.83	-.57	6	M	1.00	.31	E				
conseq	E 7	3.50	1.87	-.64	6	4.17	1.47	-.01	6	M	-.67	.52	UE				
consist	E 7	3.00	1.67	-.31	6	3.83	1.17	.36	6	NM	-.83	.25	UE	UD			
constr	E 7	2.83	.98	.55	6	4.33	1.21	.31	6	M	-1.50	.05	UE				
strength	**E 7**	4.17	1.33	.17	6	4.50	1.87	-.64	6	**M**	**-.33**	.51	**UE**		MN	N	No, undesirable
clarity	E 8	4.17	1.72	-.39	6	2.50	.84	.67	6	NM	1.67	.06	UE	D			
conseq	E 8	6.33	.52	.88	6	5.17	1.83	-.57	6	M	1.17	.18	E				
consist	E 8	4.33	1.21	.31	6	2.50	.84	.67	6	NM	1.83	.02	UE	D			
constr	E 8	5.67	1.51	-.06	6	3.67	1.97	-.81	6	M	2.00	.06	E				
strength	**E 8**	4.17	2.04	-.95	6	2.67	1.03	.50	6	**M**	**1.50**	.19	**E**		**Y**	**Y**	**No, but desirable**

Table 1 (cont'd)

		Strong				Weak											
Dim[1]	**Item[2]**	**M**	**SD**	**r_{wg}**	**N**	**M**	**SD**	**r_{wg}**	**N**	**Strength Facet (M/NM)[3]**	**Mean Diff.**	**Mann Whitney U sig.**	**Nature of Diff.[6]**	**Nature of Non-Manip. Diff.[7]**	**Manip. Dimensions in Expected Direction?[8]**	**Strength Diff. in Expected Direction?[8]**	**Non-manip. Dimensions in Expected Direction?**
clarity	E 9	5.17	1.60	-.20	6	4.17	1.60	-.20	6	NM	1.00	.06	UE	D			
conseq	E 9	6.33	.82	.69	6	5.67	1.86	-.62	6	M	.67	.60	E				
consist	E 9	5.00	.89	.63	6	4.83	.75	.74	6	NM	.17	.73	UE	D			
constr	E 9	5.50	1.76	-.45	6	3.83	2.23	-1.32	6	M	1.67	.14	E				
strength	**E 9**	4.67	2.07	-.99	6	4.20	2.17	-1.20	5	**M**	**.47**	.55	**E**		**Y**	**Y**	**No, but desirable**
clarity	E 10	4.67	1.75	-.43	6	4.50	1.64	-.26	6	M	.17	.74	E				
conseq	E 10	5.50	1.52	-.07	6	4.83	1.60	-.20	6	M	.67	.46	E				
consist	E 10	2.40	1.14	.39	5	4.50	1.38	.11	6	NM	-2.10	.03	UE	UD			
constr	E 10	4.00	1.79	-.50	6	3.67	1.21	.31	6	M	.33	.74	E				
strength	**E 10**	3.67	1.97	-.81	6	4.50	1.64	-.26	6	**M**	**-.83**	.41	**UE**		Y	N	No, undesirable

Note. [1]Dimensions (conseq = consequences, consist = consistency, constr = constraints). [2]Item (C = Citizenship, A = Adaptability, P = Perseverance, E = Ethics). [3]Strength facet manipulation (M = Manipulated, NM = Non-manipulated). [4]Situational strength dimension is represented in strong version of stem but not in weak. [5]Situational strength dimension is represented in weak version of stem but not in strong. [6]The direction of the mean difference was either "expected" (abbreviated "E") or "unexpected" (abbreviated "UE"). [7]In cases where the difference was unexpected due to a difference between stem versions along a non-manipulated dimension, the difference was "desirable" (abbreviated "D") when the stronger version of the stem was perceived as higher on the non-manipulated dimension and "undesirable" (abbreviated "UD") when the stronger version of the stem was perceived as lower on the non-manipulated dimension. [8]Y = Yes, N = No, P = Partially, M = Mostly, MN = Mostly Not.

Table 2. **Ratings of SJT stem pairs**

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	C 1	.00	.82	4	.69		NM	B	expected				
conseq	C 1	1.50	1.0	4	.53		M	B	expected				
consist	C 1	-.75	1.5	4	-.05		NM	B	unexpected	UD			
constr	C 1	.50	1.0	4	.53		NM	B	unexpected	D			
strength	**C 1**	.25	.96	4	.57	.45	**M**	B	**expected**		Y	Y	Mostly not; 1 desirable, 1 undesirable
clarity	C 2	-1.00	1.4	4	.07		NM	A	unexpected	D			
conseq	C 2	-1.25	.96	4	.57		M	A	expected				
consist	C 2	-1.00	1.6	4	-.25		NM	A	unexpected	D			
constr	C 2	-1.50	1.0	4	.53		M	A	expected				
strength	**C 2**	-1.00	1.4	4	.07	.20	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	C 3	-1.00	1.8	4	-.56		M	A	expected				
conseq	C 3	-1.25	.96	4	.57		M	A	expected				
consist	C 3	-.50	1.2	4	.22		NM	A	unexpected	D			
constr	C 3	-.75	.50	4	.88		NM	A	unexpected	D			
strength	**C 3**	-.75	1.2	4	.26	.28	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	C 4	.25	1.5	4	-.05		M	B	expected				
conseq	C 4	1.25	.96	4	.57		NM	B	unexpected	D			
consist	C 4	1.00	.82	4	.69		NM	B	unexpected	D			
constr	C 4	.25	.50	4	.88		M	B	expected				
strength	**C 4**	.50	1.2	4	.22	.46	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	C 5	-2.25	.96	4	.57		M	A	expected				
conseq	C 5	-1.75	.96	4	.57		M	A	expected				
consist	C 5	-2.00	.82	4	.69		NM	A	unexpected	D			
constr	C 5	-1.00	1.15	4	.38		NM	A	unexpected	D			
strength	**C 5**	-2.00	.82	4	.69	.58	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	C 6	1.00	.82	4	.69		NM	B	unexpected	D			
conseq	C 6	2.00	.82	4	.69		M	B	expected				
consist	C 6	1.00	.82	4	.69		NM	B	unexpected	D			
constr	C 6	.25	.50	4	.88		NM	B	unexpected	D			
strength	**C 6**	1.00	.82	4	.69	.73	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	C 7	.75	.50	4	.88		M	B	expected				
conseq	C 7	.75	.96	4	.57		M	B	expected				
consist	C 7	-.25	.96	4	.57		NM	B	unexpected	UD			
constr	C 7	1.00	1.00	3	.53		M	B	expected				
strength	**C 7**	.75	.50	4	.88	.69	**M**	B	**expected**		Y	Y	No, undesirable
clarity	C 8	-1.75	.50	4	.88		M	A	expected				
conseq	C 8	-.75	.96	4	.57		NM	A	unexpected	D			
consist	C 8	-1.25	.96	4	.57		NM	A	unexpected	D			
constr	C 8	-1.00	.00	4	1.00		M	A	expected				
strength	**C 8**	-1.00	.82	4	.69	.74	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	C 9	.50	.58	4	.84		NM	B	unexpected	D			
conseq	C 9	.50	.58	4	.84		NM	B	unexpected	D			
consist	C 9	.25	.50	4	.88		NM	B	unexpected	D			
constr	C 9	1.75	.50	4	.88		M	B	expected				
strength	**C 9**	.75	.50	4	.88	.87	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	C 10	-.25	.50	4	.88		M	A	expected				
conseq	C 10	-1.00	.82	4	.69		NM	A	unexpected	D			
consist	C 10	.25	.96	4	.57		NM	A	unexpected	UD			
constr	C 10	-1.75	.50	4	.88		M	A	expected				
strength	**C 10**	.00	.82	4	.69	.74	**M**	A	**unexpected**		Y	N	No; 1 desirable, 1 undesirable
clarity	A 1	.00	1.41	4	.07		M	A	unexpected				
conseq	A 1	-1.75	.50	4	.88		NM	A	unexpected	D			
consist	A 1	.25	1.71	4	-.36		NM	A	unexpected	UD			
constr	A 1	-1.75	.50	4	.88		M	A	expected				
strength	**A 1**	-.50	1.73	4	-.40	.21	**M**	A	**expected**		P	Y	No; 1 desirable, 1 undesirable
clarity	A 2	.75	1.26	4	.26		M	B	expected				
conseq	A 2	.50	1.29	4	.22		M	B	expected				
consist	A 2	.50	1.29	4	.22		NM	B	unexpected	D			
constr	A 2	.50	1.91	4	-.71		M	B	expected				
strength	**A 2**	.00	1.15	4	.38	.07	**M**	B	**unexpected**		Y	N	No, but desirable

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	A 3	-1.25	.50	4	.88		M	A	expected				
conseq	A 3	-.50	.58	4	.84		M	A	expected				
consist	A 3	-.75	1.26	4	.26		NM	A	unexpected	D			
constr	A 3	-1.25	.96	4	.57		M	A	expected				
strength	**A 3**	-.50	1.29	4	.22	.56	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	A 4	-2.25	.50	4	.88		M	A	expected				
conseq	A 4	-2.00	.00	3	1.00		M[6]	A	expected				
consist	A 4	-1.50	.58	4	.84		M	A	expected				
constr	A 4	-1.25	.50	4	.88		M	A	expected				
strength	**A 4**	-1.25	.50	4	.88	.90	**M**	A	**expected**		**Y**	**Y**	**--**
clarity	A 5	-.50	.58	4	.84		M	A	expected				
conseq	A 5	-1.75	.50	4	.88		M	A	expected				
consist	A 5	-.25	.96	4	.57		NM	A	unexpected	D			
constr	A 5	-.75	.96	4	.57		M	A	expected				
strength	**A 5**	-1.00	.82	4	.69	.71	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	A 6	1.75	.50	4	.88		M	B	expected				
conseq	A 6	2.00	.00	4	1.00		NM	B	unexpected	D			
consist	A 6	1.50	.58	4	.84		NM	B	unexpected	D			
constr	A 6	1.50	.58	4	.84		M	B	expected				
strength	**A 6**	1.75	.50	4	.88	.89	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	A 7	1.25	1.26	4	.26		NM	B	unexpected	D			
conseq	A 7	2.75	.50	4	.88		M	B	expected				
consist	A 7	.25	.50	4	.88		NM	B	unexpected	D			
constr	A 7	1.50	1.29	4	.22		M	B	expected				
strength	**A 7**	1.50	.58	4	.84	.62	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	A 8	-.25	.50	4	.88		NM	A	unexpected	D			
conseq	A 8	-.50	1.00	4	.53		NM	A	unexpected	D			
consist	A 8	.00	.00	4	1.00		NM	A	expected				
constr	A 8	-1.33	.58	3	.84		M	A	expected				
strength	**A 8**	-1.00	.00	4	1.00	.85	**M**	A	**expected**		**Y**	**Y**	**Mostly not, but desirable**
clarity	A 9	-.25	.50	4	.88		NM	A	unexpected	D			
conseq	A 9	-1.25	.50	4	.88		M	A	expected				
consist	A 9	.25	.50	4	.88		NM	A	unexpected	UD			
constr	A 9	-1.50	1.00	4	.53		M	A	expected				
strength	**A 9**	-1.00	.00	4	1.00	.84	**M**	A	**expected**		Y	Y	No; 1 desirable, 1 undesirable
clarity	A 10	-.25	1.26	4	.26		NM	B	unexpected	UD			
conseq	A 10	1.50	1.29	4	.22		M	B	expected				
consist	A 10	-.50	1.29	4	.22		M	B	unexpected				
constr	A 10	1.25	.96	4	.57		NM	B	unexpected	D			
strength	**A 10**	.50	1.29	4	.22	.30	**M**	B	**expected**		P	Y	No; 1 desirable, 1 undesirable

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	A 11	.75	.96	4	.57		M	B	expected				
conseq	A 11	.75	.50	4	.88		NM	B	unexpected	D			
consist	A 11	.00	1.41	4	.07		NM	B	expected				
constr	A 11	.50	1.00	4	.53		M	B	expected				
strength	**A 11**	.00	.82	4	.69	.55	**M**	B	**unexpected**		Y	N	Partially, but desirable
clarity	P 1	-.50	1.29	4	.22		M	A	expected				
conseq	P 1	-2.25	.50	4	.88		M	A	expected				
consist	P 1	-.25	1.26	4	.26		NM	A	unexpected	D			
constr	P 1	-.50	1.00	4	.53		NM	A	unexpected	D			
strength	**P 1**	-1.00	.82	4	.69	.52	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	P 2	-.50	.58	4	.84		NM	B	unexpected	UD			
conseq	P 2	1.00	.82	4	.69		M	B	expected				
consist	P 2	-.50	1.29	4	.22		M	B	unexpected				
constr	P 2	1.50	1.29	4	.22		M	B	expected				
strength	**P 2**	-.25	1.26	4	.26	.45	**M**	B	**unexpected**		M	N	No, undesirable
clarity	P 3	1.00	1.41	4	.07		NM	B	unexpected	D			
conseq	P 3	1.25	.96	4	.57		M	B	expected				
consist	P 3	.75	1.50	4	-.05		NM	B	unexpected	D			
constr	P 3	1.00	.82	4	.69		M	B	expected				
strength	**P 3**	1.25	.96	4	.57	.37	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	P 4	-.25	1.50	4	-.05		M[6]	A	expected				
conseq	P 4	-.75	.96	4	.57		M[7]	A	expected				
consist	P 4	.50	1.29	4	.22		NM	A	unexpected	UD			
constr	P 4	-1.25	.50	4	.88		M	A	expected				
strength	**P 4**	.25	.96	4	.57	.44	**M**	A	**unexpected**		Y	N	No, undesirable
clarity	P 5	-1.25	.96	4	.57		NM	A	unexpected	D			
conseq	P 5	-1.75	.96	4	.57		M	A	expected				
consist	P 5	-1.25	1.50	4	-.05		NM	A	unexpected	D			
constr	P 5	-1.75	.96	4	.57		M	A	expected				
strength	**P 5**	-1.25	.96	4	.57	.45	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**
clarity	P 6	.75	.96	4	.57		M[6]	B	expected				
conseq	P 6	.50	1.00	4	.53		M	B	expected				
consist	P 6	.75	.96	4	.57		M[7]	B	expected				
constr	P 6	.25	.96	4	.57		NM	B	unexpected	D			
strength	**P 6**	1.00	.82	4	.69	.59	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	P 7	-1.00	1.15	4	.38		NM	A	unexpected	D			
conseq	P 7	-1.75	.50	4	.88		NM	A	unexpected	D			
consist	P 7	-1.25	.50	4	.88		NM	A	unexpected	D			
constr	P 7	-1.50	.58	4	.84		M	A	expected				
strength	**P 7**	-1.50	.58	4	.84	.77	**M**	A	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	P 8	1.25	.50	4	.88		M	B	expected				
conseq	P 8	1.00	.00	4	1.00		NM	B	unexpected	D			
consist	P 8	1.00	.82	4	.69		NM	B	unexpected	D			
constr	P 8	.50	1.29	4	.22		M	B	expected				
strength	**P 8**	1.00	.82	4	.69	.70	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	P 9	1.00	1.15	4	.38		NM	B	unexpected	D			
conseq	P 9	1.75	1.26	4	.26		M	B	expected				
consist	P 9	1.00	1.15	4	.38		NM	B	unexpected	D			
constr	P 9	1.50	1.00	4	.53		M	B	expected				
strength	**P 9**	1.25	.96	4	.57	.42	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	P 10	.25	.50	4	.88		NM	A	unexpected	UD			
conseq	P 10	-.75	.50	4	.88		NM	A	unexpected	D			
consist	P 10	-.25	.50	4	.88		NM	A	unexpected	D			
constr	P 10	-1.00	.00	4	1.00		M	A	expected				
strength	**P 10**	.25	.96	4	.57	.84	**M**	A	**unexpected**		Y	N	No, mostly desirable
clarity	E 1	1.50	1.29	4	.22		M	B	expected				
conseq	E 1	1.50	1.29	4	.22		NM	B	unexpected	D			
consist	E 1	1.25	.96	4	.57		M	B	expected				
constr	E 1	.25	.50	4	.88		NM	B	unexpected	D			
strength	**E 1**	1.50	.58	4	.84	.55	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	E 2	.00	.82	4	.69		NM	B	expected				
conseq	E 2	2.25	.96	4	.57		M	B	expected				
consist	E 2	.25	.50	4	.88		M	B	expected				
constr	E 2	1.25	1.26	4	.26		M	B	expected				
strength	**E 2**	.50	1.29	4	.22	.52	**M**	B	**expected**		**Y**	**Y**	**Yes**
clarity	E 3	-.50	1.00	4	.53		M	A	expected				
conseq	E 3	-1.75	1.26	4	.26		M	A	expected				
consist	E 3	-.50	1.73	4	-.40		M	A	expected				
constr	E 3	-1.75	1.26	4	.26		M	A	expected				
strength	**E 3**	-.25	1.71	4	-.36	.06	**M**	A	**expected**		**Y**	**Y**	**Yes**
clarity	E 4	.50	.58	4	.84		NM	B	unexpected	D			
conseq	E 4	2.00	.82	4	.69		M	B	expected				
consist	E 4	.25	.96	4	.57		M	B	expected				
constr	E 4	1.25	.96	4	.57		M	B	expected				
strength	**E 4**	.50	1.00	4	.53	.64	**M**	B	**expected**		**Y**	**Y**	**No, but desirable**
clarity	E 5	.00	.82	4	.69		NM	A	expected				
conseq	E 5	-.75	.96	4	.57		NM	A	unexpected	D			
consist	E 5	.25	.50	4	.88		NM	A	unexpected	UD			
constr	E 5	-1.25	.50	4	.88		M	A	expected				
strength	**E 5**	.25	1.26	4	.26	.66	**M**	A	**unexpected**		Y	N	Partially; 1 desirable, 1 undesirable

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	E 6	.75	.96	4	.57		NM	B	unexpected	D			
conseq	E 6	.33	1.53	3	-.09		M	B	expected				
consist	E 6	.25	1.26	4	.26		NM	B	unexpected	D			
constr	E 6	-.25	1.26	4	.26		M	B	unexpected				
strength	**E 6**	.75	1.50	4	-.05	.19	**M**	B	**expected**		P	Y	No, but desirable
clarity	E 7	.75	.96	4	.57		M	B	expected				
conseq	E 7	.75	.96	4	.57		M	B	expected				
consist	E 7	.25	.96	4	.57		NM	B	unexpected	D			
constr	E 7	-.25	.50	4	.88		M	B	unexpected				
strength	**E 7**	.75	.96	4	.57	.63	**M**	B	**expected**		M	Y	No, but desirable
clarity	E 8	-.50	.58	4	.84		NM	A	unexpected	D			
conseq	E 8	-2.25	.96	4	.57		M	A	expected				
consist	E 8	.25	.50	4	.88		NM	A	unexpected	UD			
constr	E 8	-1.25	.50	4	.88		M	A	expected				
strength	**E 8**	-.25	1.50	4	-.05	.63	**M**	A	**expected**		Y	Y	No; 1 desirable, 1 undesirable
clarity	E 9	.25	.50	4	.88		NM	B	unexpected	D			
conseq	E 9	1.25	.50	4	.88		M	B	expected				
consist	E 9	-.50	.58	4	.84		NM	B	unexpected	UD			
constr	E 9	.50	.58	4	.84		M	B	expected				
strength	**E 9**	.00	.82	4	.69	.83	**M**	B	**unexpected**		Y	N	No; 1 desirable, 1 undesirable

Table 2 (cont'd)

Dim[1]	Item[2]	M[3]	SD	N	r_{wg}	Avg Item r_{wg}[4]	Strength Facet (M/NM)[5]	Which Stem (A or B) was Strong?	Nature of Diff.	Nature of Non-Manipulated Difference[8]	Manipulated Dimensions in Expected Direction?[9]	Strength Difference in Expected Direction?[9]	Non-manipulated Dimensions in Expected Direction?
clarity	E 10	-.75	.50	4	.88		M	A	expected				
conseq	E 10	-1.25	.96	4	.57		M	A	expected				
consist	E 10	.00	.82	4	.69		NM	A	expected				
constr	E 10	-1.25	.50	4	.88		M	A	expected				
strength	**E 10**	-.50	1.00	4	.53	.71	**M**	A	**expected**		**Y**	**Y**	**Yes**

Note. [1]Dimensions (conseq = consequences, consist = consistency, constr = constraints). [2]Item (C = Citizenship, A = Adaptability, P = Perseverance, E = Ethics). [3]SMEs' ratings were coded from -3 ("A Extremely Higher") to +3 ("B Extremely Higher") such that a rating of "A and B Equal" corresponded to a value of 0. [4]Average r_{wg} value for the item (across ratings of clarity, consequences, consistency, constraints, and global strength). [5]Strength facet manipulation (M = Manipulated, NM = Non-manipulated). [6]Situational strength dimension is represented in strong version of stem but not in weak. [7]Situational strength dimension is represented in weak version of stem but not in strong. [8]In cases where the difference was unexpected due to a difference between stem versions along a non-manipulated dimension, the difference was "desirable" (abbreviated "D") when the stronger version of the stem was perceived as higher on the non-manipulated dimension and "undesirable" (abbreviated "UD") when the stronger version of the stem was perceived as lower on the non-manipulated dimension. [9]Y = Yes, N = No, P = Partially, M = Mostly.

Table 3. **SJT stems kept based on SMEs' ratings of strength**

	Task Version								
	Independent Ratings of Weak/Strong Item Stems				Ratings of Stem Pairs				
Item[1]	Manipulated Dimensions in Expected Direction?	Strength Difference in Expected Direction?	Non-manipulated Dimensions in Expected Direction?[3]	Keep?	Manipulated Dimensions in Expected Direction?	Strength Difference in Expected Direction?	Non-manipulated Dimensions in Expected Direction?[3]	Keep?	Avg Item r_{wg}[2]
C 3	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Maybe	.28
C 5	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.58
C 7	Yes	Yes	No, undesirable	Yes	Yes	Yes	No, undesirable	Maybe	.69
C 8	Yes	Yes	No, undesirable	Yes	Yes	Yes	No, but desirable	Yes	.74
C 9	Yes	Yes	No, but mostly desirable	Yes	Yes	Yes	No, but desirable	Yes	.87
A 3	Partially	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Maybe	.56
A 4	Partially	Yes	--	Yes	Yes	Yes	--	Yes	.90
A 5	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.71
A 7	Partially	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.62
P 1	Partially	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.52
P 3	Yes	Yes	No; 1 desirable, 1 undesirable	Yes	Yes	Yes	No, but desirable	Yes	.37
P 5	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.45
P 6	Mostly	Yes	Yes	Yes	Yes	Yes	No, but desirable	Yes	.59
P 7	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.77
P 8	Partially	Yes	No, undesirable	Yes	Yes	Yes	No, but desirable	Yes	.70
E 1	Yes	Yes	No; 1 desirable, 1 undesirable	Yes	Yes	Yes	No, but desirable	Yes	.55

Table 3 (cont'd)

	Task Version								
	Independent Ratings of Weak/Strong Item Stems				Ratings of Stem Pairs				
Item[1]	Manipulated Dimensions in Expected Direction?	Strength Difference in Expected Direction?	Non-manipulated Dimensions in Expected Direction?[3]	Keep?	Manipulated Dimensions in Expected Direction?	Strength Difference in Expected Direction?	Non-manipulated Dimensions in Expected Direction?[3]	Keep?	Avg Item r_{wg}[2]
E 4	Yes	Yes	No, but desirable	Yes	Yes	Yes	No, but desirable	Yes	.64
E 8	Yes	Yes	No, but desirable	Yes	Yes	Yes	No; 1 desirable, 1 undesirable	Maybe	.63

Note. [1]Item (C = Citizenship, A = Adaptability, P = Perseverance, E = Ethics). [2]Average r_{wg} value for the item (across ratings of clarity, consequences, consistency, constraints, and global strength). [3]In cases where the difference was unexpected due to a difference between stem versions along a non-manipulated dimension, the difference was "desirable" when the stronger version of the stem was perceived as higher on the non-manipulated dimension and "undesirable" when the stronger version of the stem was perceived as lower on the non-manipulated dimension.

Table 4. **Rater agreement by strength dimension for final set of 18 SJT items**

	Ratings of Weak Item Stems		**Ratings of Strong Item Stems**		**Ratings of Stem Pairs**	
Strength Dimension	***M*** **r_{wg}**	***SD*** **r_{wg}**	***M*** **r_{wg}**	***SD*** **r_{wg}**	***M*** **r_{wg}**	***SD*** **r_{wg}**
Clarity	.04	.50	.23	.51	.56	.40
Consequences	.37	.42	.25	.44	.70	.21
Consistency	.24	.66	.32	.38	.55	.30
Constraints	-.08	.58	-.19	.54	.65	.23
Strength	-.02	.54	.28	.49	.62	.16

Scoring Key Development

The scoring keys for the two versions of the SJT are shown in Tables 5 (scoring for weak version of the SJT) and 6 (scoring for strong version of SJT). Scoring keys were determined based on the effectiveness ratings and best/worst responses provided by college juniors and seniors. See Appendix E for the item keying procedure. Survey items 6 (Adaptability 3), 10 (Perseverance 1), 11 (Perseverance 3), 14 (Perseverance 7), 16 (Ethics 1), 17 (Ethics 4), and 18 (Ethics 8) were the only ones that had all response options keyed the same way across the two versions of the SJT.

Table 5. **Scoring for weak version of the SJT**

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 3. A student on your floor is always organizing "social" activities, including trips to the local bars. The individual is drinking three or more drinks at least three times a week. He/she is still attending classes and completing assignments but the situation could get worse. You do not know whether anyone else is aware or concerned about the person. This individual is not a close friend. What should you do?					
1. Talk to the person about easing up on the alcohol, explaining that it will not help with his/her classes, which should be the main reason why he/she is in college.	2.97	1.22	16.95	18.64	0
2. Use humor to broach the topic and offer alternatives to the individual's usual "social" activities.	3.19	1.09	11.86	6.78	1
3. Bring up the situation with the floor's resident assistant.	2.80	1.32	11.86	25.42	0
4. Try to get the individual involved in other activities.	3.76	.90	20.34	1.69	1
5. Talk to the person to subtly determine if there are other issues that need to be addressed, and refer him/her to help if appropriate.	3.44	1.15	22.03	1.69	1
6. Talk to other people on the floor, and discuss ways to address the situation.	2.47	1.22	5.08	32.20	0
7. Ask the individual once about this behavior and see where the discussion leads, then leave the person to his/her own course of action.	3.00	1.17	11.86	13.56	0
Citizenship 5. You hear about a situation in which your acquaintance was mistreated, but it is not likely to happen to others. Several individuals know about the situation and may or may not have done something about it. You are not aware of any University policies regarding such situations. What should you do?					
1. Do whatever you can to keep it from happening again.	3.76	.82	.00	5.08	0
2. Help the person who was mistreated any way you can.	4.29	.62	11.86	.00	0
3. Inform the appropriate authorities.	4.03	.90	15.25	3.39	0
4. Punish the person who mistreated the individual.	2.00	1.03	1.69	83.05	-1
5. Try to identify the causes of the event and correct them.	3.21	1.06	5.08	8.47	0
6. Talk to the person who was mistreated to see if he/she wants to do anything about the incident and let him/her know you are willing to help.	4.48	.90	66.10	.00	1

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 7. Your roommate is having problems studying for an exam in a class that you took last semester. You had done very well in the class. You have finished your assignments for the night and were planning on going out to dinner with a couple of friends. Your roommate asks you for help in studying for the exam. She tells you she is having trouble understanding several concepts that could be on the test. She has another friend she could ask to help her study. You don't know how much time helping her would take. What should you do?					
1. Explain that you already made plans, but that when you get home from dinner you can review some material with her.	4.25	.76	61.02	.00	1
2. Call your friends and cancel dinner. Stay home to help your roommate and have dinner with her instead.	3.07	1.21	15.25	3.39	0
3. Delay your dinner plans for an hour. Spend the time helping your roommate and then go out to dinner.	3.78	.97	20.34	.00	0
4. Explain that you would have liked to help her, but you already have plans that cannot be broken.	2.31	1.04	1.69	6.78	0
5. Invite the roommate to go to dinner with you. You can help her on your drive to and from dinner.	1.92	1.04	1.69	22.03	0
6. Tell her that you would like to help, but convince her that you would be of no real help to her.	1.48	.73	.00	67.80	-1
Citizenship 8. At the beginning of the semester, your professor asks for a volunteer to type out the lecture notes after each lecture for a hearing impaired student. The task would require you to type out the lecture notes sometime before the following class (several days later) and e-mail them to the student so he has time to review them. The position is not paid. You are not sure how long and detailed lecture notes will tend to be. You also consider that your workload might increase later in the semester. What should you do?					
1. Volunteer to type the notes. Do so the entire semester.	3.00	1.02	5.08	33.90	0
2. Volunteer to type the notes. If you realize half-way through the semester that you can't do it anymore, ask the professor to find someone else.	3.63	1.00	20.34	11.86	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
3. Don't volunteer. You realize that it will take up too much time and someone else can probably do it better.	3.10	1.08	10.17	35.59	0
4. Volunteer to type half the notes if the professor can find someone to share the duty with you. If the professor can't find anyone else you wouldn't be able to do it.	3.68	1.04	32.20	15.25	0
5. Volunteer to type the notes. Ask a friend in class to share the responsibility with you so that you can take turns.	3.97	1.05	32.20	3.39	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 9. Your professor announces in class that undergraduate students are needed starting next week to help run subjects for his research study. While you would not receive any extra credit, the professor would appreciate any volunteers. This semester you are taking several classes but your workload has been manageable. Your midterms are still three weeks away. What should you do?					
1. Examine your schedule and offer to volunteer a couple of hours a week when it is personally convenient.	4.36	.58	67.80	1.69	1
2. Examine your schedule and offer to volunteer as many hours as you can.	3.69	.98	22.03	3.39	0
3. Wait to see how many other people volunteer; agree to help a couple hours a week only if no one else volunteers.	2.97	1.06	1.69	.00	0
4. Volunteer only if a friend of yours volunteers and asks you to do it as well; volunteer for as many hours as your friend.	2.31	.99	.00	11.86	0
5. Realize that you would have to give up some of your free time and choose not to volunteer.	2.41	1.05	6.78	11.86	0
6. Offer to run subjects only if you are paid.	1.73	1.10	1.69	71.19	-1
Adaptability 3. Because of family problems, you find out that your parents can no longer support you financially at the same level as they have. You do not have enough money to continue in school for two more years. You don't know if your parents really want you to finish college— it never seemed to bother them that your brother did not finish college. You might find a job that does not require a college degree. What plans should you make?					
1. Apply for student loans or get a part-time job.	4.73	.45	89.66	.00	1
2. Ask other, more wealthy family members for money to finish school.	3.10	.96	.00	8.62	0
3. Drop out of school and save money for going back.	2.27	1.16	.00	27.59	0
4. Take fewer classes because of the lower level of finances.	3.31	.99	1.72	3.45	0
5. Rely on your parents to figure something out—you should be able to count on them.	1.64	.85	.00	60.34	-1
6. Transfer to a less expensive, community college.	3.59	.98	8.62	.00	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Adaptability 4. You share a dorm room with three other students. One hour before you are expecting a guest, you get home to find the place completely trashed. There is no sign of any of your roommates. You don't know how your guest would feel about the mess. You do not know the individual very well. Your roommates sometimes react negatively to criticism. What should you do?					
1. Clean up the mess as much as possible before the guest arrives. Then speak with your roommates immediately upon their return, so your guest knows how concerned you were about the mess.	2.93	1.22	10.17	25.42	0
2. Leave the mess and explain the situation to your guest.	2.15	1.01	.00	45.76	-1
3. Leave the mess and take the guest somewhere else.	3.27	1.11	3.39	6.78	0
4. Clean up the mess as much as possible before the guest arrives. Then, without the guest around, ask the roommates why the place was trashed so badly and what can be done in the future to avoid this situation.	4.75	.44	84.75	.00	1
5. Call the guest and ask to reschedule your meeting. When your roommates arrive, ask them to clean up the mess.	2.69	1.07	1.69	22.03	-1
Adaptability 5. You have an exam in one of your classes the next day. But this exam will not be weighted too heavily in your grade and you received high marks on the previous two exams. Two hours after you settle down to study, you realize that students in an adjacent room have organized a party. You have not actually received an invitation from them but could still join the party. You know that the noise will make studying difficult. What should you do?					
1. Find an alternative place to study, like the library.	4.71	.46	91.53	.00	1
2. Tell yourself that you have to spend the next hour studying, and then you can go to the party.	3.15	1.17	5.08	32.20	0
3. Ask them to keep quiet, or wait until you are done studying if they really want to party.	2.29	1.07	.00	55.93	-1
4. Put your headphones on and try to study anyway.	3.25	.99	.00	5.08	0
5. Put in earplugs to block out some of the noise and try to study anyway.	3.41	.93	3.39	6.78	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Adaptability 7. In your first semester, you realized that one of your professors did not keep track of attendance and that you would not have to go to class. You skipped class when you had trouble waking up in the morning and when you had other commitments. You could have asked your friend in class for the notes you missed but thought that reviewing the textbook would be enough. You received a "C" on the first test. There are two tests left, each worth 20 percent of your course grade. You need a decent grade in the course so that your GPA does not suffer too much. What action should you take?					
1. Attend all of your classes from this time on.	4.54	.54	3.39	.00	0
2. Start going to all classes; contact the TA to help clarify problems with the exam, and get help with the new material.	4.88	.38	86.44	1.69	1
3. Go to class and study very hard.	4.53	.68	6.78	.00	0
4. Make sure you go to all classes and ask your professor to reward class attendance.	2.88	1.05	.00	20.34	-1
5. Try to skip class less frequently; ask someone for notes when you sleep in.	2.88	1.16	1.69	57.63	-1
6. Spend more time reading the textbook to do better on the next exam.	3.33	1.07	1.69	20.34	0
Perseverance 1. Your professor has just given you a project worth 10% of your overall grade that will probably require a good part of the semester to complete. She gave you some of the details you need to get started, and you are not sure how the project should proceed. She may or may not intend to give you any more information in class. How should you proceed?					
1. Work out the project to the best of your ability and approach the professor if you get stuck.	3.95	.85	3.39	11.86	0
2. Generate some ideas, and then go to office hours to see how the professor responds to them.	4.51	.70	49.15	1.69	1
3. Ask the professor about the project after class.	4.27	.72	5.08	1.69	0
4. Visit the professor or a teaching assistant during office hours to discuss the project.	4.53	.63	38.98	.00	1
5. Talk to other students to get an idea of what they are doing.	3.80	.76	1.69	37.29	-1
6. Try to get an idea of whether or not other students seem confused. If so, bring the issue up with the professor during class.	3.39	.85	1.69	47.46	-1

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 3. You attend a large lecture class with 200 students. You find the class dull and boring, and are having difficulty staying awake. The professor does not seem to care about class participation as he rarely asks questions. Passing class exams is the only thing you have to worry about. What should you do?					
1. Do what you can to stay awake, such as drinking caffeine or sitting toward the front of the class.	3.92	.92	23.73	.00	1
2. Read the class material beforehand to make the lecture more interesting.	4.10	.88	49.15	1.69	1
3. During the lecture, do some studying that is required for the course.	3.34	1.03	6.78	.00	0
4. Make sure you are getting enough sleep every school night.	4.10	.92	20.34	5.08	1
5. Skip the class if it is that dull and boring to you.	1.64	.94	.00	76.27	-1
6. Sit at the back of the class so that the professor does not notice your eyes closing.	1.93	1.11	.00	16.95	0
Perseverance 5. You have a professor you find "monstrously boring." He seems rude and longwinded and often reads material from the textbook. He does not require attendance but gives a few extra points on exams for good attendance. He posts general outlines of his lectures online. You need to pass the course. What should you do?					
1. Find a seat in the back and focus on some other work during the class.	2.29	1.13	1.69	11.86	0
2. Continue to attend, working hard to pay attention. You know you can hang in there until the end of the semester.	4.36	.69	49.15	1.69	1
3. Focus on taking as many notes as you can to make sure you completely understand the material.	4.23	.85	35.59	.00	1
4. Approach the professor and ask whether he could incorporate more interactive exercises into the lesson.	3.31	1.09	10.17	8.47	0
5. Read something interesting or play games to keep yourself occupied during the boring parts.	1.97	1.05	1.69	55.93	-1
6. Make an effort to attend class but allow yourself to skip it on days when you are extra tired.	2.31	.97	1.69	22.03	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 6. You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. You often want to raise your hand in class to ask the professor to explain something again. Some professors are happy to clarify but others act annoyed when you ask too many questions. You are not the only one struggling—other students also ask for further clarification. So you do not stand out. What should you do?					
1. Talk with the professors and TAs to get help on how to study.	4.54	.57	59.32	.00	1
2. Find a study partner and work on homework and class material together.	4.42	.67	15.25	.00	1
3. Get advice from your parents or close friends.	3.41	.91	.00	28.81	0
4. Study hard, try your best, and don't worry about it.	3.02	1.01	3.39	61.02	-1
5. Ask your advisor and professors if there are study groups or review sessions you can attend.	4.34	.66	13.56	.00	0
6. Hire a tutor for the difficult classes.	3.88	.97	8.47	10.17	0
Perseverance 7. You decided early in the term to do a paper on a topic you expected to be very interesting to you. After you read more about your topic, you realized that it is less interesting than you expected. In addition, your job has taken more time than you wanted, and you have had more work in your other courses than you anticipated. The paper is due in one week and it seems like you may have to engage in several "all-nighters" to complete it on time. You want a good grade. What should you do?					
1. Seek help from other students who may have had a similar experience.	3.31	.95	1.69	28.81	-1
2. Pick a topic that can be completed quicker. An "A" is an "A".	3.61	.97	11.86	20.34	0
3. Set up a schedule on which you can complete all of the other work you need to do, spend as much time on the paper as possible, and meet with the instructor to discuss what you have so far and get suggestions.	4.51	.65	71.19	.00	1
4. Do whatever it takes to complete the paper, including "all-nighters".	3.41	1.16	6.78	35.59	-1
5. Talk to the instructor about the situation and ask for advice.	4.02	.84	8.47	8.47	0
6. Make the paper a priority, but take into account how much the paper is worth in the class.	3.71	.81	.00	6.78	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 8. You are half way through a summer sports camp. You are finding the training difficult and tedious (e.g., long runs, difficult and lengthy exercises) even though you are in very good shape. You have not learned anything new about the sport in which you are interested. It is not clear to you why the coaches find the intensive training necessary. You also don't know how open they are to feedback. How should you react?					
1. Continue to train, because that is what is needed to perform.	4.03	.59	30.51	.00	1
2. See whether others at the camp feel the same way, and tell the coaches.	3.81	.99	30.51	1.72	1
3. Consider not going to the camp again because it was a disappointing experience.	2.42	1.10	3.39	29.31	0
4. Be upset and discouraged about the camp experience, though others may be learning something.	1.63	.72	.00	67.24	-1
5. Suggest new activities to the coaches to make the training more interesting.	3.91	.82	35.59	1.72	1
Ethics 1. You and your friends know that a mutual acquaintance has been dating someone for six months. However, one of your friends tries her best to get a date with this individual. She tells you about all her "schemes" but has not mentioned if she's expecting your help. You have heard the couple talk of moving in together but your friend says they also argue a lot. How should you react?					
1. If the acquaintance is in a happy relationship, tell your friend to wait and rethink it. If not, it is OK to get a date.	2.51	1.09	6.78	1.69	0
2. Support my friend. Good for her.	1.98	1.03	3.39	62.71	-1
3. Tell your friend to wait until the person is single or to just forget about the person.	4.02	1.03	35.59	1.69	0
4. Tell your friend that it is inappropriate to interfere in the relationship.	4.37	.85	52.54	1.69	1
5. Just be annoyed at your friend. Do not get involved.	2.53	1.23	1.69	32.20	0

Table 5 (cont'd)

Weak SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Ethics 4. One of your friends takes the same class you do, but at an earlier time. On the day of the midterm, he comes to your room and offers you a copy of the exam he just took. You know that the professor gives the same exam to all sections. Some students say that this professor's tests are difficult; others say the tests are manageable. You have spent five hours studying and have four hours left. The exam is worth 20 percent of your final grade. What should you do?					
1. Take the exam to study from, but not to memorize.	3.36	1.17	13.56	27.12	0
2. Tell your friend what you have studied and ask him whether that will be sufficient.	3.98	.90	23.73	5.08	1
3. Look the exam over so that you can study the areas that you don't know well, using it as a study guide.	3.69	1.20	28.81	6.78	1
4. Thank your friend, and politely decline.	2.97	1.34	20.34	13.56	0
5. Decline the exam, and suggest to the professor that in the future she should have different forms of the exam.	2.47	1.33	13.56	47.46	0
Ethics 8. You are taking a particularly difficult midterm and finding that you cannot answer some of the multiple choice questions. You realize it is very easy to see the paper of a student who you know is doing well in the course. When you look at this student's paper, she notices and pulls her paper away. She seems like a nice person so it is unlikely that she will tell the professor you were trying to cheat and get you into trouble. What action should you take?					
1. Don't show that you are embarrassed, and see if you can look at another paper.	1.68	1.02	5.08	57.63	-1
2. Just finish your exam as usual and then leave the room.	4.00	.87	35.59	5.08	1
3. Pretend like it didn't happen, make an effort not to look there again, and finish the test.	4.03	.69	25.42	1.69	1
4. Finish the exam, and approach or email the student after class to apologize.	3.40	1.29	33.90	1.69	0
5. Finish the exam, approach the student after class, and beg her not to tell on you.	2.14	1.02	.00	33.90	0

Table 6**. Scoring for strong version of the SJT**

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 3. A student on your floor is always organizing "social" activities, including trips to local bars. Aside from the fact that he/she is underage and failing several classes, you realize that the individual is drinking half a dozen or more drinks at least four times a week. No one else seems aware or concerned about the person. This individual is a close friend of yours. What should you do?					
1. Talk to the person about easing up on the alcohol, explaining that it will not help with his/her classes, which should be the main reason why he/she is in college.	4.15	1.04	49.06	1.85	1
2. Use humor to broach the topic and offer alternatives to the individual's usual "social" activities.	3.43	1.06	9.43	11.11	0
3. Bring up the situation with the floor's resident assistant.	3.02	1.30	3.77	24.07	0
4. Try to get the individual involved in other activities.	3.93	.84	20.75	1.85	1
5. Talk to the person to subtly determine if there are other issues that need to be addressed, and refer him/her to help if appropriate.	3.89	.77	9.43	.00	1
6. Talk to other people on the floor, and discuss ways to address the situation.	2.74	1.08	1.89	18.52	-1
7. Ask the individual once about this behavior and see where the discussion leads, then leave the person to his/her own course of action.	2.39	1.09	5.66	42.59	-1
Citizenship 5. You hear about a situation in which your friend was physically hurt and it can easily happen to other people. Nobody else knows about the situation so nobody will have done anything about it. You know the University encourages individuals to file prompt complaints in such situations. What should you do?					
1. Do whatever you can to keep it from happening again.	3.89	.95	3.70	1.85	0
2. Help the person who was mistreated any way you can.	4.28	.60	12.96	.00	1
3. Inform the appropriate authorities.	4.54	.79	38.89	3.70	1
4. Punish the person who mistreated the individual.	1.93	.93	.00	85.19	-1
5. Try to identify the causes of the event and correct them.	3.55	1.12	3.70	7.41	0
6. Talk to the person who was mistreated to see if he/she wants to do anything about the incident and let him/her know you are willing to help.	4.32	.78	40.74	1.85	1

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 7. Your roommate is having problems studying for an exam in a class that you took last semester and barely managed a passing grade in. You have finished your assignments for the night and were planning on going to dinner with a couple of friends. Your roommate asks you for help in studying for the exam. She says that she will get a failing grade in the class if she does not pass this exam. There is no one else she could ask for help. She asks for exactly two hours of your time. What should you do?					
1. Explain that you already made plans, but that when you get home from dinner you can review some material with her.	3.76	.95	31.48	1.85	1
2. Call your friends and cancel dinner. Stay home to help your roommate and have dinner with her instead.	3.74	.99	27.78	.00	1
3. Delay your dinner plans for an hour. Spend the time helping your roommate and then go out to dinner.	3.87	.89	37.04	.00	1
4. Explain that you would have liked to help her, but you already have plans that cannot be broken.	2.09	.90	.00	18.52	-1
5. Invite the roommate to go to dinner with you. You can help her on your drive to and from dinner.	2.13	1.17	1.85	40.74	-1
6. Tell her that you would like to help, but convince her that you would be of no real help to her.	2.09	1.05	1.85	38.89	-1
Citizenship 8. At the beginning of the semester, your professor asks for a volunteer to type out the lecture notes after each lecture for a hearing impaired student. The task would require you to type out the lecture notes right after class and e-mail them to the student so he has time to review them before the following class. The position is not paid. You know this professor presents lots of information during lectures and you can't always keep up. You have a heavy workload this semester and work part time. What should you do?					
1. Volunteer to type the notes. Do so the entire semester.	2.48	1.19	5.56	51.85	-1
2. Volunteer to type the notes. If you realize half-way through the semester that you can't do it anymore, ask the professor to find someone else.	3.22	1.14	12.96	20.37	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
3. Don't volunteer. You realize that it will take up too much time and someone else can probably do it better.	3.85	1.09	46.30	18.52	0
4. Volunteer to type half the notes if the professor can find someone to share the duty with you. If the professor can't find anyone else you wouldn't be able to do it.	3.50	1.21	16.67	5.56	0
5. Volunteer to type the notes. Ask a friend in class to share the responsibility with you so that you can take turns.	3.61	1.02	18.52	3.70	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Citizenship 9. Your professor announces in class that undergraduate students are needed started next week to help run subjects for his research study. While you would not receive any extra credit, the professor would appreciate any volunteers. This semester you are taking several difficult classes that consume most of your time. You have two midterms next week. What should you do?					
1. Examine your schedule and offer to volunteer a couple of hours a week when it is personally convenient.	4.15	.74	74.07	.00	1
2. Examine your schedule and offer to volunteer as many hours as you can.	3.13	1.17	14.81	5.56	0
3. Wait to see how many other people volunteer; agree to help a couple hours a week only if no one else volunteers.	3.22	1.00	3.70	.00	0
4. Volunteer only if a friend of yours volunteers and asks you to do it as well; volunteer for as many hours as your friend.	1.93	.84	.00	27.78	-1
5. Realize that you would have to give up some of your free time and choose not to volunteer.	2.70	1.08	5.56	12.96	0
6. Offer to run subjects only if you are paid.	1.89	.98	1.85	53.70	-1
Adaptability 3. Because of family problems, you find out that your parents can no longer support you financially at the same level as they have. You do not have enough money to continue in school for three more years. Your parents regret not being able to put you through college but expect that you will take out loans and work to support yourself. You won't be able to realize your career plans without a college degree. What plans should you make?					
1. Apply for student loans or get a part-time job.	4.74	.45	96.30	.00	1
2. Ask other, more wealthy family members for money to finish school.	2.81	1.13	.00	3.70	0
3. Drop out of school and save money for going back.	1.81	.95	.00	22.22	0
4. Take fewer classes because of the lower level of finances.	3.15	1.17	.00	3.70	0
5. Rely on your parents to figure something out—you should be able to count on them.	1.48	.64	.00	66.67	-1
6. Transfer to a less expensive, community college.	3.72	1.01	3.70	3.70	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Adaptability 4. You share a dorm room with three other students. Fifteen minutes before you are expecting a guest, you get home to find the place completely trashed. There is no sign of any of your roommates. You know that the guest would judge you for the mess. The individual's opinion is very important to you. Your roommates always react negatively to criticism. What should you do?					
1. Clean up the mess as much as possible before the guest arrives. Then speak with your roommates immediately upon their return, so your guest knows how concerned you were about the mess.	3.37	1.15	9.26	18.52	0
2. Leave the mess and explain the situation to your guest.	2.11	1.02	.00	55.56	-1
3. Leave the mess and take the guest somewhere else.	3.57	.98	11.11	.00	0
4. Clean up the mess as much as possible before the guest arrives. Then, without the guest around, ask the roommates why the place was trashed so badly and what can be done in the future to avoid this situation.	4.56	.77	75.93	1.85	1
5. Call the guest and ask to reschedule your meeting. When your roommates arrive, ask them to clean up the mess.	2.63	.96	3.70	24.07	0
Adaptability 5. You have a final exam in one of your major classes the next day. This exam will be weighted heavily in your grade and you want to do well because you intend to ask the professor for a letter of recommendation. Just as you settle down to study, you find out friends in an adjacent room have organized a party. They have invited you so you could join them. You know that the noise will make studying difficult. What should you do?					
1. Find an alternative place to study, like the library.	4.91	.29	98.15	.00	1
2. Tell yourself that you have to spend the next hour studying, and then you can go to the party.	2.34	1.13	1.85	57.41	-1
3. Ask them to keep quiet, or wait until you are done studying if they really want to party.	2.39	1.16	.00	18.52	-1
4. Put your headphones on and try to study anyway.	2.65	1.18	.00	18.52	-1
5. Put in earplugs to block out some of the noise and try to study anyway.	2.83	1.24	.00	5.56	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Adaptability 7. In your first semester, you realized that one of your professors did not record attendance and that you would not have to go to class. You skipped class when you had trouble waking up in the morning and when you had other commitments. You did not have a friend in class to ask for notes and thought that reviewing the textbook would be enough. You received a D on the first test. There is one test left and it is worth 40 percent of your grade. You cannot afford to fail the course; you will lose your scholarship if you do not maintain your GPA. What action should you take?					
1. Attend all of your classes from this time on.	4.67	.48	7.41	.00	0
2. Start going to all classes; contact the TA to help clarify problems with the exam, and get help with the new material.	4.85	.36	79.63	.00	1
3. Go to class and study very hard.	4.76	.43	11.11	.00	1
4. Make sure you go to all classes and ask your professor to reward class attendance.	2.93	1.20	.00	14.81	0
5. Try to skip class less frequently; ask someone for notes when you sleep in.	2.46	1.09	.00	68.52	-1
6. Spend more time reading the textbook to do better on the next exam.	2.96	1.16	1.85	16.67	0
Perseverance 1. Your professor has just given you a project worth 60% of your overall grade that will probably require a good part of the semester to complete. She gave you all the details you need to get started, but you want to be sure that you understood everything correctly. You know she does not plan to answer questions about the project in class. How should you proceed?					
1. Work out the project to the best of your ability and approach the professor if you get stuck.	3.80	.90	7.41	14.81	0
2. Generate some ideas, and then go to office hours to see how the professor responds to them.	4.70	.46	51.85	3.70	1
3. Ask the professor about the project after class.	4.22	.63	1.85	5.56	0
4. Visit the professor or a teaching assistant during office hours to discuss the project.	4.69	.47	37.04	.00	1
5. Talk to other students to get an idea of what they are doing.	3.69	.80	.00	25.93	-1
6. Try to get an idea of whether or not other students seem confused. If so, bring the issue up with the professor during class.	3.13	1.12	1.85	50.00	-1

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 3. You attend a lecture class with 30 students. You find the class dull and boring, and are having difficulty staying awake. The professor often randomly picks a name from the roster when he wants someone to answer a question. He also keeps students on their toes with pop quizzes. What should you do?					
1. Do what you can to stay awake, such as drinking caffeine or sitting toward the front of the class.	4.24	.73	31.48	1.85	1
2. Read the class material beforehand to make the lecture more interesting.	4.30	.82	50.00	1.85	1
3. During the lecture, do some studying that is required for the course.	2.98	1.07	.00	.00	0
4. Make sure you are getting enough sleep every school night.	4.30	.57	12.96	.00	1
5. Skip the class if it is that dull and boring to you.	1.31	.61	3.70	79.63	-1
6. Sit at the back of the class so that the professor does not notice your eyes closing.	1.56	.77	1.85	16.67	0
Perseverance 5. You have a professor you find "monstrously boring." He seems rude and longwinded and often reads material from the textbook. He requires attendance and takes points away for absences. He refuses to post any lecture notes online. The course is required for your major, so you need to do well. What should you do?					
1. Find a seat in the back and focus on some other work during the class.	1.72	1.02	1.85	22.64	-1
2. Continue to attend, working hard to pay attention. You know you can hang in there until the end of the semester.	4.50	.54	44.44	1.89	1
3. Focus on taking as many notes as you can to make sure you completely understand the material.	4.55	.64	35.19	3.77	1
4. Approach the professor and ask whether he could incorporate more interactive exercises into the lesson.	3.09	1.09	12.96	3.77	0
5. Read something interesting or play games to keep yourself occupied during the boring parts.	1.55	.75	5.56	33.96	-1
6. Make an effort to attend class but allow yourself to skip it on days when you are extra tired.	1.85	.86	.00	33.96	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 6. You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. You often want to raise your hand in class to ask the professor to explain something again. However, you have found that professors expect you to ask for extra help outside of class and to join available study groups. Other students do not seem to be struggling. So you really stand out. What should you do?					
1. Talk with the professors and TAs to get help on how to study.	4.63	.52	44.44	.00	1
2. Find a study partner and work on homework and class material together.	4.46	.64	14.81	1.85	1
3. Get advice from your parents or close friends.	3.44	.96	.00	35.19	0
4. Study hard, try your best, and don't worry about it.	2.70	1.09	5.56	57.41	-1
5. Ask your advisor and professors if there are study groups or review sessions you can attend.	4.52	.54	22.22	.00	1
6. Hire a tutor for the difficult classes.	4.17	.80	12.96	5.56	0
Perseverance 7. You decided early in the term to do a paper on a topic you expected to be very interesting to you. While you are still interested in your topic, you have found it very difficult to find information. In addition, your job has taken more time than you wanted, and you have had more work in your other courses than you anticipated. The paper is due in three days and it seems like you may have to engage in several "all-nighters" to complete it on time. You really need an "A". What should you do?					
1. Seek help from other students who may have had a similar experience.	3.35	.89	1.85	33.96	-1
2. Pick a topic that can be completed quicker. An "A" is an "A".	3.78	1.00	20.37	26.42	0
3. Set up a schedule on which you can complete all of the other work you need to do, spend as much time on the paper as possible, and meet with the instructor to discuss what you have so far and get suggestions.	4.53	.54	57.41	.00	1
4. Do whatever it takes to complete the paper, including "all-nighters".	3.61	1.11	5.56	30.19	-1
5. Talk to the instructor about the situation and ask for advice.	3.89	.79	9.26	3.77	0
6. Make the paper a priority, but take into account how much the paper is worth in the class.	3.85	.86	5.56	5.66	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Perseverance 8. You are half way through a summer sports camp. You are finding the training exceptionally difficult and tedious (e.g., long runs, difficult and lengthy exercises), particularly because you aren't in great shape. You have not learned anything new about the sport in which you are interested. The coaches say they want to get everyone into better shape—they believe this is key. They have also mentioned being open to suggestions. How should you react?					
1. Continue to train, because that is what is needed to perform.	4.06	.79	33.33	1.85	1
2. See whether others at the camp feel the same way, and tell the coaches.	3.62	.95	11.11	1.85	0
3. Consider not going to the camp again because it was a disappointing experience.	2.81	1.23	3.70	14.81	0
4. Be upset and discouraged about the camp experience, though others may be learning something.	1.78	.95	.00	79.63	-1
5. Suggest new activities to the coaches to make the training more interesting.	4.31	.82	51.85	1.85	1
Ethics 1. You and your friends know that a mutual acquaintance has been dating someone for over two years. However, one of your friends tries her best to get a date with this individual. She wants you to play an active role in her "schemes" to split the couple up. You have observed the couple get along very well and heard them talk about getting married after college. How should you react?					
1. If the acquaintance is in a happy relationship, tell your friend to wait and rethink it. If not, it is OK to get a date.	2.41	1.14	3.77	7.41	0
2. Support my friend. Good for her.	1.57	.84	.00	70.37	-1
3. Tell your friend to wait until the person is single or to just forget about the person.	4.07	.95	30.19	1.85	0
4. Tell your friend that it is inappropriate to interfere in the relationship.	4.43	.86	62.26	.00	1
5. Just be annoyed at your friend. Do not get involved.	2.48	1.18	3.77	20.37	0

Table 6 (cont'd)

Strong SJT Stem and Response Options	*M*	*SD*	% for "Best"	% for "Worst"	Scoring
Ethics 4. One of your friends takes the same class you do, but at an earlier time. On the day of the midterm, he comes to your room and offers you a copy of the exam he just took. You know that the professor gives the same exam to all sections. You have heard from many students that this professor's tests are very difficult. You have spent only a couple of hours studying and have two hours left. The exam is worth 45 percent of your grade. What should you do?					
1. Take the exam to study from, but not to memorize.	3.67	.97	16.67	20.37	0
2. Tell your friend what you have studied and ask him whether that will be sufficient.	3.83	.88	22.22	3.70	1
3. Look the exam over so that you can study the areas that you don't know well, using it as a study guide.	4.06	.83	48.15	5.56	1
4. Thank your friend, and politely decline.	2.78	1.30	1.85	16.67	0
5. Decline the exam, and suggest to the professor that in the future she should have different forms of the exam.	2.35	1.23	11.11	53.70	0
Ethics 8. You are taking a particularly difficult final exam and finding that you cannot answer most of the essay questions. You realize it is very easy to see the paper of a student who you know is doing well in the course. When you look at this student's paper, she notices and pulls her paper away. Chances are high that she will get you into trouble by telling the professor you were attempting to cheat off her. You would lose your scholarship. What action should you take?					
1. Don't show that you are embarrassed, and see if you can look at another paper.	1.54	.86	.00	72.22	-1
2. Just finish your exam as usual and then leave the room.	3.80	.81	33.33	1.85	1
3. Pretend like it didn't happen, make an effort not to look there again, and finish the test.	4.04	.85	33.33	.00	1
4. Finish the exam, and approach or email the student after class to apologize.	3.39	1.23	33.33	9.26	0
5. Finish the exam, approach the student after class, and beg her not to tell on you.	2.30	1.13	.00	16.67	0

Lab Study

Data preparation. Standardized test scores were requested from the university's registrar's office. Most students had taken the ACT (composite score comprised of English, reading, natural science, and mathematics sections). In cases where SAT scores (composite score comprised of critical reading, mathematics, and writing sections) were provided, these were converted into the metric of ACT scores (see Appendix M for the score conversion chart). In instances where an individual had taken a standardized test more than once, the highest composite score was retained for analyses.

Attention check failures. A total of 106 individuals who were otherwise eligible to be included in analyses (they were freshmen), failed one or more of the five attention checks and were excluded. Among these individuals, the average number of attention checks failed was 1.17 (*SD* = .49). T-tests and chi-square difference tests were done to examine whether the freshmen excluded from the analyses differed from the freshmen who passed all the attention checks and were included in analyses.

ACT scores were significantly higher for freshmen who were included in the analyses (*M* = 24.84, *SD* = 3.23) than for excluded freshmen (*M* = 23.60, *SD* = 3.24), $t(498) = -3.43$, $p = .001$. The ML/LL-based total SJT score was slightly higher for freshmen who were included (*M* = 17.77, *SD* = 5.06) than for excluded freshmen (*M* = 16.47, *SD* = 6.26), $t(131.22) = -1.90$, $p = .06$. Freshmen who were included perceived marginally lower constraints associated with the SJT (*M* = 3.23, *SD* = .89) than excluded freshmen (*M* = 3.41, *SD* = .82), $t(512) = 1.94$, $p = .053$. Minority freshmen were more likely to get excluded from analyses than White freshmen, χ^2 (1, *N* = 495) = 9.94, $p = .002$.

As shown in Table 7, there were no significant differences between included and excluded freshmen on the personality, biographical data, BARS, organizational citizenship, manipulation checks for clarity, consistency, and consequences, and impression management measures. There were also no significant differences on the ratings-based total SJT score, high school GPA, and first year college GPA. Surprisingly, there was no differences in the amount of time included (M = 45.87, SD = 9.00) and excluded freshmen (M = 44.73, SD = 9.36) took to complete the survey, $t(503) = -1.13$, *ns*. Likelihood of exclusion did not differ for male and female freshmen, χ^2 (1, N = 511) = .00, *ns*, and for those who completed the strong versus the weak version of the SJT, χ^2 (1, N = 514) = .26, *ns*.

Table 7. **Variable mean differences as a function of inclusion versus exclusion from analyses**

	Included Freshmen		Excluded Freshmen			
Variable	*M*	*SD*	*M*	*SD*	*df*	*t*
Openness	3.69	.54	3.63	.54	512	-1.07
Extraversion	3.31	.78	3.34	.72	512	.36
Conscientiousness	3.75	.62	3.75	.68	512	-.08
Agreeableness	4.12	.53	4.04	.55	512	-1.44
EmoStability	3.10	.69	3.03	.66	512	-.83
Biodata.Citizenship	3.32	.65	3.32	.62	512	-.01
Biodata.Adaptability	3.37	.44	3.30	.47	512	-1.47
Biodata.Perseverance	3.55	.43	3.51	.44	512	-.72
Biodata.Ethics	3.87	.48	3.79	.52	512	-1.59
Biodata.Overall	3.53	.33	3.48	.34	512	-1.29
BARS	5.02	.67	4.97	.80	145.64	-.50
Citizenship	3.68	.60	3.65	.63	512	-.35
Manip.Check_Overall	3.62	.58	3.64	.66	512	.39
Manip.Check_Clarity	3.74	.72	3.72	.79	512	-.36
Manip.Check_Consistency	3.71	.75	3.71	.84	512	-.01
Manip.Check_Constraints	3.23	.89	3.41	.82	512	1.94
Manip.Check_Consequences	3.81	.76	3.74	.78	512	-.85
Impression.Management	2.98	.50	2.96	.47	511	-.27
ML/LL-based Total SJT Score	17.77	5.06	16.47	6.26	131.22	-1.90
Ratings-based Total SJT Score	126.29	9.86	127.78	10.16	412	1.18
Ratings-based Intrapersonal SJT Score	55.37	6.53	56.05	6.41	447	.85
Ratings-based Interpersonal SJT Score	67.04	5.72	66.72	6.43	470	-.48
ACT Score	24.84	3.23	23.60	3.24	498	-3.43***
HS GPA	3.66	.31	3.67	.29	487	.29
College GPA	3.31	.53	3.22	.61	505	-1.55
Time to Complete Survey	45.87	9.00	44.73	9.36	503	-1.13

Note. *$p < .05$, **$p < .01$, ***$p < .001$. Included freshmen passed all attention checks; excluded freshmen failed one or more attention checks.

Descriptives. The intercorrelations, means, standard deviations, and reliabilities for study variables are provided in Table 8 (note that explanations of variable names and other notes about the table are presented on page 112).

Correlations indicate that across SJT versions (weak, strong), overall SJT scores were related to BARS ($r_{\text{ML/LL-based scores}} = .39$, $p < .001$; $r_{\text{ratings-based scores}} = .47$, $p < .001$), citizenship ($r_{\text{ML/LL-based scores}} = .26$, $p < .001$; $r_{\text{ratings-based scores}} = .24$, $p < .001$), and first year college GPA ($r_{\text{ML/LL-based scores}} = .19$, $p < .001$; $r_{\text{ratings-based scores}} = .20$, $p < .001$). With regard to personality and related variables, overall SJT scores (ML/LL-based scores and/or ratings-based scores) were related to openness ($r_{\text{ML/LL-based scores}} = .07$, *ns*; $r_{\text{ratings-based scores}} = .19$, $p < .001$), extraversion ($r_{\text{ML/LL-based scores}} = .14$, $p < .01$; $r_{\text{ratings-based scores}} = .09$, *ns*), conscientiousness ($r_{\text{ML/LL-based scores}} = .23$, $p < .001$; $r_{\text{ratings-based scores}} = .37$, $p < .001$), agreeableness ($r_{\text{ML/LL-based scores}} = .24$, $p < .001$; $r_{\text{ratings-based scores}} = .37$, $p < .001$), and impression management ($r_{\text{ML/LL-based scores}} = .32$, $p < .001$; $r_{\text{ratings-based scores}} = .33$, $p < .001$), and unrelated to emotional stability ($r_{\text{ML/LL-based scores}} = .07$, *ns*; $r_{\text{ratings-based scores}} = .06$, *ns*). With regard to background information and test scores, overall SJT scores were related to citizenship (biodata) ($r_{\text{ML/LL-based scores}} = .23$, $p < .001$; $r_{\text{ratings-based scores}} = .27$, $p < .001$), adaptability (biodata) ($r_{\text{ML/LL-based scores}} = .17$, $p < .001$; $r_{\text{ratings-based scores}} = .18$, $p < .001$), perseverance (biodata) ($r_{\text{ML/LL-based scores}} = .33$, $p < .001$; $r_{\text{ratings-based scores}} = .45$, $p < .001$), ethics (biodata) ($r_{\text{ML/LL-based scores}} = .40$, $p < .001$; $r_{\text{ratings-based scores}} = .43$, $p < .001$),

and high school GPA ($r_{\text{ML/LL-based scores}} = .16$, $p < .01$; $r_{\text{ratings-based scores}} = .22$, $p < .001$), and unrelated to ACT scores ($r_{\text{ML/LL-based scores}} = -.06$, *ns*; $r_{\text{ratings-based scores}} = -.03$, *ns*). Finally, with regard to demographics, overall SJT scores were related to gender, such that women tended to get higher scores ($r_{\text{ML/LL-based scores}} = -.19$, $p < .001$; $r_{\text{ratings-based scores}} = -.21$, $p < .001$), and unrelated to race ($r_{\text{ML/LL-based scores}} = .03$, *ns*; $r_{\text{ratings-based scores}} = .07$, *ns*).

As compared to overall SJT scores, the more cognitive predictors showed lower correlations with BARS ($r_{\text{ACT scores}} = .03$, *ns*; $r_{\text{HS GPA}} = .11$, $p < .05$), and citizenship ($r_{\text{ACT scores}} = .05$, *ns*; $r_{\text{HS GPA}} = .17$, $p < .001$), and higher correlations with first year college GPA ($r_{\text{ACT scores}} = .42$, $p < .001$; $r_{\text{HS GPA}} = .44$, $p < .001$). These predictors also had higher correlations with race, such that Whites tended to have higher scores on average than minorities ($r_{\text{ACT scores}} = .32$, $p < .001$; $r_{\text{HS GPA}} = .24$, $p < .001$). ACT scores (but not high school GPA; $r_{\text{gender}} = -.02$, *ns*) were significantly related to gender ($r = .14$, $p < .01$), but unlike overall SJT scores, favored men rather than women.

SJT dimension scores are not included in Table 8 as these dimensions had low reliabilities (ranging from .12 to .44 for ML/LL-based scores; ranging from .10 to .64 for ratings-based scores) and were not used in analyses. These low reliabilities for ratings-based and ML/LL-based dimension scores are provided in Table 9.

Several variables were related to strength perceptions of the situations in the SJT (i.e., the manipulation check). Minority respondents (relative to Whites) reported higher perceptions of

clarity (r = -.16, p < .01) and constraints (r = -.16, p < .01), more agreeable respondents reported higher perceptions of consequences (r = .19, p < .001), and respondents with higher ACT scores reported lower perceptions of clarity (r = -.16, p < .01) and constraints (r = -.11, p < .05).

Interestingly, the relation between perceptions of clarity and consistency was considerably stronger (r = .77, p < .001) relative to the other relations between perceived strength dimensions ($r_{\text{clarity-constraints}}$ = .26, p < .001; $r_{\text{clarity-consequences}}$ = .35, p < .001; $r_{\text{consistency-constraints}}$ = .31, p < .001; $r_{\text{consistency-consequences}}$ = .38, p < .001; $r_{\text{constraints-consequences}}$ = .41, p < .001). Thus, individuals who thought the situations in the SJT provided more consistent information, also tended to think that the situations provided clearer expectations. This strong association between perceptions of clarity and consistency is in line with earlier research ($r_{\text{clarity-consistency}}$ = .81, p < .01; Meyer et al., in press).

Table 8. **Intercorrelations, means, standard deviations, and reliabilities for study variables**

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. SJT.s[a]	17.77	5.06	*.55*												
2. SJT.rating.interpersonal.s[b]	67.04	5.72	**.45**	*.65*											
3. SJT.rating.intrapersonal.s[b]	55.37	6.53	**.70**	**.31**	*.69*										
4. SJT.rating.s[c]	126.29	9.86	**.69**	**.78**	**.83**	*.72*									
5. SJT Version (0=Weak; 1=Strong)	.49	.50	**.24**	**-.13**	**.47**	**.21**	--								
6. Openness	3.69	.54	.07	**.29**	.07	**.19**	**-.13**	*.81*							
7. Extraversion	3.31	.78	**.14**	.07	.06	.09	-.03	**.25**	*.89*						
8. Conscientiousness	3.75	.62	**.23**	**.31**	**.32**	**.37**	-.08	**.22**	.04	*.83*					
9. Agreeableness	4.12	.53	**.24**	**.40**	**.22**	**.37**	.01	**.24**	**.24**	**.21**	*.80*				
10. EmoStability	3.10	.69	.07	.08	.04	.06	-.02	**.20**	**.20**	**.12**	**.14**	*.84*			
11. Biodata.Citizenship	3.32	.65	**.23**	**.24**	**.17**	**.27**	.00	**.18**	**.22**	**.11**	**.25**	**.11**	*.76*		
12. Biodata.Adaptability	3.37	.44	**.17**	**.12**	**.18**	**.18**	-.05	**.25**	**.29**	**.46**	**.11**	**.41**	**.17**	*.65*	
13. Biodata.Perseverance	3.55	.43	**.33**	**.34**	**.37**	**.45**	-.09	**.37**	**.27**	**.42**	**.28**	.07	**.27**	**.42**	*.70*
14. Biodata.Ethics	3.87	.48	**.40**	**.40**	**.36**	**.43**	.01	**.17**	-.03	**.23**	**.38**	.06	**.25**	**.14**	**.34**
15. Biodata.Overall	3.53	.33	**.42**	**.41**	**.40**	**.50**	-.04	**.35**	**.28**	**.42**	**.38**	**.23**	**.72**	**.59**	**.71**
16. BARS	5.02	.67	**.39**	**.39**	**.38**	**.47**	-.08	**.43**	**.29**	**.29**	**.35**	**.17**	**.37**	**.35**	**.50**
17. Citizenship	3.68	.60	**.26**	**.24**	**.17**	**.24**	-.06	**.10**	**.35**	**.14**	**.36**	**.10**	**.39**	**.23**	**.34**
18. Impression.Management	2.98	.50	**.32**	**.23**	**.32**	**.33**	.03	**.12**	-.05	**.25**	**.30**	**.21**	**.10**	**.18**	**.21**
19. Manip.Check_Clarity	3.74	.72	.06	.03	.07	.05	.03	.06	.09	.08	.01	-.08	.04	.06	**.12**
20. Manip.Check_Consistency	3.71	.75	.07	.05	.04	.05	-.01	.08	.06	.06	.07	-.02	.07	.09	**.11**
21. Manip.Check_Constraints	3.23	.89	-.06	**-.12**	-.06	**-.11**	.03	.04	-.04	-.02	-.04	-.05	-.01	-.09	.04
22. Manip.Check_Consequences	3.81	.76	.07	**.14**	**.12**	**.17**	.06	.08	-.04	.04	**.19**	-.02	.07	.03	**.14**
23. Manip.Check_Overall[d]	3.62	.58	.04	.02	.05	.04	.03	.08	.02	.05	.06	-.06	.05	.02	**.13**
24. ACT Score	24.84	3.23	-.06	**.11**	**-.15**	-.03	-.09	**.13**	**-.19**	.01	.06	.02	**.15**	.09	.02
25. HS GPA	3.66	.31	**.16**	**.12**	**.19**	**.22**	-.04	.01	-.01	**.22**	.06	.04	**.21**	**.21**	**.23**

Table 8 (cont'd)

	M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
26. College GPA	3.31	.53	**.19**	.07	**.21**	**.20**	.01	-.01	**-.11**	**.20**	.05	-.07	.09	**.25**	**.22**
27. Gender (0=Female; 1=Male)	.21	.41	**-.19**	**-.26**	**-.11**	**-.21**	-.04	.07	-.03	-.09	**-.14**	**.22**	**-.17**	.03	**-.11**
28. Race (0=Minority; 1=White)	.83	.38	.03	.09	-.02	.07	.03	.01	.02	.07	**.14**	.04	.02	**.13**	.04
29. Time[e]	45.87	9.00	.08	.07	.06	.04	.04	.03	**-.15**	-.06	.07	.06	.00	**-.12**	**-.11**

Table 8 (cont'd)

	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
14. Biodata.Ethics	*.64*															
15. Biodata.Overall	**.63**	*.78*														
16. BARS	**.43**	**.61**	*.77*													
17. Citizenship	**.22**	**.45**	**.29**	*.77*												
18. Impression.Management	**.49**	**.35**	**.38**	.05	*.76*											
19. Manip.Check_Clarity	.03	.09	**.11**	**.12**	-.02	*.82*										
20. Manip.Check_Consistency	.05	**.12**	**.12**	**.14**	-.03	**.77**	*.88*									
21. Manip.Check_Constraints	-.04	-.03	-.05	.01	-.05	**.26**	**.31**	*.88*								
22. Manip.Check_Consequences	**.19**	**.16**	**.13**	**.11**	.04	**.35**	**.38**	**.41**	*.87*							
23. Manip.Check_Overall[d]	.06	**.10**	.09	**.12**	-.02	**.78**	**.81**	**.70**	**.69**	*.92*						
24. ACT Score	.09	**.14**	.03	.05	.02	**-.16**	-.07	**-.11**	.06	**-.10**	--					
25. HS GPA	**.14**	**.29**	**.11**	**.17**	**.16**	-.04	-.02	-.01	**.14**	.02	**.31**	--				
26. College GPA	.07	**.22**	.08	.09	**.15**	-.03	-.02	-.07	**.12**	-.01	**.42**	**.44**	--			
27. Gender (0=Female; 1=Male)	**-.14**	**-.16**	-.07	-.06	.00	.01	.00	.06	.04	.04	**.14**	-.02	.09	--		
28. Race (0=Minority; 1=White)	.01	.06	-.09	**.12**	-.05	**-.16**	-.06	**-.16**	.00	**-.13**	**.32**	**.24**	**.23**	.06	--	
29. Time[e]	**.11**	-.03	.08	**-.16**	**.23**	-.06	-.07	-.07	.05	-.05	-.10	.02	-.02	.03	-.08	--

Note. Correlations in bold are significant at at least the .05 level (2-tailed). Internal consistency reliabilities are shown along the diagonal. *N*s range from 339 (for SJT.rating.s) to 408. [a]SJT.s is the sum of item scores based on Most Likely and Least Likely responses (excluding item 4—citizenship 8). [b]SJT.rating.interpersonal.s and SJT.rating.intrapersonal.s are comprised of several items each and represent interpersonal and intrapersonal SJT dimensions, respectively. These are based on Likert ratings. [c]SJT.rating.s is the sum of item scores based on Likert ratings (excluding item 4—citizenship 8). [d]Manipulation.Check_Overall is an average of all the manipulation check items. [e]Time represents the amount of time it took participants to complete the survey.

Table 9. **Reliabilities for SJT dimension scores**

	Internal Consistency Reliability (α)	
Dimension	ML/LL-based SJT scores	Ratings-based SJT scores
Citizenship	.30	.10
Adaptability	.12	.41
Perseverance	.44	.64
Ethics	.19	.31

Factor analysis. Since the a priori dimensional structure did not produce reliable scales, all the scored items (except for item 4—citizenship 8, for which no responses were keyed as best or worst for the weak version of the SJT, and item 17—ethics 4, which was removed due to low loadings on all extracted factors) were subjected to an exploratory factor analysis (EFA) to determine if a more interpretable structure existed. Maximum likelihood estimation with varimax rotation was used for the EFA on the ratings-based scores and the ML/LL-based scores in turn. Five factors were extracted for the ratings-based item scores and explained 36.18% of the variance. This model adequately fit the data, $\chi^2(50) = 59.59$, *ns*. However, upon further examination, all the extracted factors had poor reliability ranging from -.60 to .62 and one of the factors consisted of one item. Six factors were extracted for the ML/LL-based scores and explained 32.25% of the variance. This model adequately fit the data, $\chi^2(39) = 28.46$, *ns*. However, upon further examination, three factors consisted of just one item each, and the other three factors had poor reliability ranging from .21 to .46.

Given that the items in the current SJT may be conceptualized as having interpersonal and intrapersonal dimensions (Oswald et al., 2004), a forced solution of two factors was also tried. Although the goodness of fit test for the ratings-based scores and the ML/LL-based scores [$\chi^2(89) = 213.06$, $p < 001$, and $\chi^2(89) = 117.39$, $p < .05$, respectively] indicated that these models did not fit the data well, the factors extracted for the ratings-based scores nearly met the .70 minimum convention of reliability. All the perseverance items, one adaptability item, and one citizenship item loaded on a single factor and appeared to have a personal motivation component ("intrapersonal" factor; $\alpha = .69$). The ethics items and majority of the citizenship and adaptability items loaded on the other factor, and had a dealing with others component

("interpersonal" factor; α = .65). The two factors were related but not so strongly as to be impractical (r = .31, p < .001) and explained 23.68% of the variance in the ratings-based scores. Loadings of items on the factors are presented in Table 10. Scale scores based on these factors were used in subsequent analyses. Correlations of these factors with other variables are included in Table 8.

The same two factors explained just 12.79% of the variance in ML/LL-based scores and had poor reliability (α = .51 for intrapersonal factor, α = .37 for interpersonal factor). Scale scores based on these factors were not used in subsequent analyses.

Table 10. **Loadings for ratings-based SJT scores**

		Factor	
Label	Item	1 ("Intrapersonal")	2 ("Interpersonal")
SJT11.rating.s	Perseverance 3	.731	.159
SJT12.rating.s	Perseverance 5	.658	.208
SJT9.rating.s	Adaptability 7	.572	.107
SJT1.rating.s	Citizenship 3	.442	-.251
SJT14.rating.s	Perseverance 7	.390	.135
SJT10.rating.s	Perseverance 1	.345	.142
SJT13.rating.s	Perseverance 6	.331	.122
SJT15.rating.s	Perseverance 8	.292	.124
SJT3.rating.s	Citizenship 7	-.057	.522
SJT5.rating.s	Citizenship 9	.088	.498
SJT2.rating.s	Citizenship 5	.105	.486
SJT8.rating.s	Adaptability 5	.212	.412
SJT6.rating.s	Adaptability 3	.186	.397
SJT7.rating.s	Adaptability 4	.201	.337
SJT16.rating.s	Ethics 1	.299	.317
SJT18.rating.s	Ethics 8	.210	.213
	Variance explained	13.95%	9.73%

Manipulation check. T-tests were used to examine whether participants recognized that the SJT situations varied in situational strength across study conditions. T-tests indicated that individuals did not perceive a difference in overall strength (perceptions averaged across the four strength dimensions), $t(406) = -.70$, *ns*, or in the individual strength dimensions of clarity, $t(406) = -.52$, ns, consistency, $t(406) = .27$, *ns*, constraints, $t(406) = -.69$, *ns*, and consequences, $t(406) = -1.22$, *ns*.

The fact that test takers did not perceive strength differences across the two versions of the SJT suggests that the strength manipulation was fairly weak on average. This is not surprising given that an effort was made to not make SJT stems overly strong. However, it is also possible that individuals would have reported different strength perceptions of particular SJT items had they been asked for perceptions of individual items comprising the test as opposed to summary judgments of the entire set of items. Accordingly, the situational strength manipulation may have anticipated effects on certain items but not others. Thus, analyses were still conducted to address the hypotheses and research questions in this study, though results should be regarded as exploratory.

Hypothesis 1a. Hypothesis 1a predicted an interaction of SJT dimension scores and SJT strength on criterion-related validities such that validities will be lower for a strong SJT. The ratings-based interpersonal and intrapersonal SJT dimensions were used for these analyses instead of the adaptability, citizenship, perseverance, and ethics dimensions, as the latter set of SJT dimensions proved to have particularly low internal consistency reliabilities. The SJT dimension score and situational strength were entered in the first step of the regression. The interaction of the SJT dimension score and situational strength was entered in the second step.

SJT dimension scores were centered prior to entry in the regression. Organizational citizenship, first year college GPA, and BARS were used as outcomes.

There was no statistically significant interaction between interpersonal SJT scores and situational strength on organizational citizenship scores, $\Delta R^2 = .00$, $\Delta F(1, 373) = .08$, *ns.* There was likewise no statistically significant interaction between intrapersonal SJT scores and situational strength on organizational citizenship scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = .28$, *ns.*

There was no statistically significant interaction between interpersonal SJT scores and situational strength on first year college GPA, $\Delta R^2 = .00$, $\Delta F(1, 369) = .38$, *ns.* There was likewise no statistically significant interaction between intrapersonal SJT scores and situational strength on first year college GPA, $\Delta R^2 = .00$, $\Delta F(1, 357) = .35$, *ns.*

There was only a marginally significant interaction between interpersonal SJT scores and situational strength on the BARS, $\Delta R^2 = .01$, $\Delta F(1, 373) = 3.04$, $p = .08$. Counter to expectations, the relation between interpersonal SJT and BARS scores was actually somewhat stronger ($r = .44$, $p < .001$) for the strong SJT relative to the weak ($r = .35$, $p < .001$). There was no statistically significant interaction between intrapersonal SJT scores and situational strength on the BARS, $\Delta R^2 = .00$, $\Delta F(1, 361) = .51$, *ns.* Thus, Hypothesis 1a was not supported for any of the outcomes.

Hypothesis 1b. Hypothesis 1b predicted an interaction of the overall SJT score and situational strength on criterion-related validities such that validities will be lower for a strong SJT. Two versions of the SJT score (score based on Most Likely/Least Likely responses; score based on ratings) and three different outcomes (organizational citizenship; first year college

GPA; BARS) were examined. The SJT score and situational strength were entered in the first step of the regression. The interaction of SJT score and situational strength was entered in the second step. SJT scores were centered prior to entry in the regression.

There was no statistically significant interaction between ML/LL-based SJT scores and situational strength on organizational citizenship scores, $\Delta R^2 = .00$, $\Delta F(1, 380) = .48$, *ns*. There was likewise no statistically significant interaction between ratings-based SJT scores and situational strength on organizational citizenship scores, $\Delta R^2 = .00$, $\Delta F(1, 335) = .32$, *ns*.

There was no statistically significant interaction between ML/LL-based SJT scores and situational strength on first year college GPA, $\Delta R^2 = .00$, $\Delta F(1, 375) = .58$, *ns*. There was likewise no statistically significant interaction between ratings-based SJT scores and situational strength on first year college GPA, $\Delta R^2 = .00$, $\Delta F(1, 332) = .00$, *ns*.

There was no statistically significant interaction between ML/LL-based SJT scores and situational strength on the BARS, $\Delta R^2 = .00$, $\Delta F(1, 380) = 1.18$, *ns*. There was likewise no statistically significant interaction between ratings-based SJT scores and situational strength on the BARS, $\Delta R^2 = .00$, $\Delta F(1, 335) = .11$, *ns*. Thus, Hypothesis 1b was not supported for any of the outcomes.

Hypothesis 2a. Hypothesis 2a predicted that the convergent validities of the SJT dimension scores with the biographical data dimension scores will be moderated by SJT strength such that validities will be lower for a strong SJT. This hypothesis could not be tested because the SJT dimensions (i.e., adaptability, citizenship, perseverance, and ethics) that corresponded with the biodata dimensions had low internal consistency reliabilities.

Hypothesis 2b. Hypothesis 2b predicted that the convergent validity of the overall SJT score with the overall score on the biographical data measure will be moderated by SJT strength such that validity will be lower for a strong SJT. Two versions of the SJT score (score based on Most Likely/Least Likely responses; score based on ratings) were examined as dependent variables in the regression. Biodata scores and situational strength were entered in the first step of the regression. The interaction of biodata scores and situational strength was entered in the second step. Biodata scores were centered prior to entry in the regression. There was no statistically significant interaction of biodata scores and situational strength on ML/LL-based SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 380) = .80$, *ns*. There was likewise no statistically significant interaction of biodata scores and situational strength on ratings-based SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 335) = .71$, *ns*. Thus, Hypothesis 2b was not supported.

Hypothesis 3. Hypothesis 3 predicted that the discriminant validities of the SJT dimension scores with the biographical data dimension scores will be moderated by SJT strength such that validities will be higher (i.e., less favorable) for a strong SJT. This hypothesis could not be tested because the SJT dimensions (i.e., adaptability, citizenship, perseverance, and ethics) that corresponded with the biodata dimensions had low internal consistency reliabilities.

Research question 1. Research question 1 asked whether situational strength would moderate the interrelations of the SJT dimensions with various personality scales. The ratings-based interpersonal and intrapersonal SJT dimensions were used for these analyses. A personality scale score (openness, extraversion, conscientiousness, agreeableness, or emotional stability) and situational strength were entered in the first step of the regression. The interaction of the personality scale score and situational strength was entered in the second step. Personality

scale scores were centered prior to entry in the regression. Rating-based interpersonal and intrapersonal SJT scores were used as outcomes.

There was no statistically significant interaction of openness and situational strength on the intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = .73$, *ns*. There was, however, a statistically significant interaction of openness and situational strength on the interpersonal SJT scores, $\Delta R^2 = .02$, $\Delta F(1, 374) = 10.16$, $p < .01$. The simple slope for the weak SJT ($b = 4.60$, $SE_b = .74$, $t(373) = 6.25$, $p < .001$) was positive and significant, whereas the simple slope for the strong SJT ($b = 1.26$, $SE_b = .74$, $t(373) = 1.71$, $p = .09$) was not. The interaction is illustrated in Figure 4. The form of this interaction is consistent with the situational strength hypothesis—the weak SJT's interpersonal dimension was a more valid measure of openness than the strong SJT's interpersonal dimension.

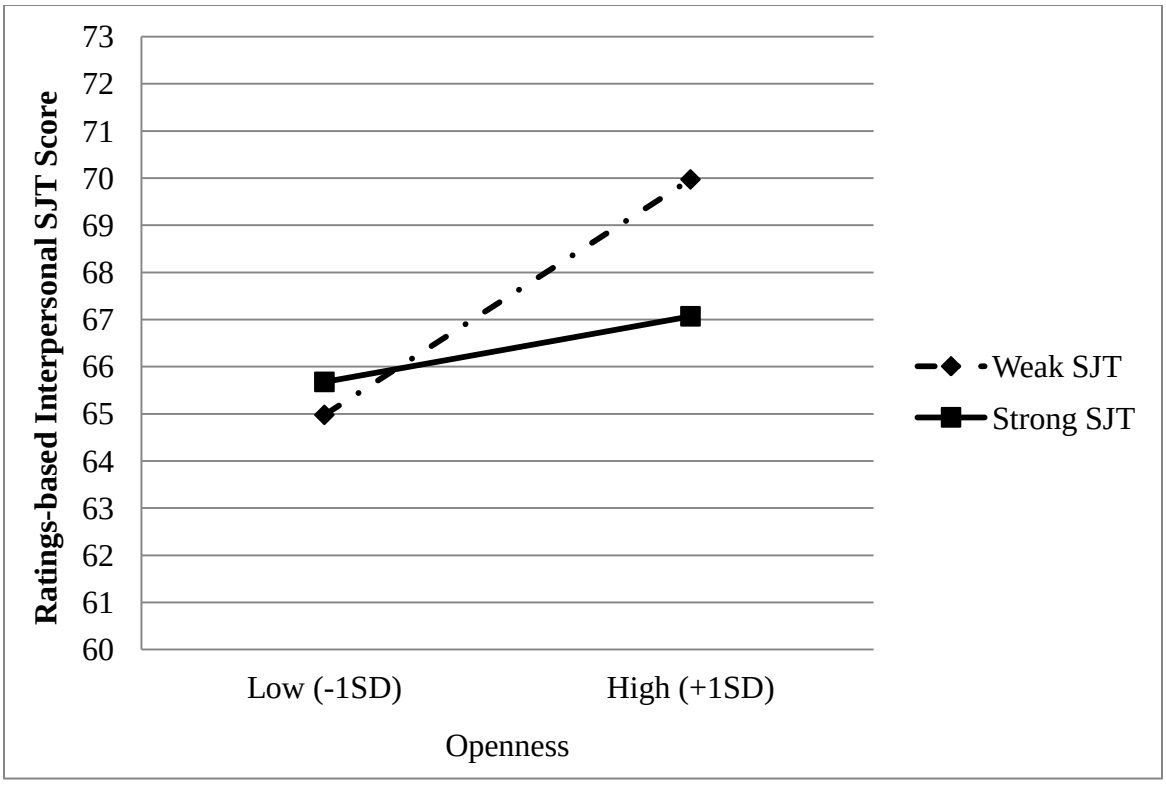

Figure 4. **Interaction of openness and situational strength on interpersonal SJT scores.**

To better understand why the weak SJT's interpersonal dimension was a better measure of openness than the strong SJT's interpersonal dimension, SJT items comprising this dimension were examined to identify those that measured openness significantly better on the weak SJT relative to the strong. One citizenship item (test item 5), two adaptability items (test items 6 and 8), and one ethics item (test item 16) met this criterion, $\Delta R^2 = .03$, $\Delta F(1, 396) = 13.26$, $p < .001$, $\Delta R^2 = .01$, $\Delta F(1, 403) = 4.16$, $p < .05$, $\Delta R^2 = .01$, $\Delta F(1, 399) = 5.45$, $p < .05$, and $\Delta R^2 = .01$, $\Delta F(1, 400) = 4.24$, $p < .05$, respectively.

Item 5, a citizenship item which has to do with volunteering to help a professor run subjects for his research study, can be used as an illustrative example. As shown in Table 11 below, ratings on five of the six response options on the weak situation, which describes a

manageable workload (low constraints), were either positively or negatively associated with openness whereas none of the same response option ratings on the strong situation, which describes a heavy workload (high constraints) were significantly associated with openness. The response options for this item represent various degrees of willingness to volunteer (citizenship) and when constraints are low, individuals higher in openness respond in ways that represent greater willingness to volunteer. When constraints are high, however, openness is unrelated to willingness to volunteer. Likelihood ratings that individuals provided for the various response options offer additional insight. On average, individuals responding to the strong situation rated response option 5 (Realize that you would have to give up some of your free time and choose not to volunteer), which represents a low level of citizenship, higher (M = 2.95, SD = 1.20) than did individuals responding to the weak situation (M = 2.63, SD = 1.14), $t(405) = -2.74$, $p < .01$. On the other hand, individuals responding to the strong situation rated option 2 (Examine your schedule and offer to volunteer as many hours as you can), which represents a high level of citizenship, lower (M = 2.75, SD = 1.29) than did individuals responding to the weak situation (M = 3.17, SD = 1.20), $t(405) = 3.47$, $p < .001$. Individuals responding to the strong situation likewise rated option 1 (Examine your schedule and offer to volunteer a couple of hours a week when it is personally convenient), which represents a fairly high level of citizenship as well, lower (M = 3.71, SD = 1.22) than did individuals responding to the weak situation (M = 4.06, SD = .97), $t(380.38) = 3.22$, $p < .01$. In light of the heavier workload illustrated in the strong situation, the situational press to not volunteer, regardless of one's level of openness makes sense and is consistent with the situational strength hypothesis.

It should be noted that ratings on response options 1 (keyed as correct for both the weak and strong SJT), 4 (keyed as incorrect for the strong SJT), and 6 (keyed as incorrect for both the

weak and strong SJT) were the only ones that had implications for individuals' SJT scores. Because ratings on the remaining options are not scored (these options are keyed as '0'), the fact that the option ratings differ (as a function of situational strength) in the degree to which they represent openness is ultimately not reflected in the item *scores*.

Table 11. **Means and correlations (with openness) for SJT response ratings**

	Weak SJT				Strong SJT					
Response Option	Key	*r* with Openness	*M*	*SD*	Key	*r* with Openness	*M*	*SD*	*df*	*t*
SJT5.Op1	1	.21**	4.06	.97	1	-.07	3.71	1.22	380.38	3.22**
SJT5.Op2	0	.21**	3.17	1.20	0	.10	2.75	1.29	405	3.47***
SJT5.Op3	0	.04	3.13	1.13	0	-.12	3.30	1.16	405	-1.45
SJT5.Op4	0	-.17*	2.91	1.20	-1	-.05	2.71	1.17	402	1.69
SJT5.Op5	0	-.17*	2.63	1.14	0	-.02	2.95	1.20	405	-2.74**
SJT5.Op6	-1	-.33***	1.78	1.07	-1	-.07	1.74	1.05	403	.38
SJT6.Op1	1	.20**	4.86	.37	1	-.05	4.82	.48	375.37	1.07
SJT6.Op2	0	-.02	2.50	1.41	0	.03	2.08	1.19	397.84	3.31**
SJT6.Op3	0	-.04	1.72	1.06	0	-.08	1.61	.92	401.82	1.19
SJT6.Op4	0	.16*	2.90	1.19	0	-.07	2.58	1.18	403	2.74**
SJT6.Op5	-1	-.28***	2.05	1.14	-1	-.17*	2.04	1.05	405	.16
SJT6.Op6	0	-.02	3.67	1.13	0	-.02	3.43	1.20	405	2.08*
SJT8.Op1	1	.13	4.30	.88	1	.14	4.76	.49	324.47	-6.46***
SJT8.Op2	0	-.16*	3.29	1.24	-1	-.02	2.67	1.25	404	5.03***
SJT8.Op3	-1	-.26***	1.72	.86	-1	-.04	1.86	.85	404	-1.64
SJT8.Op4	0	.04	3.75	1.07	-1	-.07	3.17	1.28	388.97	5.04***
SJT8.Op5	0	.02	3.08	1.31	0	-.10	2.86	1.28	402	1.70
SJT16.Op1	0	.03	2.84	1.24	0	-.04	2.55	1.31	405	2.28*
SJT16.Op2	-1	-.21**	2.06	1.06	-1	-.09	1.63	.89	405	4.52***
SJT16.Op3	0	.08	4.38	.80	0	.00	4.49	.72	406	-1.34
SJT16.Op4	1	.23***	4.33	.82	1	.07	4.59	.66	389.70	-3.55***
SJT16.Op5	0	-.03	2.63	1.07	0	-.08	3.08	1.17	405	-4.09***

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

Next, there was no significant interaction of agreeableness and situational strength on the intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = .17$, *ns*. There was, however, a statistically significant interaction of agreeableness and situational strength on the interpersonal SJT scores, $\Delta R^2 = .02$, $\Delta F(1, 373) = 10.26$, $p < .01$. The simple slopes for both the weak SJT ($b = 6.00$, $SE_b = .71$, $t(373) = 8.42$, $p < .001$) and strong SJT ($b = 2.78$, $SE_b = .71$, $t(373) = 3.91$, $p < .001$) were positive and significant, but the former was steeper. The interaction is illustrated in Figure 5. The form of this interaction, like the one between openness and SJT strength above, is consistent with the situational strength hypothesis.

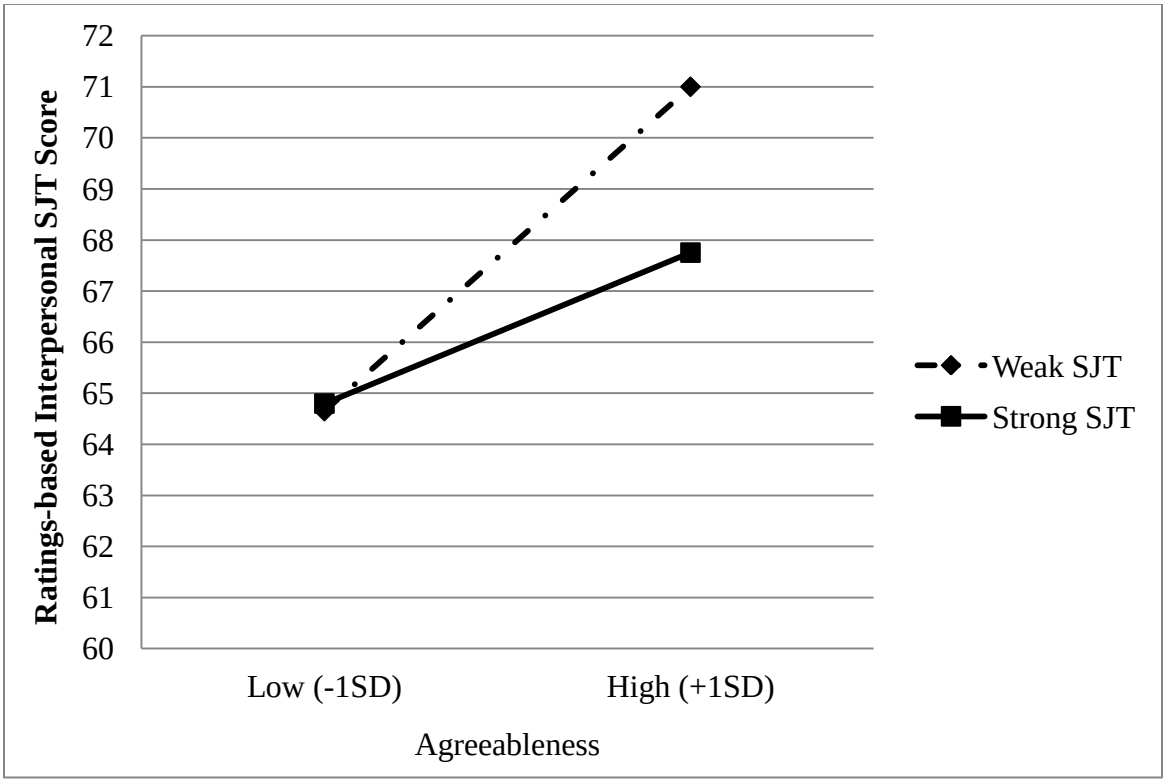

Figure 5. **Interaction of agreeableness and situational strength on interpersonal SJT scores.**

To better understand why the weak SJT's interpersonal dimension was a better measure of agreeableness than the strong SJT's interpersonal dimension, SJT items comprising this dimension were examined to identify one that was a significantly better measure of agreeableness on the weak SJT relative to the strong. Three adaptability items (test items 6, 7, and 8) and one ethics item (test item 16) met or nearly met this criterion, $\Delta R^2 = .01$, $\Delta F(1, 403) = 3.53$, $p = .06$, $\Delta R^2 = .02$, $\Delta F(1, 398) = 8.53$, $p < .01$, $\Delta R^2 = .01$, $\Delta F(1, 399) = 5.4$, $p < .05$, $\Delta R^2 = .01$, $\Delta F(1, 400) = 3.55$, $p = .06$, respectively.

Item 7, an adaptability item, which requires deciding what to do about a messy dorm room in anticipation of a guest's visit, can be used as an illustrative example. As shown in Table 12 below, ratings on four of the five response options on the weak situation were significantly positively or negatively associated with agreeableness, whereas none of the same response option ratings on the strong situation were significantly associated with agreeableness (but note that ratings on one response option were). For the weak situation, where the guest is arriving in an hour, the guest's perspective on messes is unknown, and the roommates who created the mess may or may not be open to criticism, more agreeable respondents reported higher willingness to clean up the mess as much as possible and later confront the roommates (option 4). On the other hand, more agreeable respondents were less willing to leave the mess and explain it to the guest (option 2), leave the mess and take the guest somewhere else (option 3), and call the guest to reschedule and ask the roommates to clean up the mess (option 5). For the strong situation, where the guest is arriving in 15 minutes, is expected to judge the person for the mess, and the roommates are known to react negatively to criticism, more agreeable respondents were less willing to clean up as much as possible and then speak to the roommates upon their return to show the guest that they were concerned about the mess (option 1).

Although there were no significant differences in how the overall sample rated any of the response options as a function of situational strength, those low on agreeableness (1 SD below the mean) who responded to the strong situation rated calling the guest to reschedule and asking the roommates to clean up the mess (option 5) as less likely (M = 1.75, SD = .91) than those low on agreeableness who responded to the weak situation (M = 2.62, SD = 1.23), $t(60.48) = 3.34$, $p < .001$. Furthermore, those high on agreeableness (1 SD above the mean) who responded to the strong situation rated leaving the mess and explaining it to the guest (option 2) as more likely (M = 2.16, SD = 1.08) than did respondents high on agreeableness who responded to the weak situation (M = 1.50, SD = .92), $t(58) = -2.51$, $p < .05$. In light of the higher time constraints associated with the strong version of the situation, the situational press felt by at least certain groups to not try to reschedule the meeting and to leave the mess (and explain it to the guest) makes sense and is consistent with the situational strength hypothesis.

As pointed out above, not all the situational strength effects observed at the level of the response ratings are ultimately reflected in the item scores. For item 7, options 1 and 3 are keyed '0' for both the weak and strong SJT, so the fact that ratings on these options differentially reflect agreeableness as a function of test strength is not captured by the scores for item 7.

Table 12. **Means and correlations (with agreeableness) for SJT response ratings**

	Weak SJT				Strong SJT					
Response Option	Key	*r* with Agreeableness	*M*	*SD*	Key	*r* with Agreeableness	*M*	*SD*	*df*	*t*
SJT6.Op1	1	.26***	4.86	.37	1	.21**	4.82	.48	375.37	1.07
SJT6.Op2	0	-.11	2.50	1.41	0	.00	2.08	1.19	397.84	3.31**
SJT6.Op3	0	-.03	1.72	1.06	0	-.24***	1.61	.92	401.82	1.19
SJT6.Op4	0	.17*	2.90	1.19	0	-.10	2.58	1.18	403	2.74**
SJT6.Op5	-1	-.22**	2.05	1.14	-1	-.03	2.04	1.05	405	.16
SJT6.Op6	0	.130	3.67	1.13	0	.05	3.43	1.20	405	2.08*
SJT7.Op1	0	-.10	3.07	1.21	0	-.16*	3.24	1.27	403	-1.38
SJT7.Op2	-1	-.20**	1.83	1.08	-1	.00	1.93	1.03	405	-.99
SJT7.Op3	0	-.20**	2.95	1.30	0	-.06	2.88	1.27	405	.52
SJT7.Op4	1	.28***	4.46	.86	1	.13	4.41	.91	403	.56
SJT7.Op5	-1	-.27***	2.07	1.13	0	.06	2.01	.99	399.58	.59
SJT8.Op1	1	.19**	4.30	.88	1	.13	4.76	.49	324.47	-6.46***
SJT8.Op2	0	-.08	3.29	1.24	-1	.08	2.67	1.25	404	5.03***
SJT8.Op3	-1	-.15*	1.72	.86	-1	.04	1.86	.85	404	-1.64
SJT8.Op4	0	-.05	3.75	1.07	-1	-.12	3.17	1.28	388.97	5.04***
SJT8.Op5	0	-.04	3.08	1.31	0	.07	2.86	1.28	402	1.70
SJT16.Op1	0	-.10	2.84	1.24	0	-.05	2.55	1.31	405	2.28*
SJT16.Op2	-1	-.25***	2.06	1.06	-1	-.21**	1.63	.89	405	4.52***
SJT16.Op3	0	.22**	4.38	.80	0	-.01	4.49	.72	406	-1.34
SJT16.Op4	1	.32***	4.33	.82	1	.13	4.59	.66	389.70	-3.55***
SJT16.Op5	0	.020	2.63	1.07	0	-.08	3.08	1.17	405	-4.09***

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

It is appropriate to consider more broadly the mechanism underlying the statistically significant interaction of openness and agreeableness with situational strength on the interpersonal SJT scores. It was proposed that there should be more variability in individuals' responses to a weak SJT relative to a strong. Table 13 shows the variability and reliability of SJT scores based on the weak test relative to those based on the strong test. Consistent with expectations, Levene's test for equality of variances indicates that there is significantly more variability in interpersonal SJT scores for the weak version of the test relative to the strong (F = 8.09, $p < .01$). The interpersonal SJT score based on weak situations is also slightly more reliable ($\alpha = .70$) than the interpersonal SJT score based on strong situations ($\alpha = .60$). Looking at the intrapersonal SJT dimension, variability and reliability were not higher for the weak SJT than for the strong.

Table 13. **SJT score variability and reliability as a function of situational strength**

	SD		Internal Consistency (α)	
Variable	Weak SJT	Strong SJT	Weak SJT	Strong SJT
Rating-based Interpersonal SJT Score	6.23	5.02	.70	.60
Rating-based Intrapersonal SJT Score	5.84	5.73	.64	.71
Rating-based Overall SJT Score	9.98	9.31	.73	.76

Note. *N* ranges from 161 to 196.

There was no statistically significant interaction of extraversion and situational strength on either the interpersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 373) = .06$, *ns*, or the intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = .92$, *ns*. There was also no statistically significant interaction of emotional stability and situational strength on either the interpersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 373) = .00$, *ns*, or the intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = .00$, *ns*.

The low correlations of SJT scores with extraversion ($r_{\text{extraversion-interpersonal SJT score}} = .08$, *ns*; $r_{\text{extraversion-intrapersonal SJT score}} = .04$, *ns*) and with emotional stability ($r_{\text{emotional stability-interpersonal SJT score}} = .08$, *ns*; $r_{\text{emotional stability-intrapersonal SJT score}} = .04$, *ns*) indicate that overall, the current SJT did not capture extraversion or emotional stability very well. Looking at the correlations of various response option ratings with extraversion and emotional stability, ratings on either one or both versions of the SJT were significantly correlated with extraversion in only 28% of cases (i.e., for 29 out of the 102 total response options) and emotional stability in only 22% of cases (i.e., for 22 response options). It was, however, somewhat more common for response option ratings on the weak SJT to be significantly correlated with extraversion or emotional stability and for response option ratings on the strong SJT not to be (16% of cases for extraversion and 14% of cases for emotional stability) than for the reverse to be true (9% of cases for extraversion and 4% of cases for emotional stability). These effects at the level of the response ratings—for the 16% of cases for extraversion and the 14% of cases for emotional stability—that are consistent with the situational strength hypothesis are not all reflected in SJT scores; 38% of the cases for extraversion and 43% of the cases for emotional stability that support the situational strength hypothesis are associated with response options keyed '0' for both the weak and strong SJT. Thus, not only were the responses poor measures of extraversion and emotional stability to begin with, but the scoring keys further reinforce that the more extraverted or emotionally stable responses are often not the more "correct" ones. All in all, the SJT design elements were stacked against finding interactions of situational strength with extraversion and with emotional stability and for predicting SJT scores.

Finally, there was no statistically significant interaction of conscientiousness and situational strength on either the interpersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 373) = .24$, *ns*, or the intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 361) = 1.38$, *ns*. The reasons for this lack of effects are apparently different than those for extraversion and emotional stability described above. SJT scores *do* reflect conscientiousness ($r_{\text{conscientiousness-interpersonal SJT score}} = .31$, $p < .001$; $r_{\text{conscientiousness-intrapersonal SJT score}} = .32$, $p < .001$). Looking at the correlations of various response option ratings with conscientiousness, ratings on either one or both versions of the SJT were significantly correlated with conscientiousness in 47% of cases. However, it was unexpectedly somewhat more common for response option ratings on the *strong* SJT to be significantly correlated with conscientiousness and for response option ratings on the weak SJT not to be (17% of cases) than for the reverse to be true (10% of cases). These effects at the level of the response ratings—for the 17% of cases—that are *inconsistent* with the situational strength hypothesis, were not all translated into SJT scores; 41% of these cases that are inconsistent with the situational strength hypothesis are associated with response options keyed '0' for both the weak and strong SJT. It seems possible then, that had the more conscientious responses been more often keyed as "correct" (or even incorrect), conscientiousness may have significantly interacted with situational strength to predict SJT scores. However, the stronger SJT could have unexpectedly been the better measure of conscientiousness.

Research question 2. Research question 2 asked whether situational strength would moderate the relation between scores on the SJT and scores on a measure of socially desirable responding. Impression management and situational strength were entered in the first step of the regression. The interaction of impression management and situational strength was entered in the

second step. Impression management scores were centered prior to entry in the regression. Two versions of the overall SJT score (score based on Most Likely/Least Likely responses; score based on ratings) as well the ratings-based interpersonal and intrapersonal SJT scores were examined as outcomes.

There was no statistically significant interaction between impression management scores and situational strength on ML/LL-based SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 379) = .03$, *ns* , on ratings-based SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 334) = .90$, *ns*, on interpersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 372) = .76$, *ns*, and on intrapersonal SJT scores, $\Delta R^2 = .00$, $\Delta F(1, 360) = .72$, *ns*. Thus, situational strength apparently did not moderate the relation between SJT scores (overall or dimension-specific) and socially desirable responding.

Incremental validity of the SJT. Hierarchical regressions were used to examine the incremental validity of the SJT above HS GPA and SAT/ACT scores for predicting criteria. Two versions of the SJT score (score based on Most Likely/Least Likely responses; score based on ratings) and three different outcomes (organizational citizenship; first year college GPA; BARS) were examined. HS GPA and ACT scores were entered in the first step of the regression. HS GPA, ACT scores, and SJT scores were entered in the second step.

HS GPA and ACT scores alone explained 2.8% of the variance in organizational citizenship scores, $F(2, 365) = 5.29$, $p < .01$. HS GPA was positively related to organizational citizenship scores ($b = .17$, $p < .01$), whereas ACT scores were not significantly related to organizational citizenship scores ($b = .00$, *ns*). ML/LL-based SJT scores had incremental validity over HS GPA and ACT scores for predicting organizational citizenship scores, $\Delta R^2 = .06$, $\Delta F(1,$

364) = 23.10, $p < .001$. ML/LL-based SJT scores were positively related to organizational citizenship scores ($b = .25$, $p < .001$).

When ratings-based SJT scores were entered into the regression in place of ML/LL-based SJT scores, SJT scores still had incremental validity over HS GPA and ACT scores for predicting organizational citizenship scores, $\Delta R^2 = .05$, $\Delta F(1, 323) = 15.88$, $p < .001$. Ratings-based SJT scores were positively related to organizational citizenship scores ($b = .22$, $p < .001$).

HS GPA and ACT scores alone explained 28.4% of the variance in first year college GPA, $F(2, 365) = 72.30$, $p < .001$. HS GPA and ACT scores were both positively related to first year college GPA ($b = .35$, $p < .001$ and $b = .31$, $p < .001$, respectively). ML/LL-based SJT scores had incremental validity over HS GPA and ACT scores for predicting first year college GPA, $\Delta R^2 = .03$, $\Delta F(1, 364) = 13.01$, $b = .16$, $p < .001$. When ratings-based SJT scores were entered into the regression in place of ML/LL-based SJT scores, SJT scores still had incremental validity over HS GPA and ACT scores for predicting first year college GPA, $\Delta R^2 = .02$, $\Delta F(1, 323) = 9.02$, $b = .14$, $p < .01$.

HS GPA and ACT scores alone did not explain a significant amount of the variance in BARS scores, $R^2 = .01$, $F(2, 365) = 2.42$, $p = .09$. ML/LL-based SJT scores had incremental validity over HS GPA and ACT scores for predicting BARS scores, $\Delta R^2 = .14$, $\Delta F(1, 364) = 60.00$, $b = .38$, $p < .001$. When ratings-based SJT scores were entered into the regression in place of ML/LL-based SJT scores, SJT scores still had incremental validity over HS GPA and ACT scores for predicting BARS scores, $\Delta R^2 = .21$, $\Delta F(1, 323) = 86.91$, $b = .47$, $p < .001$.

Relative mean score differences. T-tests were used to examine gender- and race-based score differences on the ML/LL-based and ratings-based SJT scores relative to group score differences on the more cognitive predictors. There were no significant difference between men (M = 3.65, SD = .32) and women (M = 3.67, SD = .30) on HS GPA, $t(387) = .35$, *ns*. On average, men (M = 25.74, SD = 2.77) had higher ACT scores than women (M = 24.61, SD = 3.30), $t(395) = -2.86$, $p < .01$, $d = -.35$. In contrast, women (M = 18.26, SD = 4.99) scored higher than men (M = 15.91, SD = 4.95) on the ML/LL-based SJT, $t(379) = 3.74$, $p < .001$, $d = .47$. Women (M = 127.42, SD = 9.82) likewise scored higher than men (M = 122.48, SD = 9.25) on the ratings-based SJT, $t(334) = 3.92$, $p < .001$, $d = .51$.

On average, White respondents (M = 3.70, SD = .27) had significantly higher HS GPA than minority respondents (M = 3.50, SD = .40), $t(376) = -4.79$, $p < .001$, $d = -.68$. White respondents (M = 25.37, SD = 2.89) likewise had significantly higher ACT scores than minority respondents (M = 22.60, SD = 3.72), $t(383) = -6.69$, $p < .001$, $d = -.91$. In contrast, there were no significant differences between White (M = 17.89, SD = 4.91) and minority (M = 17.44, SD = 5.70) respondents' ML/LL-based SJT scores, $t(368) = -.66$, *ns*, and between White (M = 126.58, SD = 9.71) and minority (M = 124.75, SD = 10.76) respondents' ratings-based SJT scores, $t(325) = -1.18$, *ns*.

Mean score differences on weak and strong SJT. T-tests were used to examine score differences on the weak and strong versions of the SJT separately. When examining the weak version of the SJT, there were marginally significant differences between women's (M = 16.90, SD = 4.77) and men's (M = 15.42, SD = 4.00) ML/LL-based SJT scores, $t(192) = 1.86$, $p = .06$, $d = .32$, and between women's (M = 125.11, SD = 10.24) and men's (M = 121.77, SD = 9.06) ratings-based SJT scores, $t(173) = 1.88$, $p = .06$, $d = .33$. When examining the strong version of

the SJT, women's ML/LL-based SJT scores (M = 19.62, SD = 4.86) were significantly higher than men's (M = 16.49, SD = 5.88), $t(185) = 3.37$, $p < .01$, $d = .62$; women's ratings-based SJT scores (M = 129.89, SD = 8.74) were also significantly higher than men's (M = 123.28, SD = 9.53), $t(159) = 3.91$, $p < .001$, $d = .74$.

When examining the weak version of the SJT, there was no significant difference between White (M = 16.59, SD = 4.38) and minority (M = 16.29, SD = 5.62) respondents' ML/LL-based SJT scores, $t(185) = -.35$, *ns*, and between White (M = 124.61, SD = 9.59) and minority (M = 122.78, SD = 11.70) respondents' ratings-based SJT scores, $t(167) = -.86$, *ns*. When examining the strong version of the SJT, there was likewise no significant difference between White (M = 19.20, SD = 5.07) and minority (M = 18.74, SD = 5.58) respondents' ML/LL-based SJT scores, $t(181) = -.46$, *ns*, and between White (M = 128.64, SD = 9.43) and minority (M = 127.19, SD = 9.16) respondents' ratings-based SJT scores, $t(156) = -.66$, *ns*.

Exploratory analyses. Given some of the group differences in perceptions of situational strength (see discussion of zero order correlations above), further analyses were undertaken to examine whether the strength manipulation may have worked on certain subgroups (defined by race, agreeableness levels, or ACT levels) within the sample. That is, it is possible that certain groups *did* perceive the strong SJT as being significantly higher in strength than the weak SJT, despite the fact that no differences were seen in the overall sample.

First, t-tests indicated no differences in the effectiveness of the manipulation for Whites and minorities. Although overall strength perception differences were in the expected direction for both Whites and minorities, neither group perceived significant differences between the weak (M = 3.55, SD = .56 and M = 3.73, SD = .57, for Whites and minorities respectively) and strong

SJT (*M* = 3.60, *SD* = .62 and *M* = 3.85, *SD* = .53, for Whites and minorities respectively), *t*(323) = -.73, *ns*, and *t*(66) = -.86, *ns*, respectively.

Second, t-tests indicated that respondents below average on agreeableness (1.50-4.11 range) reported higher overall strength perceptions for the strong SJT (*M* = 3.66, *SD* = .55) than for the weak SJT (*M* = 3.50, *SD* = .47), *t*(197) = -2.20, *p* < .05, whereas respondents above average on agreeableness (4.20-5.00 range) did not report different overall strength perceptions of the weak (*M* = 3.69, *SD* = .62) and strong SJT (*M* = 3.61, *SD* = .66), *t*(207) = .83, *ns*. Thus, the strength manipulation worked as expected only for those below average on agreeableness.

Third, t-tests indicated that respondents with above average ACT scores (25-36 range) reported higher overall strength perceptions for the strong SJT (*M* = 3.72, *SD* = .63) than for the weak SJT (*M* = 3.49, *SD* = .53), *t*(220) = -2.91, *p* < .01, whereas respondents with below average ACT scores (15-24 range) reported slightly higher overall strength perceptions for the weak SJT (*M* = 3.74, *SD* = .57) than for the strong SJT (*M* = 3.58, *SD* = .54), *t*(176) = 1.94, *p* = .054. Thus, the strength manipulation worked as expected only for those with above average ACT scores.

ACT scores, in contrast to agreeableness scores, had low or non-significant associations with the personality variables (*r* with openness = .13, *p* < .05; *r* with extraversion = -.19, *p* < .01; *r* with conscientiousness = .01, *ns*; *r* with agreeableness = .06, *ns*; *r* with emotional stability = .02, *ns*), rating-based SJT scores (*r* with intrapersonal SJT score = -.15, *p* < .01; *r* with interpersonal SJT score = .11, *p* < .05), and two of the three outcome variables in this study (*r* with citizenship = .05, *ns*; *r* with BARS = .03, *ns*). In light of this information, interactions of personality scores and situational strength on SJT scores and interactions of SJT scores and situational strength on outcome scores could be calculated on just the individuals with high ACT scores (*N* = 222; for whom the strength manipulation "worked") with minimal repercussions for

indirect range restriction on the personality, SJT score, and outcome variables of interest. Indeed, as can be seen in Table 14, restricting the sample to just those with above average ACT scores has minimal impact on the means, standard deviations, and reliabilities of the personality, SJT, and outcome measures. As expected, the largest reduction in score variability occurs for first year college GPA, given that college GPA is strongly associated with ACT scores in this study.

Considering that ACT scores were related to the amount of attention respondents paid to the survey—freshmen with higher ACT scores failed fewer attention checks (r = -.16, p < .001)—excluding additional respondents from analyses based on a below average ACT score may serve to screen out those who were responding to the survey inattentively but nonetheless noticed and responded correctly to the attention check questions.

Table 14. **Variables' measurement properties for high ACT group as compared to the full sample**

	Full Sample			High ACT Group			
Variable	*M*	*SD*	α	*M*	*SD*	α	$SD_{High\ ACT}/SD_{Full}$
Openness	3.69	.54	.81	3.76	.53	.80	.98
Extraversion	3.31	.78	.89	3.22	.81	.90	1.04
Conscientiousness	3.75	.62	.83	3.77	.66	.86	1.07
Agreeableness	4.12	.53	.80	4.16	.54	.83	1.04
Emotional Stability	3.10	.69	.84	3.11	.72	.85	1.04
Rating-based Interpersonal SJT Score	67.04	5.72	.65	67.66	5.68	.66	.99
Rating-based Intrapersonal SJT Score	55.37	6.53	.69	54.80	6.42	.65	.98
BARS	5.02	.67	.77	5.03	.67	.78	1.00
Citizenship	3.68	.60	.77	3.72	.59	.78	.98
First Year College GPA	3.31	.53	--	3.46	.47	--	.88

Note. *N* ranges from 365 to 408 for the full sample and from 200 to 222 for the High ACT Group. High is defined as above average.

Hypothesis 1b, and research questions 1 and 2 were reexamined using just the subsample of respondents with above average ACT scores. Results did not differ from those using the larger sample. Hypothesis 1b, which predicted interactions between SJT scores and situational strength on criteria, was still not supported: neither interpersonal SJT scores nor intrapersonal SJT scores interacted with situational strength to predict BARS scores, citizenship scores, or first year college GPA. Pertaining to research question 1, there was again a significant interaction of openness and situational strength on interpersonal SJT scores, $\Delta R^2 = .02$, $\Delta F(1, 206) = 5.57$, $p < .05$, and a significant interaction of agreeableness and situational strength on interpersonal SJT scores, $\Delta R^2 = .03$, $\Delta F(1, 206) = 7.56$, $p < .01$. The shapes of these interactions match those of the interactions computed using the larger sample. Finally analyses to address research question 2, which concerned the interactions of impression management scores and situational strength on SJT scores, again showed no significant interaction of impression management and situational strength on SJT scores. Next, study findings are summarized and discussed in terms of their theoretical and practical implications.

DISCUSSION

Summary of Study Findings

The goal of the current study was to extend existing knowledge and the practical value of SJTs by considering whether the situational strength of scenarios comprising an SJT has implications for the test's construct and criterion-related validity. Related to construct validity, results indicate that changing the strength of a situation (by changing its clarity, consistency, constraints, and/or consequences) can change the interpretation of that situation and its associated response options. This is evidenced by the fact that, as per experts' judgments, over half of the final eighteen SJT items had at least some response options keyed differently depending on the SJT version, as well as the fact that the pattern of correlations between respondents' likelihood ratings on the various response options and personality variables changed as a function of items' situational strength. These effects, however, did not translate into differences in SJT scores' criterion-related validity for the outcomes in this study. The results related to the overall SJT and situational strength effects on the SJT scores' relations with other variables are summarized and discussed below.

SJT construct validity and dimensionality. Looking across test versions, the SJT in this study was a relatively better measure of openness, agreeableness, and conscientiousness than of extraversion and emotional stability. SJT scores also had strong associations with the various biodata scales designed to tap the same dimensions as the SJT, particularly with the perseverance- and ethics-related biodata scales. A significant association between SJT and impression management scores suggests that the SJT may have been capturing individual differences in impression management as well. This is not surprising given that SJTs with

behavioral tendency instructions (which measure personality more so than knowledge) can be susceptible to impression management (Hooper et al., 2006).

The behavioral dimensions that the SJT was designed to assess—adaptability, citizenship, perseverance, and ethics—had low internal consistency reliabilities and an alternative factor solution was sought. The task of coming up with a common (across SJT versions) interpretable factor solution for the scored SJT items was made particularly difficult by the fact that the two versions of the SJT did not capture the same individual differences or to the same extent. Interpersonal and intrapersonal SJT dimensions were extracted from the data for use in analyses as these dimensions made theoretical sense, had close to acceptable reliability, and had interpretable associations with individual difference variables, but the two factor solution did not fit the data particularly well.

SJT criterion-related validity. Consistent with earlier research (Schmitt et al., 2009; Schmitt et al., 2010), the current SJT had significant criterion-related validities with both academic and non-academic criteria. Out of the three criteria, SJT scores (ratings-based) explained the largest percent of the variance (22.1%) in BARS scores, followed by citizenship scores (6.0%), and first year college GPA (4.1%). The opposite trend was seen for the more cognitive predictors. As a set, high school GPA and ACT scores explained the largest percent of variance (28.4%) in first year college GPA, followed by citizenship scores (2.8%), and BARS scores (1.3%, $p = .08$). These findings are consistent with research suggesting that although cognitive measures do a great job predicting initial academic performance (Buyse, 2011; Robbins et al., 2004), they may not effectively predict students' performance in the non-academic areas (e.g., integrity and social responsibility) that contribute to their success in school (Schmitt et al., 2009; Shultz & Zedeck, 2011).

SJT scores (ratings-based) explained an additional 1.9% ($p < .01$) of the variance in first year college GPA beyond high school GPA and ACT scores. As discussed earlier, McDaniel et al.'s (2007) meta-analysis estimated that the incremental validity of SJTs above cognitive ability falls within the 3% to 5% range. Taking high school GPA (past performance) out of the equation and leaving just ACT (proxy for cognitive ability) and SJT scores, results in the current SJT accounting for an additional 4.6% of the variance in first year college GPA above ACT scores.

Overall SJT group differences. Whereas men in this study had an advantage over women on ACT scores, consistent with existing research (Whetzel et al., 2008), women tended to score higher on the SJT than did men. Interestingly, women had only a marginal advantage over men on the weak SJT, but a significant advantage over men on the strong SJT. It is possible that women were more attentive and responsive to the cues in the strong version of the SJT relative to men, as research does indicate that women may be more sensitive to situational cues (e.g., Grand, Golubovich, Ryan, & Schmitt, 2013; Mael, Connerley, & Morath, 1996). However, men and women did not report significantly different strength perceptions of the SJT they completed (for either the weak or strong SJT), and women who completed the strong SJT did not report higher strength perceptions than women who completed the weak SJT. These results might be seen as counteracting the argument that women may have been more attentive to the cues in the strong SJT. Importantly though, because the strength perception items referred to the entire *set* of situations in the SJT, these items simply may have failed to capture differences in how men and women perceived the strength of the *individual* SJT items. This particular issue is further discussed below.

Consistent with research showing minimal racial group differences in SJT scores (Whetzel et al., 2008), there were no significant differences between Whites' and minorities'

scores on the current SJT. Neither the weak nor the strong version of the test significantly disadvantaged minorities. Minority-White score differences on SJTs tend to differ depending on the specific minority group being considered (Whetzel et al., 2008), but there were not enough minorities in the current sample to examine specific groups. Whereas the SJT did not disadvantage minorities, both high school GPA and ACT scores did.

Manipulation check. The effectiveness of the situational strength manipulation in this study was checked by comparing the situational strength perceptions of individuals who responded to the weak SJT with the perceptions of those who responded to the strong SJT. The overall sample in this study did not perceive differences between the two tests in global strength or any of the individual facets of strength. On the one hand, this indicates that the manipulation was fairly weak. This was to be expected as an effort was made not to make situations so strong as to make the SJT items impractical for use in a *test*. On the other hand, as the manipulation check asked about the test items overall, it may not have been fine-grained enough to capture individuals' differential strength perceptions of the individual items in the SJT. Given that items within a test form varied with regard to each particular facet of strength (e.g., consistency was not even manipulated in most items), responding about the items as a set required that respondents aggregate across their perceptions of individual items to make a summary judgment about them. It is unknown how respondents formed these judgments and to what extent they were able to recall the various items well enough to make accurate judgments.

Ultimately, situational strength effects were found in spite of the fact that the manipulation check failed to work, indicating that, even though respondents were unable to report perceiving strength differences, they nonetheless apparently reacted to them. Furthermore, when those with higher ACT scores, for whom the manipulation check *did* work, were isolated

and analyzed separately, results did not differ from those based on the larger sample, for whom the manipulation check did not work. Situational strength effects in this study are discussed next.

Situational strength effects on construct validity. In partial support of the situational strength hypothesis that weak situations better allow for the expression of individual differences than do strong situations, weak interpersonal SJT scores were better measures of openness and agreeableness than were strong interpersonal SJT scores. There were, however, no significant differences between the extent to which weak and strong interpersonal SJT scores reflected extraversion, emotional stability, or conscientiousness, and no significant differences in the extent to which weak and strong intrapersonal SJT scores reflected any of the measured personality traits. SJT scores on the weak test (relative to the strong one) also failed to demonstrate a stronger relation with a biodata measure designed to assess the same dimensions as the SJT.

Although the likelihood ratings that respondents provided for the various SJT response options tended to better reflect individual differences in openness, agreeableness, extraversion, and/or emotional stability for those completing the weak SJT relative to the strong, these effects could only be translated into SJT scores in cases where the particular response option subject to these effects was keyed as being correct or incorrect. Recall that ratings-based SJT scores were computed by summing the ratings of response options keyed as being correct with the reverse coded ratings of response options keyed as incorrect; ratings of responses keyed as neither correct nor incorrect were not used in computing SJT scores. The results seem to indicate that the current SJT was a good enough measure of openness and agreeableness that enough of the situational strength effects at the level of response option ratings were ultimately translated in item scores to observe significant effects of situational strength in the scored items. In contrast,

the current SJT was not a good measure of emotional stability and extraversion to begin with, so this may have made it more difficult to find situational strength effects for these personality variables.

Conscientiousness was unique from the other four personality variables in that likelihood ratings individuals provided for responses on the strong SJT tended to reflect their individual differences in conscientiousness better than did ratings individuals provided for response options on the weak SJT. Although the interaction of conscientiousness and situational strength on SJT scores was not statistically significant, the trend toward effects counter to the situational strength hypothesis seems worthy of discussion because these results also run counter to meta-analytic findings indicating that conscientiousness predicts behavior better in weak situations than in strong ones (Meyer et al., 2009).

Trait activation theory (e.g., Tett & Burnett, 2003) may help to explain the current findings with regard to conscientiousness. It argues that a trait will not become manifested in behavior unless cues are available to indicate that the trait is relevant for that particular situation (e.g., Tett & Burnett, 2003). Given that conscientious individuals tend to be more achievement-oriented, responsible, and willing to comply with rules and expectations (Barrick & Mount, 1991; Roberts, Chernyshenko, Stark, & Goldberg, 2005), it is possible that conscientiousness, in order to be "activated", requires that a situation provide a fairly high level of clarity around rules and performance expectations. In other words, conscientious individuals may need to know what is expected of them before they can direct their resources toward meeting those expectations. For this reason, items that tried to manipulate clarity along with the other facets of strength rather than providing an adequate level of clarity for both the weak and strong version of the situation

may have contributed to the unexpected finding that stronger (more clear) situations tended to elicit conscientiousness somewhat better than the weaker ones.

Meyer et al. (2009), who reported that the criterion-related validity of conscientiousness is higher for weak occupations relative to strong had actually focused their analyses on just the constraints and consequences dimensions of strength. It is probably the case that most organizations tend to set out fairly clear rules and performance expectations for their workers (thereby providing for an adequate baseline level of clarity), whereas there may be considerably more variability across jobs (and occupations) in how constrained and consequential workers' workplace behaviors are.

Situational strength effects on the variability and reliability of SJT scores. The situational strength hypothesis would suggest that there should be more variability in individuals' responses to a weak SJT relative to a strong (Meyer et al., 2009). Although this was not evident at the level of the overall SJT score, the interpersonal score, which reflected agreeableness and openness to a higher degree for the weak test relative to the strong, relatedly had better variability and internal consistency for the weak SJT than for the strong. The intrapersonal SJT score, on the other hand, which was a slightly better measure of conscientiousness for the strong test relative to the weak, relatedly had a slightly better internal consistency for the strong SJT than for the weak. Thus, there was partial support for the mechanism believed to underlie situational strength effects.

Situational strength effects on criterion-related validity. Situational strength did not moderate the relations of SJT scores with BARS, citizenship behaviors, or first year college GPA. The non-significant interactions are relatively more surprising for interpersonal SJT scores

(versus intrapersonal scores) as these were found to be more reliable for the weak test relative to the strong and exhibited more variability.

The lack of a significant interaction between interpersonal SJT scores and situational strength on citizenship scores is particularly notable in light of the strong relation between agreeableness and citizenship scores in this study, supporting research showing the importance of agreeableness for prosocial behavior (Graziano & Eisenberg, 1997; Smith & Nelson, 1975), and the fact that the weak interpersonal dimension was a better measure of agreeableness relative to the strong. Relatedly, given the strong relations of openness and agreeableness with BARS scores in the current study, as well as the fact that the weak interpersonal SJT score was a better measure of both these personality traits relative to the strong, interpersonal SJT scores should also have interacted with situational strength in predicting BARS scores. However, this interaction was only marginally significant and, counter to expectations, the interpersonal SJT score-BARS relation was actually stronger for the strong version of the SJT relative to the weak.

Situational strength effects on impression management. Even though there was a significant relation between SJT and impression management scores, this relation did not vary as a function of situational strength. This suggests that at the overall test level, the situations on the strong test may not have been more susceptible to faking than those on the weak test. Given that being able to discern the relevant performance dimensions is likely a prerequisite for the ability to manage one's impression in a particular situation (Jansen et al., 2013), higher strength may not have made the relevant performance dimensions of the strong SJT more easily discernible. Changing the strength of SJT situations may not affect the inherent multidimensionality of those situations and the associated response options.

Individual differences in perceptions of strength. The current study evidenced some individual differences in perceptions of situational strength. Minority respondents (relative to Whites) reported higher perceptions of clarity and constraints, more agreeable respondents reported higher perceptions of consequences, and respondents with higher ACT scores reported lower perceptions of clarity and constraints.

Although minorities tended to perceive the situations in the SJT to be stronger, they did not have more "correct" perceptions of strength than Whites. That is, neither group perceived significant strength differences between the weak and strong SJT. On the other hand, individuals of below average agreeableness and those with above average ACT scores did have more correct perceptions of strength than their corresponding groups. That is, respondents below average on agreeableness and those with above average ACT scores perceived the strong SJT to be significantly stronger than the weak SJT. Corresponding groups—respondents above average on agreeableness and those with below average ACT scores—did not report significantly different overall strength perceptions of the weak and strong SJT.

The fact that students with higher ACT scores were more accurate at judging the situational strength of the SJT (as defined by expert consensus) is consistent with a body of research indicating that smarter people tend to be better at the information-processing activities involved in effectively reading and understanding social situations (Jansen et al., 2013). Findings with regard to the role of agreeableness and other personality variables in situational assessment are more inconsistent (Jansen et al., 2013). Results pertaining to minorities' higher perceptions of situational clarity and constraints but inability to perceive the differences between the weak and strong SJT can be interpreted in light of existing research indicating that minorities may be more reactive to ambiguous cues in test items and not perceive these items the way they are

consensually defined (Grand et al., 2013). The theoretical and practical implications of this study's findings are discussed next.

Theoretical Implications

Scholars in the area of individual differences have often called for more studies that approach the prediction of behavior from an interactionist perspective, where the target behavior is conceptualized as being a function of both the person and the particular situation (e.g., Endler & Magnusson, 1976; Hattrup & Jackson, 1996; Pervin, 1985). Although there is still no agreed upon taxonomy of situations that would help scholars navigate this research domain (Johns, 2006), the importance of situational strength is something that scholars do seem to agree on.

Earlier research has tended to examine situational strength's role as a moderator of relations between individual differences and criteria (e.g., Beaty et al., 2001; Withey et al., 2005; Smithikrai, 2008). The current study adds to this body of research by examining situational strength's role as a moderator in the measurement of individual differences via an SJT and that measure's criterion-related validity. This constitutes an extension of research on situational strength to SJTs, as the implications of situational strength for SJTs or even related measurement methods like situational interviews have received little research attention to date.

Relatively little is known about the variables that moderate SJTs' validity. While a large number of studies have sought to delineate the design factors that can impact test takers' SJT item interpretation and response process, relatively more attention has been paid to the implications of test-level (e.g., response instructions, delivery medium, rating format) as opposed to item-level characteristics and cues (e.g., item specificity, item complexity). The current study suggests that situational strength can affect test takers' interpretations of and responses to an SJT. Relatedly, manipulating the strength of a situation can change the nature of the construct(s)

targeted by the SJT item. When there is uncertainty regarding the construct(s) SJT items measure to begin with, manipulations of strength can present test designers with additional interpretive challenges. Current findings suggest that strength's implications for SJTs' construct and criterion-related validity deserve further attention.

Much of the existing research on situational strength has been criticized for failing to explicitly manipulate or measure the construct as it was conceptualized by Mischel (1977) and for frequently using situational strength for convenient post hoc interpretation of findings (Cooper & Withey, 2009). The current study drew on recent work delineating the facets of strength (Hattrup & Jackson, 1996; Meyer et al., 2010) and providing for a means of measuring them (Meyer et al., in press) to explicitly manipulate the strength of an SJT as well as to measure respondents' perceptions of that strength.

An advantage of the situational strength manipulation in this study was that every attempt was made to hold the length and readability of corresponding SJT stems in the two versions of the test approximately constant. This was a way to ensure that item strength would not be confounded with item length or readability, as items that are longer and more difficult to read are likely to be more cognitively loaded and have higher criterion-related validity (at least with a measure of academic performance). This enhances the internal validity of the study.

While the situational strength hypothesis implies that a weaker SJT should be superior because it will better allow for the expression of individual differences in behavioral intentions, researchers have speculated (e.g., Brooks & Highhouse, 2006; Jansen et al., 2013), and some studies have found (e.g., Reynolds et al., 1999, 2000; Lievens, De Corte, & Schollaert, 2008; Schmit, Ryan, Stierwalt, & Powell, 1995) that lower ambiguity and greater contextualization (at least up to a point) are preferable as they tend to result in assessments with better criterion-

related validity. It is argued that respondents will be better able to draw on relevant personal experiences and accurately present themselves if they interpret items in the way the test developer, in attempting to design a predictor with conceptual overlap with criteria, or experts interpreted them (Brooks & Highhouse, 2006; Jansen et al., 2013; Lievens et al., 2008).

One way to potentially reconcile predictions about the advantages of lower strength, one dimension of which is clarity, with the idea that it is desirable to have a fairly high level of clarity in assessments is to examine whether the clarity facet of strength should be thought of and treated differently than strength's other dimensions. Specifically, it may be the case that consequences and constraints most impact on the expression of individual differences (conscientiousness is particular) when situations provide a fairly high level of clarity around rules and expectations. Indeed, the majority of research in the area of situational strength has apparently focused on the constraints (often operationalized as job autonomy) and consequences dimensions of strength (Meyer et al., 2009).

The consistency dimension of strength may require special treatment as well. Given the strong relation between clarity and consistency (respondents who perceived the SJT to have higher consistency also perceived it to have higher clarity; $r = .77$; $p < .001$)—a relation that is also evident in earlier research (Meyer et al., in press)—manipulations of consistency may work more effectively as a way to affect situational clarity rather than as an independent dimension of strength in its own right.

Practical Implications

Validity and fairness of SJTs. The current study provides additional evidence to suggest that SJTs may effectively supplement more cognitive measures in an admissions context. Current findings add to an ever growing body of research that demonstrates SJTs' validity for predicting

academic performance as well as performance in other nonacademic areas like social responsibility and leadership that contribute to students' overall success (e.g., Lievens et al., 2005; Lievens & Sacket, 2011; Mumford et al., 2008; Oswald et al., 2004; Schmitt et al., 2009; Schmitt et al., 2010). In line with earlier research, the current SJT showed significant relations with first year college GPA, self-rated academic and nonacademic performance, and citizenship behaviors. The SJT was also incrementally valid above high school GPA (past performance) and standardized test scores, indicating that there can be value in adding an SJT to a selection battery that includes these more traditional predictors of college student success.

An attractive quality of SJTs is their tendency to evidence lower subgroup differences than more cognitive measures (Whetzel et al., 2008). This is particularly true of SJTs that use behavioral tendency, as opposed to knowledge, response instructions as the former result in better assessment of personality than cognitive ability (McDaniel et al., 2007; Whetzel et al., 2008). The current study contributes to the body of research on SJTs with behavioral tendency response instructions, showing that it is possible to design these tests to not have significant racial subgroup score differences.

One of the tradeoffs potentially made when opting for an SJT with behavioral tendency instructions for reasons that include minimizing racial subgroup differences, is that the test is likely to be more fakable than one with knowledge-based instructions (Hooper et al., 2006). As such, the current SJT did exhibit a significant correlation with impression management scores.

Research has also demonstrated that moderate gender-based differences favoring women are likely for SJTs to the extent that they measure personality variables that women tend to score higher on than men (e.g., agreeableness, conscientiousness) (Whetzel et al., 2008). Evidence of this was seen in the current study as well.

Overall, this study offers additional support for the utility of SJTs as a standardized method of collecting and scoring relevant non-cognitive information about applicants (Oswald et al., 2004). More interesting, however, are the implications of current findings for effective SJT design.

Implications of situational strength for SJT design. In practice, there is arguably a relatively narrow range of situational strength, particularly with regard to the clarity facet of strength, within which test developers have room to maneuver when designing SJT stems. Specifically, a test developer would not want a situation to be so ambiguous as to make interpretation difficult (Brooks & Highhouse, 2006) nor so clear as to make the "correct" response obvious (thereby resulting in items with little variability in responses). While the current study indicates that situational strength can be an important design element to consider when creating an SJT with behavioral response instructions, deliberately varying situational strength in a meaningful way in light of practical constraints can be challenging[6]. The fact that fewer than half of the initial pool of 41 SJT item pairs that were designed with these considerations in mind "passed" expert judgments of differential strength and were subsequently included in the experiment, attests to the challenging nature of this undertaking.

Based on experiences creating the stimuli materials for this study, its findings, as well as the sense-making process associated with interpreting these findings in light of earlier research, some tentative practical recommendations regarding the manipulation of situational strength and the circumstances under which it is likely to matter most are made below:

1. SJT items that are academically-oriented in nature (e.g., describe situations requiring accomplishing a work goal, achieving a good grade, etc.) and thus should elicit

[6] An additional challenge is that it may be difficult to manipulate the strength of a situation without changing the nature of the construct targeted by the SJT item.

conscientiousness, may be ones that particularly require a fairly high level of clarity around rules and expectations. In light of the strong relation between clarity and consistency, low consistency in these items may be detrimental for individuals' ability to express their conscientiousness.

2. The constraints and consequences facets of strength are probably the most effective mechanisms for the manipulation of situational strength. Examples of ways to manipulate constraints are: 1. Adjusting information about the availability of other helpers (for items assessing citizenship for example); 2. Specifying the extent to which one has the requisite ability or preparation for an important task; and 3. Modifying the amount of time remaining for getting a task accomplished. Examples of ways to manipulate consequences are: 1. Delineating implications of some behavior (e.g., cheating) for one's work product, test grade, class grade, job prospects, etc.; 2. Indicating the likelihood of getting into trouble or experiencing embarrassment as a result of one's behavior; and 3. Indicating the likelihood that a relevant other will get harmed if action is not taken.
3. Enhancing the constraints or consequences associated with behavior in a particular scenario is likely to have greater implications for the way individuals will interpret and respond to the situation when the specified constraints or consequences have direct bearing on one's ability or willingness to execute the behaviors provided in the response options. For example, if a scenario has to do with volunteering for an optional extracurricular activity, specifying that the individual has a heavy workload in school that semester is probably going to affect one's willingness to commit to a certain number of hours more so than it will affect one's willingness to approach a friend about joining the extracurricular activity together.

4. The current study indicates that situational strength might impact the construct validity of an SJT with behavioral tendency instructions. Situational strength may not function the same way when it is manipulated for SJTs with knowledge-based instructions.
5. Using an expert key for scoring an SJT can tend to diffuse the effects of situational strength (as certain response options may be interpreted differently as a function of strength but not count toward an individual's test score). It is possible that a theoretical key, which rewards behaviors consistent with higher levels of the targeted trait, may better facilitate situational strength effects on test scores.
6. The situational strength of SJT scenarios is likely going to matter more for criterion-related validity when predictor constructs and criterion facets are well-matched. But since the current study did not find effects of situational strength on criterion-related validity, this recommendation should definitely be subjected to further examination.

As already indicated, a lot of the above are tentative recommendations and should therefore be examined more closely in future research. Accordingly, study limitations and specific recommendations for additional research are discussed next.

Limitations and Future Directions

The current study has a number of limitations. These limitations have to do with poor internal consistency reliability of anticipated SJT dimensions, smaller sample sizes than desired to meet power requirements (given the large number of respondents cut due to inattentiveness), the self-reported nature of most of the collected measures, the way the SJT's susceptibility to impression management was evaluated, generalizability of findings to a selection context, generalizability of findings to other SJT formats, and the focus on the effects of global situational strength rather than on the effects of the individual facets of strength. These limitations provide a

point of departure for recommending how future research may expand and improve upon the current study.

Reliability of SJT dimensions. The behavioral dimensions that the SJT was designed to assess—adaptability, citizenship, perseverance, and ethics—were not measured reliably enough to examine this study's hypotheses pertaining to SJT dimensions (e.g., the convergent and discriminant validities of SJT dimension scores with biodata dimension scores). Not surprisingly, SJT dimension scores based on Most Likely and Least Likely responses had particularly low internal consistency reliability. Likewise not surprisingly, the perseverance SJT dimension, which consisted of six items, had relatively higher reliability than the other three SJT dimensions, which had fewer items. To enable more effective examination of construct validity issues, future research in this area should attempt to maximize the reliability of hypothesized SJT dimensions by collecting ratings of all response options and increasing the number of items per SJT dimension.

Sample size and power. The weak effects associated with situational strength and the relatively low sample sizes available for testing these hypotheses meant that power for finding the hypothesized effects was often low. It should be noted that exclusion criteria were fairly stringent in this study, as participants were excluded from analyses for failing even one of the attention checks. However, analyses indicated that easing up on the exclusion criteria (allowing a participant one attention check failure) would not have substantially changed this study's results.

Self-report measures and timing. Most of the measures in this study were collected via self-report, which may have given rise to common method bias and artificially inflated relations between variables (e.g., between SJT scores and BARS). Although common method bias was unlikely to have produced spurious interactions of agreeableness and openness with situational

strength (Evans, 1985; Siemsen, Roth, & Oliveira, 2010), future research should re-examine associations found in this study using data collected from multiple sources (e.g., other reports of personality, other ratings of non-academic performance).

The predictors (i.e., SJT, biographical data, personality) and two of the outcome measures (i.e., BARS, citizenship) in this study were collected at one point in time, limiting the ability to draw conclusions about the causal nature of observed relations. It would have been better to separate these measures in time, but attrition is typically a problem when data are collected at multiple time points, with approximately 15% to 25% of students in this participant pool dropping out of studies (e.g., Billington, 2012; Zorzie, 2012). Future research should adopt a longitudinal study design to examine predictor interactions with situational strength on nonacademic outcome measures collected at a later point in time.

Assessment of impression management. Use of an impression management scale to examine the extent to which the stronger SJT may have been more susceptible to socially desirable responding than the weaker SJT may not be ideal. It has been suggested that social desirability scales like the BIDR used in the current study may be more representative of a stable individual difference than of situational pressures to manage one's impression (e.g., McFarland & Ryan, 2006). That is, these scales might not be sensitive to response distortion caused by changes in situational cues. In order to better understand the extent to which SJTs with stronger situations may be more susceptible to impression management, future research should manipulate response instructions (to respond honestly vs. to behave like an applicant) to examine the effects on SJT scores.

Generalizability to selection context. Given that context (applicants vs. incumbents) can have implications for the construct validity of SJT scores (MacKenzie et al., 2010), current

findings based on responses from college freshmen may have limited generalizability to high school students applying to colleges (the target population). One consideration is differences between applicants' and college freshmen's level of experience. Second semester freshmen who participated in the study had had one semester to gain some experience with college situations, and may have been more knowledgeable than college applicants in that regard. As such, college freshmen may be able to rely on their experience with situations similar to those on the SJT to more effectively discern expectations and respond in an effective manner. College applicants are likely to have less relevant knowledge to rely on and, in attempting to impression manage, may find it relatively more difficult to discern "correct" responses. Future research should examine what implications these differences may have for the generalizability of situational strength effects to applicants.

A second consideration is differences in current college students' and applicants' level of motivation to respond to measures in a socially desirable manner. Applicants' higher motivation to impression manage in their high stakes situation is believed to be a contributing factor to findings that, on average, applicants receive higher scores on personality measures and respond with less variability relative to incumbents (Hough, 1998; MacKenzie et al., 2010). Relatedly, there may be less variability in applicants' (relative to college freshmen's) scores on an SJT designed to measure personality and this might make it more difficult to find situational strength effects. On the other hand, given the multidimensional nature of typical SJTs, applicants may not be able to impression manage as much as they are able to on unidimensional personality scales (Hooper et al., 2006). Future research should examine the implications of situational strength for SJTs' construct and criterion-related validity with applicant samples to address the extent to which current findings generalize to less experienced, yet more motivated respondents.

Generalizability to SJTs with alternative design features. Current findings for a paper-and-pencil based SJT with behavioral tendency response instructions and an expert-derived scoring key(s) may or may not generalize to SJTs with alternative design features. Given that response instructions can affect the construct and criterion-related validity of SJTs (Ployhart & Ehrhart, 2003), current findings based on behavioral tendency response instructions may not generalize to SJTs with alternative instructions. Differences in the extent to which weak and strong SJT items assess personality may be less likely to emerge when an SJT is not a good measure of personality to begin with. In that case, current findings with regard to interactions of openness and agreeableness with situational strength may not generalize to an SJT with knowledge-based instructions as such a test would probably assess cognitive ability more so that personality (McDaniel et al., 2007).

Even though the SJT in the current study was originally developed using a construct-oriented approach (Oswald et al., 2004), it does not have unidimensional response options representing the range of a single trait. Relatedly, expert-based as opposed to theoretical scoring (which would reward behaviors representing higher levels of a particular trait, such as conscientiousness, with higher scores) was used. These factors may have made it more difficult to find situational strength effects on SJT scores. As researchers continue trying to develop SJTs using theoretical, construct-oriented approaches (e.g., Cooper et al., 2013; Mumford et al., 2008; Ployhart, 1999), future research should examine the extent to which situational strength impacts upon SJT items developed and scored using these methods.

Test delivery format is another factor worth discussing when considering the generalizability of current findings. Since video-based SJTs, relative to paper-and-pencil SJTs, can provide more contextual information and higher fidelity, can better engage test takers, and do

not rely on reading comprehension in conveying information (Jones & DeCotiis, 1986; Weekley & Jones, 1997), the likelihood of finding situational strength effects for video-based SJTs may be higher than for SJTs administered via paper-and-pencil. Whereas a test taker may accidently skip over an important piece of information pertaining to situational strength provided in a written situation, or simply fail to process it due to a high cognitive load, the same information provided via visual and auditory cues could be easier to attend and respond to. Relatedly, media richness theory (Daft & Lengel, 1984) suggests that richer media can communicate ambiguous information more effectively. Thus, future research should delve into the extent to which delivery medium matters for situational strength effects in SJTs.

Manipulation of multiple dimensions of strength. A weak and strong version of an SJT was created in this study by manipulating multiple dimensions of strength to make a particular stem seem weaker or stronger. In retrospect, a more informative study design for understanding the potentially unique effects of clarity, consistency, consequences, and constraints may have been to create weak and strong stem pairs by manipulating a single dimension of situational strength at one time. In light of the above discussion regarding the potentially unique roles of clarity as a prerequisite for finding expected effects for consequences and constraints, and of consistency as a means of manipulating clarity, it would have been beneficial to be able to: 1) Examine whether experts report perceiving clarity differences in item pairs where only consistency had been manipulated, and 2) Compare strength effects seen for SJT pairs for which a particular strength dimension was manipulated with strength effects seen for SJT pairs where another strength dimension was manipulated. In particular, future research addressing the implications of situational strength for SJTs should examine how the effects of consequences and constraints on SJTs' construct and criterion-related validity differ depending on levels of clarity.

Conclusion

The current study applied an interactionist lens to SJTs and examined the extent to which items on an SJT may function differently depending on cues that make a particular situation appear weaker or stronger. Partial support was found for the idea that strong situations should constrain the expression of individual differences in behavior (in this case, behavioral intentions). No evidence was found to indicate that situational strength affects the criterion-related validity of SJTs. It is recommended, however, that future research should consider the four facets of strength individually to better address the implications of situational strength for the psychometric properties of SJTs.

APPENDICES

APPENDIX A: Unaltered SJT Items by Dimension

Citizenship

1. After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. How would you react in this situation?

a. Encourage others to donate blood.
b. Donate money to the Red Cross instead.
c. Volunteer your time to generate money for the Red Cross.
d. Volunteer to give out cookies and help at the blood drives.
e. Ask the Red Cross if you could help them in any other way.

2. A fellow student allows you to listen to threatening phone calls that have been placed on the person's answering machine by another student. The student does not want you to tell anyone, but thinks the caller may be capable of causing physical harm. What would you do?

a. Try to talk them into calling the police and warn them not to walk around alone.
b. Talk to the resident assistant about it.
c. Contact the police yourself if you think there is any real threat of physical harm.
d. Find out who is making the calls, if it is another student, confront them – singly or jointly.
e. Unless the friend knows something that they're not saying, there is no reason NOT to call the police – so call them if your friend won't.
f. Have the friend change their phone number, and have it unlisted.

3. A friend on your floor is always organizing "social" activities including trips to local bars. Aside from the fact that this person is underage and failing some classes, you realize that the individual is drinking half a dozen or more drinks at least three or four times a week. No one else seems to know or be concerned about the person. What would you do?

a. Talk to him/her about easing up on the alcohol, explaining that it will not help with his/her classes, which should be the main reason why he/she is in college.
b. Use humor to broach the topic and offer alternatives to his/her usual "social" activities.
c. Bring up the situation with the floor's resident assistant.
d. Try to get him/her involved in other activities.
e. Talk to the person to subtly determine if there are other issues that need to be addressed, and refer him/her to help if appropriate.
f. Talk to other people on the floor, and discuss ways to address the situation.
g. Ask him/her once about this behavior and see where the discussion leads, then leave him/her to his/her own course of action.

4. You are on your way out for lunch and walking down a busy city street. There is an elderly lady a few yards ahead of you carrying a hand-bag and small bag of groceries. Suddenly, you see a group of young men converge on the elderly lady, knock her to the ground, and snatch her purse. What would you do?

a. First, try to get the purse back from the men. Then make sure the lady was not hurt. Finally, call the police.
b. Take careful notice of the men but not pursue them. Make sure the lady was not hurt, and then call the police.
c. Help the lady up and ask others around for help. Accompany her to the police station to file a report.
d. Since there are many other people present, you stay out of the incident.

5. You hear about a situation in which someone has been mistreated, and it makes you very angry. What would you do?

a. Whatever you could to keep it from happening again.
b. Help the person who was mistreated any way you can.
c. Inform the appropriate authorities.
d. Punish the person who mistreated the individual.
e. Try to identify the causes of the event and correct them.
f. Talk to the person who was mistreated to see he wants to do anything about the incident and let him know you are willing to help him.

6. After a football game, you and your friends go to a party. A friend who is driving gets severely intoxicated and insists on driving. In fact, he gets violent when someone tries to take his keys. How would you react to this situation?

a. Do what you can to grab the friend's keys so he can't drive.
b. Try to calm him down, and explain that you are concerned about him.
c. If he insists, then allow him to drive without any passengers in the car.
d. Call a cab to pick him up.
e. Tell him that you will call the police if he insists on driving.

*7. Your roommate is having problems studying for an exam in a class that happens to be your major. You have finished your assignments for the night and were planning on going out to dinner with a couple of friends. However, you recognize that your roommate has helped you previously on some of your assignments. Your roommate asks you for help in studying for exam. What would you do?

a. Explain that you already made plans, but that when you get home from dinner you can review some material with her.
b. Call your friends and cancel dinner, staying home to help your roommate and have dinner with her instead.
c. Delay your dinner plans for an hour, spending the time helping your roommate, then go out to dinner.
d. Explain that you would have liked to help her, but you already have plans that cannot be broken.

*8. At the beginning of the semester, your professor asks for a volunteer to type out the lecture

notes after each lecture for a hearing-impaired student. The task would require you to type out the lecture notes and e-mail them to the student before the following class. The position is not paid. What would you do?

a. Volunteer to type the notes.
b. Volunteer at the beginning of the semester, but half-way through the semester realize you don't want to do it anymore and quit.
c. Realize that it will take up too much time so you don't volunteer.
d. Volunteer to type half the notes if the professor can find someone to share the duty with you, but if the professor can't find anyone else you wouldn't be able to do it.

*9. Your professor announces in class that undergraduate students are needed to help run subjects for his upcoming study. While you would not receive any formal sort of extra credit, the professor would appreciate any volunteers. Given the following choices, which option would you choose?

a. Examine your schedule and offer to volunteer a couple hours a week when it is personally convenient.
b. Examine your schedule and offer to volunteer as many hours as you can.
c. Realize that you would have to give up some of your free time and choose not to volunteer.
d. Offer to run subjects only if you are paid.

*10. You are an experienced employee. A new employee comes to you for assistance. You spend time showing the employee how to do a task. Next month the same thing happens and you again help the new employee do the same task. This situation continues and you finally get upset since the new employee should be able to do the task alone. What would you do?

a. Explain to the person that you do not understand what the problem is with the task but that you have helped as much as you can.
b. As long as the person was trying, continue to show the person how to do the task.
c. Ask the employee to take notes or make a copy of the product to use as a guide in the future for how to perform the task.
d. Inform the employee to pay careful attention because this is the last time you will demonstrate how to do the task.
e. Sit down with the employee to try to determine what the problem is so that you can figure out the best way to deal with the situation from here on.

**Note*. Items 8 and 9 were adapted from Salter (2009). Item 10 (a backup item) was adapted from Trippe (2002).

Adaptability

1. You have very much wanted to be a teacher, but you failed the entrance exam into the College of Education. This exam is not given again for a year. What would you do?

a. Change majors to something similar that does not require an entrance exam.
b. Take a year off and earn some money and then retake the exam.
c. If you really want to be a teacher, take additional relevant classes, and seek advice on how to best prepare for the examination the next year.
d. Take other requirements or courses of interest to you for a year, and then retake the examination next year.

2. You are the student coordinator for the gym, and it's 4:30 P.M. You have just been informed that there is no heat in the gym. As it is the middle of winter and very cold, you know this will be a problem. There is a student dance being held in the gym at 7:00 P.M., and there are no alternative facilities in which to hold the number of people expected at this event. What would you do?

a. Let everyone know that it's postponed or called off.
b. Call maintenance, and see if they can fix it.
c. Look for small heaters to fill the room.
d. Call people and check the consensus opinion about what to do.
e. Find a group of rooms as an alternative location.
f. Inform the students to dress warmly.

3. Because of family problems, you realize that your parents can no longer support you financially at the same level as they have and you do not have enough money to continue in school. What plans would you make?

a. Apply for student financial aid or get a part-time job.
b. Ask other family members for money to finish school.
c. Drop out of school and save money for going back.
d. Take fewer classes because of the lower level of finances.

4. You share a dorm room with three other students. One half-hour before you are expecting a guest, you get home to find the place completely trashed. There is no sign of any of your roommates. What would you do?

a. Clean up the mess as much as possible before the guest arrives. Then speak with your roommates immediately upon their return, so your guest knows how concerned you were about the mess.
b. Leave the mess and explain the situation to your guest.
c. Leave the mess and take the guest somewhere else.
d. Clean up the mess as much as possible before the guest arrives. Then, without the guest around, ask the roommates why the place was trashed so badly and what can be done in the future to avoid this situation.

5. You have a big exam the next day and just settle down to study when you realize friends in an adjacent room have organized a party for that night. You would love to join them and the noise makes studying difficult. What would you do?

a. Find an alternative place to study, like the library.
b. Tell yourself that you have to study for an hour, then you can go to the party.
c. Ask them to keep quiet, or wait until you are done studying if they really want to party.
d. Put your headphones on and try to study anyway.

6. You just spent two days working on a report for one of your courses. You gave it to one of the teaching assistants to review. She returns it to you with extensive suggestions for revision, primarily involving editorial changes that reflect her preferred style of writing rather than any real changes to the substance of the paper. What would you do?

a. Make changes according to her writing style because it could help your grade.
b. Make the revisions as long as the main ideas and content of the paper are the same.
c. Ignore the style changes but correct any suggestions that you thought had merit.
d. Go to someone else for a second opinion, like the professor.
e. Talk to the teaching assistant about the problem and explain why your style of writing is fine.
f. Leave the paper as is.

7. In your first semester, you realized that none of your professors took attendance and that you would not have to go to class if you didn't want to. You began skipping classes that met early in the morning or at times when there were interesting social opportunities. The first test grade you received was a 1.0. What actions would you take?

a. Attend all of your classes from this time on.
b. Start going to all classes; contact the TA to help clarify problems with the exam, and get help with the new material.
c. Go to class and study very hard.
d. Make sure you go to all classes and ask your professor to reward class attendance.
e. Start going to class or make sure you have someone else's notes and study the text when you sleep in.

8. After returning to school after an extended absence, you learn that you have a paper due the next day in one of your classes. Prior to leaving, you checked your syllabus and asked the professor what you would be missing by being out so that you could take care of it before you left. There was no mention of the paper. What would you do?

a. Write the best paper you can in the time allotted. Then, when turning it in, speak with the professor about your absence and the possible miscommunication.
b. Contact the professor as soon as you are aware of the paper, explain the situation, and ask for their suggestion.
c. Go to class the next day and argue for a week's extension.

d. Ask the professor privately for an extension.
e. Do the paper and submit it when complete, regardless of the deadline imposed while you were gone.
f. Inform the professor that the paper is not on the syllabus, and you were not made aware, then agree on a reasonable due date.

9. Half way into the first semester you realize that you have too much to do. You are taking six courses, working 20 hours a week, and your boyfriend/girlfriend is demanding that you spend more time together. You feel that you must go home to visit a relative on many weekends. You also find you are missing classes and meetings. What action(s) would you take?

a. Drop a course or take it without being graded.
b. Cut down on the hours you work.
c. Cut back a little bit on everything: classes, work, the relationship, and weekend visits.
d. Ask your boy/girlfriend to be understanding and support your need for more time for school.
e. Prioritize your responsibilities and goals. Then divide up time based on importance.
f. Try to save time by calling/writing the sick relative instead of visiting, unless it was really serious.

*10. You are going through an especially busy period at school. It is the end of the semester, you have papers due, need to prepare for exams, and coworkers at your part-time job are asking that you work more shifts to help them out. You find yourself beginning to lose track of details and are feeling overwhelmed. What do you do?

a. Decide what's important and then prioritize your responsibilities.
b. Relax and take a step back, knowing that you can't do everything at once.
c. Get organized, and start planning ahead and scheduling.
d. Apologize, decline the extra shifts, and tell your coworkers that school is your first priority.
e. Quit your job.
f. Sleep less and work harder to get things done.

*11. You are interested in finance, but do not have further finance courses for at least another semester. What would you do?

a. Wait until the next semester, and take another class then.
b. Try to register for an alternative finance course as an elective.
c. Use the semester to do some independent study so that you are well prepared for the next course.
d. Get involved in on-campus finance clubs or investment games.
e. See if you could be a TA for a finance class.

**Note*. These items are backup items.

Perseverance

1. Your professor has just given you a project that will obviously require the whole semester to complete. She gave you all the details you need to get started, but you are not sure how the project should proceed from there. She does not appear to intend to give you any more information in class. How would you proceed?

a. Work out the project to the best of your ability and approach the professor if you get stuck.
b. Generate some ideas, and then go to office hours to see how the professor responds to them.
c. Ask the professor about the project after class.
d. Visit the professor or a teaching assistant during office hours to discuss the project.
e. Talk to other students to get an idea of what they are doing.
f. Try to get an idea of whether or not other students seem confused. If so, bring the issue up with the professor during class.

2. You are collaborating with other classmates on a project. The group of you keeps running into a variety of problems that threaten to cause the project to be late. The other group members want to just plan to submit it late. Another option would be to devote much more time than planned to the project and possibly get it in on time. What would you do?

a. Try to get it done, but plan to submit it late.
b. Ask the instructor for help or for an extension. If that doesn't work, just try your best and do what you can or turn it in late.
c. Motivate the group to devote more time and work together to get it done.
d. Have the group decide what to do.
e. Work hard to finish it because there are consequences for being late and meeting deadlines is important to you.
f. Tell the instructor your situation, and ask them for advice.

3. You are finding a particular class dull and boring, and are having difficulty staying awake. What would you do?

a. Do what you can to stay awake, such as drinking caffeine or sitting toward the front of the class.
b. Read the class material beforehand to make the lecture more interesting.
c. During the lecture, do some studying that is required for the course.
d. Make sure you are getting enough sleep every school night.
e. Skip the class if it is that dull and boring to you.

4. You have been working on a research paper for a couple of weeks. You initially had trouble settling on a satisfactory topic for this paper and then you had trouble finding sufficient reference material to do the paper. Now you are having trouble organizing the material you do have. The paper is due in two weeks and you would have time to start over on a new topic which sounds easier. What approach would you take to complete this course requirement?

a. Pick an easier topic and get on with it.
b. Ask the professor what they think you should do.
c. Stick with the topic you have, but ask the professor for some help with organizing the material.
d. Stick with the topic, and take your work over to the writing center to see if you can get some help.
e. Figure that a new topic that sounds easy may not be – you could have the same problem all over again, so just battle on with what you have already started.

5. You have a professor you find "monstrously boring." He seems rude and longwinded and often reads material from the textbook. He requires attendance. His course is required for your major. What would you do?

a. Find a seat in the back and focus on some other work during the class.
b. Continue to attend, working hard to pay attention. You know you can hang in there until the end of the semester.
c. Focus on taking as many notes as you can to make sure you completely understand the material.
d. Approach the professor and ask whether he could incorporate more interactive exercises into the lesson.
e. Bring a crossword puzzle to class to keep yourself occupied during the boring parts.

6. You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. What would you do?

a. Talk with the professors and TAs to get help on how to study.
b. Find a study partner and work on homework and class material together.
c. Talk to your parents and an advisor.
d. Study hard, try your best, and don't worry about it.
e. Talk to my advisor and teachers; see if there are study groups or review sessions I can attend.
f. Hire a tutor for the difficult classes.

7. You decided early in the term to do a paper on a topic very interesting to you. However, you have found it difficult to find information on your topic, your job has taken more time than you wanted, and you have had more work in your other courses than you anticipated. Now it seems like you may have to engage in several "all-nighters" to complete your paper on time. What would you do?

a. Seek help from other students who may have had a similar experience.
b. Pick a topic that can be completed quicker. An "A" is an "A".
c. Set up a schedule on which you can complete all of the other work you need to do, spend as much time on the paper as possible, and meet with the instructor to discuss what you have so far and get suggestions.
d. Do whatever it takes to complete the paper, including "all-nighters".
e. Talk to the instructor about the situation and ask for advice.

f. Make the paper a priority, but take into account how much the paper is worth in the class.

8. During a summer sports camp, you find the training exceptionally difficult and tedious (e.g., long runs, difficult and length exercises) and that you are not yet learning anything new about the sport in which you are interested. How would you react?

a. Continue to train, because that is what is needed to perform.
b. See whether others at the camp feel the same way, and tell the coaches.
c. Consider not going to the camp again because it was a disappointing experience.
d. Be upset and discouraged about the camp experience, though others may be learning something.
e. Suggest new activities to the coaches to make the training more interesting.

9. You are required to take an exam that is required of all calculus students, but not specifically a part of your calculus course. You take and fail this exam three times. You need the calculus credit to continue in your major. What would you do?

a. Keep taking the exam in the hopes you will eventually pass.
b. Begin to think of some other options.
c. Talk with the director of the program and explain that the exam covers material that is not part of any classes, find out where you can get help with this material.
d. Try taking the class from a different instructor and join a study group.
e. Find someone to tutor you with calculus specifically as it relates to the exam material.
f. Change your major.

10. You are having problems in one of your classes. The second exam is quickly approaching, and you want to do well on it, especially considering the poor grade you received on the first exam. You have worked hard so far in college to maintain a good GPA and don't want to sacrifice it. What would you do?

a. Continue to study and not give up even when you encounter difficult concepts.
b. Make an appointment and consult the professor for advice on studying and how to approach the material.
c. Set aside more time to study than you did studying for the previous exam.
d. Accept that you cannot always get great grades and focus your time and energy on other classes that you are doing well in.

**Note*. This is a backup item. It was adapted from Salter (2009).

Ethics

1. You and your friends know that a mutual acquaintance has been dating another person for nearly a year. However, one of your friends tries her best to get a date with this individual. How would you react?

a. If the acquaintance is in a happy relationship, tell your friend to wait and rethink it. If not, it is ok to get a date.
b. Support my friend. Good for her.
c. Tell your friend to wait until the person is single or to just forget about the person.
d. Tell your friend that it is inappropriate to interfere in the relationship.
e. Just be annoyed at your friend. Do not get involved.

2. You know that a group of students in your class cheats on exams by putting formulas into scientific calculators or into cell phones. The professor has clearly warned against such activity, but you are not sure what she would do if she knew what these students were doing. What action would you take?

a. Try doing the same thing until people start getting caught.
b. Study the way you know best, don't cheat, but don't turn in the other students either.
c. You would do nothing; it's none of your business.
d. You would mention it to the professor so she can deal with the problems in the class.
e. Don't tell the professor, but make sure it is clear you are not involved in case they get caught.
f. Send the professor an anonymous message about what is going on.

3. In a class of 50 students, you discover that a group of your friends have worked out a scheme to share answers on an exam. The professor has vision problems and will likely never notice. You are not doing very well in the course. What would you do in these circumstances?

a. Avoid being around these friends.
b. It is not exactly honest but under the circumstances, the scheme is OK. You would join them.
c. Do your own work and not tell the professor about the scheme because it is not your problem.
d. Cheat and get a good grade.
e. Tell the professor about the scheme.
f. Study for the exam, but join the scheme as a backup strategy for the test.

4. One of your friends takes the same class you do, but at an earlier time. On the day of the midterm, he comes to your room and offers you a copy of the exam he just took. You know from earlier exams that the professor gives the same exam to all sections.

a. Take the exam to study from, but not to memorize.
b. Tell your friend what you have studied and ask them whether that will be sufficient.

c. Look the exam over so that you can study the areas that you don't know well, using it as a study guide.
d. Thank your friend, and politely decline.
e. Decline the exam, and suggest to the professor that in the future she has different forms of the exam.

5. You find a very expensive cell phone on the edge of a sidewalk along the street. There is no identifying information on the phone. What would you do with the phone?

a. Leave it there. The person who lost it might come back looking for it.
b. Since there is no identification, take it, and have the number changed.
c. Take it. Either wait until someone calls, or call a number on the phone. Figure out who the owner is and return it.
d. Put a 'lost and found' ad in the paper, or put up signs about the lost phone.
e. Make some effort to return it, but keep it if you do not find the owner.
f. Return it to the authorities or 'lost and found.' Or, return it to a phone store/dealer.

6. You have been having trouble in a course. On the day of the exam someone offers to sell you a copy of the exam. You have heard rumors that a large number of students in the class have purchased this exam and are afraid that if many students do well, your grade will look even worse. What would you do?

a. Purchase the exam so that you do not look worse than everybody else.
b. Take the exam with the knowledge that you have from studying.
c. Approach the instructor and explain what you have been offered.
d. Buy the exam, but try to miss some of the questions.
e. Inform the instructor anonymously in the hopes that he will change the exam.

7. One of your friends wants to copy the homework of another friend. They have the same class but different instructors so there is no chance of detection. The friend who has completed the work refuses to let your other friend copy the work and the second friend becomes very angry and complains to you. How would you react?

a. Tell her to quit complaining and just do the assignment.
b. Explain that she's better off in the long run by doing the work.
c. Suggest to your friend that she ask for help with the homework instead.
d. Tell her that the friend who refused to cheat was under no obligation. She should respect that choice.
e. Not interfere in the conflict. If anything, you would support the friend refusing to allow copying.
f. Listen to each person and try to understand his/her point of view.

8. You are taking a particularly difficult exam when you realize how easy it is to see the paper of a student who you know is doing well in the course. When you look at this student's paper, she notices you and pulls her paper away. You are deeply embarrassed. What action would you take?

a. Don't show that you are embarrassed, and see if you can look at another paper.
b. Just finish your exam as usual and then leave the room.
c. Pretend like it didn't happen, make an effort not to look there again, and finish the test.
d. Finish the exam, and apologize to the student after class.

9. You look at a classmate's paper during an exam and the person turns to you telling you to quit. When the professor notices this exchange, she asks you both to leave and gives you both a failing grade. What would you do now?

a. Make sure that you track down the other student and apologize to them.
b. Go to the professor and explain that your classmate had no part in your cheating, and was simply telling you to quit.
c. Go to the professor and apologize for trying to cheat.
d. Explain to the professor what had been going on, and beg that they not give your classmate a failing grade.

*10. You are a member of a team that has completed a class project. The professor hands back the grade and feedback on the project. In professor's comments was a pointed attack on the group for plagiarizing and half a letter grade was deducted from the project's final grade. You know that the student honor code requires you to report whomever has plagiarized. However, the plagiarizing was not related to the portion of the project that you were personally responsible for. What would you do?

a. Discuss the situation with the group and come up with a decision what to do, or not to do, together.
b. Tell the person responsible for the error that he should contact the team leader.
c. Tell the professor who was responsible for the plagiarism.
d. Accept the grade, learning from the experience and vowing never to let it happen again.

**Note*. This is a backup item. It was adapted from Salter (2009).

APPENDIX B: SJT Items at Two Levels of Situational Strength

Table 15. **Citizenship 1**

Weaker	Stronger
After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. **The Red Cross will probably get enough donors to help all the people who need blood[1]**. How would you react in this situation?	After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. **The Red Cross clearly has a shortage of blood donors[1]**. **Several of your close relatives need blood[1].** How would you react in this situation?
Note. 1 represents low *consequences*.	*Note*. 1 represents high *consequences*.
Words = 45	Words = 45
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 69.4	Flesch Reading Ease = 69.8
Flesch-Kincaid Grade Level = 6.3	Flesch-Kincaid Grade Level = 5.7

a. Encourage others to donate blood.
b. Donate money to the Red Cross instead.
c. Volunteer your time to generate money for the Red Cross.
d. Volunteer to give out cookies and help at the blood drives.
e. Ask the Red Cross if you could help them in any other way.
f. Convince your best friend to donate blood since you cannot. (added)

Table 16. **Citizenship 2**

Weaker	**Stronger**
A fellow student allows you to listen to threatening phone calls that have been left on the individual's answering machine. The individual **does not want you to tell anyone else because he/she is embarrassed about the situation**[1], and **hopes that the caller is not capable of causing physical harm**[2]. What would you do?	A fellow student allows you to listen to threatening phone calls that have been left on the individual's answering machine. The individual **makes you swear not to tell anyone because his/her own conduct was highly inappropriate**[1], but **knows the caller has a criminal record and could cause physical harm**[2]. What would you do?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 53	Words = 53
Passive Sentences = 33%	Passive Sentences = 33%
Flesch Reading Ease = 60	Flesch Reading Ease = 56.9
Flesch-Kincaid Grade Level = 9.3	Flesch-Kincaid Grade Level = 9.7

a. Try to talk the person into calling the police and warn the individual not to walk around alone.
b. Contact the police yourself if you think there is any real threat of physical harm.
c. Find out who is making the calls, if it is another student, confront the person – singly or jointly.
d. Unless the person knows something he/she is not saying, there is no reason NOT to call the police – so call the police if the person won't.
e. Talk to the resident assistant about the situation.
f. Have the person change his/her phone number, and have it unlisted.

Table 17. **Citizenship 3**

Weaker	**Stronger**
A student on your floor is always organizing "social" activities, including trips to the local bars. **The individual is drinking three or more drinks at least three times a week**. **He/she is still attending classes and completing assignments but the situation could get worse[1].** You **do not know whether anyone else is aware or concerned[2]** about the person. **This individual is not a close friend[2]**.What would you do?	A student on your floor is always organizing "social" activities, including trips to local bars. **Aside from the fact that he/she is underage and failing several classes, you realize that the individual is drinking half a dozen or more drinks at least four times a week[1]**. **No one else seems aware or concerned[2]** about the person. **This individual is a close friend of yours[2]**. What would you do?
Note. 1 represents low *consequences*; 2 represents low *clarity*.	*Note*. 1 represents high *consequences*; 2 represents high *clarity*.
Words = 68	Words = 68
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 62.8	Flesch Reading Ease = 67.7
Flesch-Kincaid Grade Level = 7.9	Flesch-Kincaid Grade Level = 7.2

a. Talk to the person about easing up on the alcohol, explaining that it will not help with his/her classes, which should be the main reason why he/she is in college.
b. Use humor to broach the topic and offer alternatives to the individual's usual "social" activities.
c. Bring up the situation with the floor's resident assistant.
d. Try to get the individual involved in other activities.
e. Talk to the person to subtly determine if there are other issues that need to be addressed, and refer him/her to help if appropriate.
f. Talk to other people on the floor, and discuss ways to address the situation.
g. Ask the individual once about this behavior and see where the discussion leads, then leave the person to his/her own course of action.

Table 18. **Citizenship 4**

Weaker	**Stronger**
You are on your way out for lunch and walking down the street. There is a **middle-aged**[1] lady a **block**[1] ahead of you carrying a hand-bag and bag of groceries. Suddenly, you see a group of hooded young men converge on the lady, knock her to the ground, and snatch her purse. **You hear her yell for help, directed at no one in particular**[1]. There **are several other people close by**[2]. What would you do?	You are on your way out for lunch and walking down the street. There is an **elderly**[1] lady a **few yards**[1] ahead of you carrying a hand-bag and small bag of groceries. Suddenly, you see a group of hooded young men converge on the elderly lady, knock her to the ground, and snatch her purse. **Seeing you, she yells for help getting back her purse**[1]. **There is one other person nearby**[2]. What would you do?
Note. 1 represents low *clarity*; 2 represents low *constraints*.	*Note*. 1 represents high *clarity**; 2 represents high *constraints*.
Words = 75	Words = 75
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 82.4	Flesch Reading Ease = 84.7
Flesch-Kincaid Grade Level = 4.8	Flesch-Kincaid Grade Level = 4.5

*****It is clearer that an elderly lady needs help getting up.

a. First, try to get the purse back from the men. Then make sure the lady was not hurt. Finally, call the police.
b. Take careful notice of the men but do not pursue them. Make sure the lady was not hurt. Finally, call the police.
c. Help the lady up and ask others around for help. **Offer to** accompany her to the police station to file a report.
d. **Since you are not the only one available to help**, stay out of the incident.
e. Quickly take out a notepad and pen and write down as much information as possible about the men. Then call the police. (added)

Table 19. **Citizenship 5**

Weaker	**Stronger**
You hear about a situation in which **your acquaintance was mistreated, but it is not likely to happen to others[1]. Several individuals know about the situation and may or may not have done something about it[2]. You are not aware of any University policies regarding such situations[2].** What would you do?	You hear about a situation in which **your friend was physically hurt and it can easily happen to other people[1]. Nobody else knows about the situation so nobody will have done anything about it[2]. You know the University encourages individuals to file prompt complaints in such situations[2]**. What would you do?
Note. 1 represents low *consequences*; 2 represents low *clarity*.	*Note*. 1 represents high *consequences*; 2 represents high *clarity*.
Words = 51	Words = 51
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 54.5	Flesch Reading Ease = 52.8
Flesch-Kincaid Grade Level = 8.8	Flesch-Kincaid Grade Level = 9.0

a. Do whatever you can to keep it from happening again.
b. Help the person who was mistreated any way you can.
c. Inform the appropriate authorities.
d. Punish the person who mistreated the individual.
e. Try to identify the causes of the event and correct them.
f. Talk to the person who was mistreated to see if he/she wants to do anything about the incident and let him/her know you are willing to help.

Table 20. **Citizenship 6**

Weaker	**Stronger**
After a football game, you and your friends go to a party. After the party, a friend who **has had a few drinks**[1] insists he can still drive. He gets **angry**[1] when someone tries to take his keys. **He insists that he never drives unless he is sure he can handle it**[1]. How would you react to this situation?	After a football game, you and your friends go to a party. After the party, a friend who **has gotten severely intoxicated**[1] insists he will be driving home. He gets **violent**[1] when someone tries to take away his keys. **He has driven drunk before and almost got into an accident that time**[1]. How would you react to this situation?
Note. 1 represents low *consequences*.	*Note*. 1 represents high *consequences*.
Words = 59	Words = 59
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 88.7	Flesch Reading Ease = 70.1
Flesch-Kincaid Grade Level = 3.8	Flesch-Kincaid Grade Level = 6.4

a. Do what you can to grab his keys so that he can't drive.
b. Try to calm him down, and explain that you are concerned about him.
c. If he insists, then allow him to drive without any passengers in the car.
d. Tell him that you will call the police if he insists on driving.
e. Call a cab to come pick him up **and drive him home**.
f. Force him into the backseat of your car, accompanied by another friend, and drive him home yourself. (added)

Table 21. **Citizenship 7**

Weaker	Stronger
Your roommate is having problems studying for an exam in a class that you took last semester. **You had done very well in the class[1].** You have finished your assignments for the night and were planning on going out to dinner with a couple of friends. Your roommate asks you for help in studying for the exam. **She tells you she is having trouble understanding several concepts that could be on the test[2]. She has another friend she could ask to help her study[1]. You don't know how much time helping her would take[3].** What would you do?	Your roommate is having problems studying for an exam in a class that you took last semester **and barely managed a passing grade in[1]**. You have finished your assignments for the night and were planning on going to dinner with a couple of friends. Your roommate asks you for help in studying for the exam. **She says that she will get a failing grade in the class if she does not pass this exam[2]. There is no one else she could ask for help[1]. She asks for exactly two hours of your time[3].** What would you do?
Note. 1 represents low *constraints;* 2 represents low *consequences*; 3 represents low *clarity*.	*Note*. 1 represents high *constraints;* 2 represents high *consequences*; 3 represents high *clarity*.
Words = 98	Words = 98
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 83.9	Flesch Reading Ease = 86.4
Flesch-Kincaid Grade Level = 4.5	Flesch-Kincaid Grade Level = 4.6

a. Explain that you already made plans, but that when you get home from dinner you can review some material with her.
b. Call your friends and cancel dinner. Stay home to help your roommate and have dinner with her instead.
c. Delay your dinner plans for an hour. Spend the time helping your roommate and then go out to dinner.
d. Explain that you would have liked to help her, but you already have plans that cannot be broken.
e. Invite the roommate to go to dinner with you. You can help her on your drive to and from dinner. (added)
f. Tell her that you would like to help, but convince her that you would be of no real help to her. (added)

Table 22. **Citizenship 8**

Weaker	Stronger
At the beginning of the semester, your professor asks for a volunteer to type out the lecture notes after each lecture for a hearing-impaired student. The task would require you to **type out the lecture notes sometime before the following class (several days later)**[1] and e-mail them to the student so he has time to review them. The position is not paid. You are **not sure how long and detailed lecture notes will tend to be**[2]. **You also consider that your workload might increase later in the semester**[1]. What would you do?	At the beginning of the semester, your professor asks for a volunteer to type out the lecture notes after each lecture for a hearing-impaired student. The task would require you to **type out the lecture notes right after class**[1] and e-mail them to the student so he has time to review them before the following class. The position is not paid. You know **this professor presents lots of information during lectures**[2] and **you can't always keep up**[1]. **You have a heavy workload this semester and work part time**[1]. What would you do?
Note. 1 represents low *constraints*; 2 represents low *clarity.*	*Note*. 1 represents high *constraints*; 2 represents high *clarity.*
Words = 92	Words = 92
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 70.1	Flesch Reading Ease = 68.9
Flesch-Kincaid Grade Level = 7.3	Flesch-Kincaid Grade Level = 7.4

a. Volunteer to type the notes. **Do so the entire semester.**
b. Volunteer **to type the notes**. **If you realize** half-way through the semester that you **can't** do it anymore, **ask the professor to find someone else**.
c. Don't volunteer. You realize that it will take up too much time **and someone else can probably do it better.**
d. Volunteer to type half the notes if the professor can find someone to share the duty with you. If the professor can't find anyone else you wouldn't be able to do it.
e. Volunteer to type the notes. Ask a friend in class to share the responsibility with you so that you can take turns. (added)

Table 23. **Citizenship 9**

Weaker	Stronger
Your professor announces in class that undergraduate students are needed starting next week to help run subjects for his research study. While you would not receive any extra credit, the professor would appreciate any volunteers. **This semester you are taking several classes but your workload has been manageable[1]. Your midterms are still three weeks away[1]**. What would you do?	Your professor announces in class that undergraduate students are needed started next week to help run subjects for his research study. While you would not receive any extra credit, the professor would appreciate any volunteers. **This semester you are taking several difficult classes that consume most of your time[1]. You have two midterms next week[1]**. What would you do?
Note. 1 represents low *constraints.*	*Note*. 1 represents high *constraints.*
Words = 59	Words = 59
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 60.0	Flesch Reading Ease = 61.5
Flesch-Kincaid Grade Level = 7.8	Flesch-Kincaid Grade Level = 7.6

a. Examine your schedule and offer to volunteer a couple of hours a week when it is personally convenient.
b. Examine your schedule and offer to volunteer as many hours as you can.
c. Wait to see how many other people volunteer; agree to help a couple hours a week only if no one else volunteers. (added)
d. Volunteer only if a friend of yours volunteers and asks you to do it as well; volunteer for as many hours as your friend. (added)
e. Realize that you would have to give up some of your free time and choose not to volunteer.
f. Offer to run subjects only if you are paid.

Table 24. **Citizenship 10**

Weaker	Stronger
You have been a research assistant in a lab for a year now and are considered "experienced". **A new RA realizes this and approaches you for assistance**[1]. You spend time showing the individual how to do a task. Next month the same thing happens and you again help the individual do the same task. **This situation continues and you start to wonder if the individual is simply lazy and prefers to have help**[2]**. You work in the lab 20 hours per month and it only takes fifteen minutes or so for you to help the individual each month**[2]. What would you do?	You have been a research assistant in a lab for a year now and are considered "experienced". **Your professor directs a new RA to you for assistance and asks you to help her out**[1]. You spend time showing the individual how to do a task. Next month the same thing happens and you again help the individual do the same task. **This situation continues and you start to think the individual has a learning disability**[2]**. You work in the lab only 8 hours per month and helping this individual is taking up an hour of your limited time**[2]. What would you do?
Note. 1 represents low *clarity*; 2 represents low *constraints.*	*Note*. 1 represents high *clarity*; 2 represents high *constraints.*
Words = 102	Words = 102
Passive Sentences = 14%	Passive Sentences = 14%
Flesch Reading Ease = 73.4	Flesch Reading Ease = 72.6
Flesch-Kincaid Grade Level = 6.6	Flesch-Kincaid Grade Level = 6.7

a. Explain to the individual that you do not understand what the problem is with the task but that you have helped as much as you can.
b. As long as the individual was trying, continue to show him/her how to do the task.
c. Ask the individual to take notes or make a copy of the product to use as a guide in the future for how to perform the task.
d. Inform the individual to pay careful attention because this is the last time you will demonstrate how to do the task.
e. Sit down with the individual to try to determine what the problem is so that you can figure out the best way to deal with the situation from here on.
f. Inform the professor you are working for that you are frustrated with the situation and he/she should intervene. (added)

Table 25. **Adaptability 1**

Weaker	Stronger
Though you have multiple talents and career possibilities[1], you have very much wanted to be a teacher. Both **of your parents are teachers but have always allowed you to choose your own path. They will probably help you pay for college regardless of your major**[2]. You just found out you failed the entrance exam into the College of Education. This exam is not given again for a year. What would you do?	You have very much wanted to be a teacher and **do not have any other interests or talents**[1]. **Your parents are teachers and have been pushing you toward that career as well. They promise to help you pay for college if you become a teacher**[2]. You just found out you failed the entrance exam into the College of Education. This exam is not given again for a year. What would you do?
Note. 1 represents low *constraints*; 2 represents low *clarity*.	*Note*. 1 represents high *constraints*; 2 represents high *clarity*.
Words = 72	Words = 72
Passive Sentences = 16%	Passive Sentences = 16%
Flesch Reading Ease = 73.6	Flesch Reading Ease = 78.3
Flesch-Kincaid Grade Level = 5.9	Flesch-Kincaid Grade Level = 5.3

a. Change majors to something similar that does not require an entrance exam.
b. Take a year off and earn some money and then **decide if you want to** retake the exam next year.
c. If you really want to be a teacher, take additional relevant classes, and seek advice on how to best prepare for next year's exam.
d. Take other requirements or courses of interest to you for a year, and then retake the exam next year.
e. Try to convince an administrator that you were having a bad day when you took the exam. Maybe you'll be allowed to retake it this year. (added)

Table 26. **Adaptability 2**

Weaker	**Stronger**
You are the student coordinator for the gym, and it's **2:00 P.M.**[1] You have just been informed that there is no heat in the gym. As it is **almost winter and pretty chilly**[2], you know this will be a problem. There is a student dance being held in the gym at 7:00 P.M., and there are no alternative facilities in which to hold the number of people expected at this event. **You have become the student coordinator only recently. You are not even sure if you are expected to handle these sorts of issues in your position.**[3] What would you do?	You are the student coordinator for the gym, and it's **6:00 P.M.**[1] You have just been informed that there is no heat in the gym. As it is **the middle of winter and very cold**[2], you know this will be problematic. There is a student dance being held in the gym at 7:00 P.M., and there are no alternative facilities in which to hold the number of people expected at this event. **You know that you are expected to handle these sorts of issues as student coordinator.**[3] **You do not have the authority to cancel events, however**[1]**.** What would you do?
Note. 1 represents low *constraints*; 2 represents low *consequences*; 3 represents low *clarity*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*; 3 represents high *clarity*.
Words = 101	Words = 101
Passive Sentences = 40%	Passive Sentences = 20%
Flesch Reading Ease = 63.2	Flesch Reading Ease = 62.3
Flesch-Kincaid Grade Level = 9.4	Flesch-Kincaid Grade Level = 9.5

a. Let everyone know that the dance is postponed or called off.
b. Call maintenance, and see if they can fix it.
c. Look for small heaters to fill the room.
d. Call **other student leaders and faculty** and check the consensus opinion about what to do.
e. Find a group of rooms as an alternative location.
f. Inform the students to dress warmly.

Table 27. **Adaptability 3**

Weaker	**Stronger**
Because of family problems, you find out that your parents can no longer support you financially at the same level as they have. You do not have enough money to continue in school for **two more years**[1]. **You don't know if your parents really want you to finish college— it never seemed to bother them that your brother did not finish college**[2]. **You might find a job that does not require a college degree**[3]. What plans would you make?	Because of family problems, you find out that your parents can no longer support you financially at the same level as they have. You do not have enough money to continue in school for **three more years**[1]. **Your parents regret not being able to put you through college but expect that you will take out loans and work to support yourself**[2]. **You won't be able to realize your career plans without a college degree**[3]. What plans would you make?
Note. 1 represents low *constraints*; 2 represents low *clarity*; 3 represents low *consequences*.	*Note*. 1 represents low *constraints;* 2 represents high *clarity*; 3 represents high *consequences*.
Words = 79	Words = 79
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 76.2	Flesch Reading Ease = 76.2
Flesch-Kincaid Grade Level = 6.5	Flesch-Kincaid Grade Level = 6.5

a. Apply for student **loans** or get a part-time job.
b. Ask other, **more wealthy** family members for money to finish school.
c. Drop out of school and save money for going back.
d. Take fewer classes because of the lower level of finances.
e. Rely on your parents to figure something out—you should be able to count on them. (added)
f. Transfer to a less expensive, community college. (added)

Table 28. **Adaptability 4**

Weaker	Stronger
You share a dorm room with three other students. **One hour**[1] before you are expecting a guest, you get home to find the place completely trashed. There is no sign of any of your roommates. **You don't know how your guest would feel about the mess. You do not know the individual very well**[2]**. Your roommates sometimes react negatively to criticism**[4]. What would you do?	You share a dorm room with three other students. **Fifteen minutes**[1] before you are expecting a guest, you get home to find the place completely trashed. There is no sign of any of your roommates. **You know that the guest would judge you for the mess**[2]**. The individual's opinion is very important to you**[3]**. Your roommates always react negatively to criticism**[4]. What would you do?
Note. 1 represents low *constraints*; 2 represents low *clarity*; 4 represents low *consistency*; Note: *consequences* not represented.	*Note*. 1 represents high *constraints*; 2 represents high *clarity*; 3 represents high *consequences*; 4 represents high *consistency*
Words = 65	Words = 65
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 84.1	Flesch Reading Ease = 76.3
Flesch-Kincaid Grade Level = 3.8	Flesch-Kincaid Grade Level = 4.9

a. Clean up the mess as much as possible before the guest arrives. Then speak with your roommates immediately upon their return, so your guest knows how concerned you were about the mess.
b. Leave the mess and explain the situation to your guest.
c. Leave the mess and take the guest somewhere else.
d. Clean up the mess as much as possible before the guest arrives. Then, without the guest around, ask the roommates why the place was trashed so badly and what can be done in the future to avoid this situation.
e. Call the guest and ask to reschedule your meeting. When your roommates arrive, ask them to clean up the mess. (added)

Table 29. **Adaptability 5**

Weaker	**Stronger**
You have **an exam in one of your classes the next day[1]. But this exam will not be weighted too heavily in your grade and you received high marks on the previous two exams[1]. Two hours after you settle down to study[2]**, you realize that **students[3]** in an adjacent room have organized a party. **You have not actually received an invitation from them but could still join the party[3]**. You know that the noise will make studying difficult. What would you do?	You have **a final exam in one of your major classes the next day[1]. This exam will be weighted heavily in your grade and you want to do well because you intend to ask the professor for a letter of recommendation[1]. Just as you settle down to study[2]**, you find out **friends[3]** in an adjacent room have organized a party. **They have invited you so you could join them[3].** You know that the noise will make studying difficult. What would you do?
Note. 1 represents low *consequences*; 2 represents low *constraints*; 3 represents low *clarity*.	*Note*. 1 represents high *consequences*; 2 represents high *constraints*; 3 represents high *clarity*.
Words = 82	Words = 82
Passive Sentences = 16%	Passive Sentences = 16%
Flesch Reading Ease = 77.4	Flesch Reading Ease = 78.4
Flesch-Kincaid Grade Level = 5.8	Flesch-Kincaid Grade Level = 5.7

a. Find an alternative place to study, like the library.
b. Tell yourself that you have to spend the next hour studying, and then you can go to the party.
c. Ask them to keep quiet, or wait until you are done studying if they really want to party.
d. Put your headphones on and try to study anyway.
e. Put in earplugs to block out some of the noise and try to study anyway. (added)

Table 30. **Adaptability 6**

Weaker	**Stronger**
You just **spent two days**[1] preparing a report for one of your courses. **You thought it would be useful to give your report to one of the teaching assistants to review**[2]. She returns it to you with extensive suggestions for revision, primarily involving editorial changes that reflect her preferred style of writing rather than any real changes to the substance of the paper. The paper is **due in 4 days**[1]. **You need to get at least a "C"**[1]. What would you do?	You just **spent a full week**[1] working on a report for one of your courses. **You followed the professor's strong recommendation and gave your report to one of the teaching assistants to review**[2]. She returns it to you with extensive suggestions for revision, primarily involving editorial changes that reflect her preferred style of writing rather than any real changes to the substance of the paper. The paper is **due in 2 days**[1]. **You need to get an "A"**[1]**.** What would you do?
Note. 1 represents low *constraints*; 2 represents low *clarity*.	*Note*. 1 represents high *constraints*; 2 represents high *clarity*.
Words = 82	Words = 82
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 70.4	Flesch Reading Ease = 65.3
Flesch-Kincaid Grade Level = 6.8	Flesch-Kincaid Grade Level = 7.5

a. Make changes according to her writing style because it could help your grade.
b. Make the revisions as long as the main ideas and content of the paper are the same.
c. Ignore the style changes but correct any suggestions that you thought had merit.
d. Go to someone else for a second opinion, like **the other teaching assistant or** the professor.
e. Talk to the teaching assistant about the problem and explain why your style of writing is fine.
f. Leave the paper as is.

Table 31. **Adaptability 7**

Weaker	Stronger
In your first semester, you realized that one of your professors did not keep track of attendance and that you would not have to go to class. You skipped class when you had trouble waking up in the morning and when you had other commitments. **You could have asked your friend in class for the notes you missed**[1] but thought that reviewing the textbook would be enough. You **received a C** on the first test. There are **two tests left, each worth 20 percent**[1] of your course grade. **You need a decent grade in the course so that your GPA does not suffer too much**[2]. What action would you take?	In your first semester, you realized that one of your professors did not record attendance and that you would not have to go to class. You skipped class when you had trouble waking up in the morning and when you had other commitments. **You did not have a friend in class to ask for notes**[1] and thought that reviewing the textbook would be enough. You **received a D** on the first test. There is **one test left and it is worth 40 percent**[1] of your grade. **You cannot afford to fail the course; you will lose your scholarship if you do not maintain your GPA**[2]. What action would you take?
Note. 1 represents low *constraints*; 2 represents low *consequences.*	*Note*. 1 represents high *constraints*; 2 represents high *consequences.*
Words = 110	Words = 110
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 88.5	Flesch Reading Ease = 87.5
Flesch-Kincaid Grade Level = 4.8	Flesch-Kincaid Grade Level = 4.4

a. Attend all of your classes from this time on.
b. Start going to all classes; contact the TA to help clarify problems with the exam, and get help with the new material.
c. Go to class and study very hard.
d. Make sure you go to all classes and ask your professor to reward class attendance.
e. **Try to skip class less frequently**; **ask someone** for notes when you sleep in.
f. Spend more time reading the textbook to do better on the next exam. (added)

Table 32. **Adaptability 8**

Weaker	Stronger
After returning to school after an extended absence, you learn that you have a **10-page paper due in three days**[1] in one of your classes. It is **on a topic you have some familiarity with**[1]. Prior to leaving, you checked your syllabus and asked the professor what you would be missing by being out so that you could take care of it before you left. There was no mention of the paper. What would you do?	After returning to school after an extended absence, you learn that you have a **25-page paper due the next day**[1] in one of your classes. It is **on a topic you have no familiarity with**[1]. Prior to leaving, you checked your syllabus and asked the professor what you would be missing by being out so that you could take care of it before you left. There was no mention of the paper. What would you do?
Note. 1 represents low *constraints.*	*Note*. 1 represents high *constraints.*
Words = 76	Words = 76
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 78	Flesch Reading Ease = 78
Flesch-Kincaid Grade Level = 6.2	Flesch-Kincaid Grade Level = 6.2

a. Write the best paper you can in the time allotted. Then, when turning it in, speak with the professor about your absence and the possible miscommunication.
b. Contact the professor as soon as you are aware of the paper, explain the situation, and ask for his/her suggestion.
c. Go to class the next day and argue for a week's extension.
d. Ask the professor privately for an extension.
e. Write the paper and submit it when complete, regardless of the deadline imposed while you were gone.
f. Inform the professor that the paper is not on the syllabus, and you were not made aware, then agree on a reasonable due date.

Table 33. **Adaptability 9**

Weaker	**Stronger**
Half way into the first semester you realize that you have too much to do. You are taking **three courses, working 15-20 hours a week**[1], and your boyfriend/girlfriend **wants to**[1] spend more time together. You feel you must go home **every other weekend**[1] to visit a sick relative. **You find that you are often arriving late to classes and receiving "Bs" and "Cs" on assignments**[2]**.** What action would you take?	Half way into the first semester you realize that you have too much to do. You are taking **four courses, working 25-30 hours a week**[1], and your boyfriend/girlfriend is **constantly demanding**[1] that you start spending more time together. You feel you must go home **every weekend**[1] to visit a sick relative. **You find that you are missing classes and receiving "Ds" and "Fs" on assignments**[2]**.** What action would you take?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 70	Words = 70
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 70.8	Flesch Reading Ease = 68.5
Flesch-Kincaid Grade Level = 6.9	Flesch-Kincaid Grade Level = 7.2

a. Drop a course or take it without being graded.
b. Cut down on the hours you work.
c. Cut back a little bit on everything: classes, work, the relationship, and weekend visits.
d. Ask your boyfriend/girlfriend to be understanding and support your need for more time for school.
e. Prioritize your responsibilities and goals. Then divide up time based on importance.
f. Try to save time by calling/writing the sick relative instead of visiting, unless it was really serious.

Table 34. **Adaptability 10**

Weaker	Stronger
You are going through an especially busy period at school. It is the end of the semester; you have papers due and need to prepare for exams. In the meantime, **coworkers at your part-time job are asking that you work more shifts to help them out. Your parents, however, want you to spend less time working and more time studying[1].** You find yourself beginning to lose track of details and are feeling overwhelmed. **You expect that if things continue this way, you will get a "C" in one of your classes[2].** What do you do?	You are going through an especially busy period at school. It is the end of the semester; you have papers due and need to prepare for exams. In the meantime, **coworkers at your part-time job are asking that you work more shifts to help them out. Your parents expect you to pay your own way through school[1].** You find yourself beginning to lose track of details and are feeling overwhelmed. **You are sure that if things continue this way, you are going to get an "F" in two of your classes[2].** What do you do?
Note. 1 represents low *consistency*; 2 represents low *consequences*.	*Note*. 1 represents high *consistency*; 2 represents high *consequences*.
Words = 95	Words = 95
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 82.5	Flesch Reading Ease = 86.1
Flesch-Kincaid Grade Level = 4.6	Flesch-Kincaid Grade Level = 4.1

a. Decide what's important and then prioritize your responsibilities.
b. Relax and take a step back, knowing that you can't do everything at once.
c. Get organized, and start planning ahead and scheduling.
d. Apologize, decline the extra shifts, and tell your coworkers that school is your first priority.
e. Quit your job.
f. Sleep less and work harder to get things done.

Table 35. **Adaptability 11**

Weaker	Stronger
You are interested in finance, but do not have further finance courses for at least another semester**. You've asked your parents for suggestions but they were unable to offer any helpful advice[1]. This will not be a very busy semester for you[2].** What would you do?	You are interested in finance, but do not have further finance courses for at least another semester. **You've spoken with your advisor and she recommended learning as much as you can about finance[1]. This will be a very busy semester for you[2].** What would you do?
Note. 1 represents low *clarity*; 2 represents low *constraints*.	*Note*. 1 represents high *clarity*; 2 represents high *constraints*.
Words = 46	Words = 46
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 64.5	Flesch Reading Ease = 64.5
Flesch-Kincaid Grade Level = 7.1	Flesch-Kincaid Grade Level = 7.1

a. Wait until the next semester, and take another class then.
b. Try to register for an alternative finance course as an elective.
c. Use the semester to do some independent study so that you are well prepared for the next course.
d. Get involved in on-campus finance clubs or investment games.
e. Find out if you could be a TA for a finance class.

Table 36. **Perseverance 1**

Weaker	Stronger
Your professor has just given you a project **worth 10% of your overall grade**[1] that will probably require a good part of the semester to complete. She gave you **some of the details**[2] you need to get started, **and you are not sure how the project should proceed**. She **may or may not intend**[2] to give you any more information in class. How would you proceed?	Your professor has just given you a project **worth 60% of your overall grade**[1] that will probably require a good part of the semester to complete. She gave you **all the details**[2] you need to get started, **but you want to be sure that you understood everything correctly**. You know she **does not plan**[2] to answer questions about the project in class. How would you proceed?
Note. 1 represents low *consequences;* 2 represents low *clarity.*	*Note*. 1 represents high *consequences;* 2 represents high *clarity.*
Words = 66	Words = 66
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 74.7	Flesch Reading Ease = 69.5
Flesch-Kincaid Grade Level = 6.9	Flesch-Kincaid Grade Level = 7.6

a. Work out the project to the best of your ability and approach the professor if you get stuck.
b. Generate some ideas, and then go to office hours to see how the professor responds to them.
c. Ask the professor about the project after class.
d. Visit the professor or a teaching assistant during office hours to discuss the project.
e. Talk to other students to get an idea of what they are doing.
f. Try to get an idea of whether or not other students seem confused. If so, bring the issue up with the professor during class.

Table 37. **Perseverance 2**

Weaker	Stronger
You are collaborating with other classmates on a project. The group of you keeps running into a variety of problems that threaten to cause the project to be late. **Group opinion is split on whether you should just plan to submit it late**[1]. You know the instructor **penalizes late work with half a letter grade for each day it is late**[2]. Another option would be to devote much more time than planned to the project and possibly get it in on time. **Your group members could probably find the time**[3]. What would you do?	You are collaborating with other classmates on a project. Your group keeps running into a variety of problems that threaten to cause the project to be late. **All the other group members want to just plan to submit the project late**[1]. But you know the instructor **penalizes late work with one letter grade for each day it is late**[2]. Another option would be to devote much more time than planned to the project and possibly get it in on time. **Your group members tell you they don't have the time**[3]**.** What would you do?
Note. 1 represents low *consistency*; 2 represents low *consequences*; 3 represents low *constraints*.	*Note*. 1 represents high *consistency*; 2 represents high *consequences*; 3 represents high *constraints*.
Words = 94	Words = 94
Passive Sentences = 14%	Passive Sentences = 0%
Flesch Reading Ease = 74.4	Flesch Reading Ease = 77.1
Flesch-Kincaid Grade Level = 6.2	Flesch-Kincaid Grade Level = 5.8

a. Try to get it done, but plan to submit it late.
b. Ask the instructor for help or for an extension. If that doesn't work, just try your best and do what you can or turn it in late.
c. Motivate the group to devote more time and work together to get it done.
d. Have the group decide what to do.
e. Work hard to finish it because there are consequences for being late and meeting deadlines is important to you.
f. Tell the instructor your situation, and ask him or her for advice.

Table 38. **Perseverance 3**

Weaker	**Stronger**
You attend a large lecture class **with 200 students**[1]. You find the class dull and boring, and are having difficulty staying awake. **The professor does not seem to care about class participation as he rarely asks questions**[1]. **Passing class exams is the only thing you have to worry about**[2]. What would you do?	You attend a lecture class **with 30 students**[1]. You find the class dull and boring, and are having difficulty staying awake. **The professor often randomly picks a name from the roster when he wants someone to answer a question**[1]**. He also keeps students on their toes with pop quizzes**[2]**.** What would you do?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 53	Words = 53
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 69.9	Flesch Reading Ease = 73.1
Flesch-Kincaid Grade Level = 6.1	Flesch-Kincaid Grade Level = 5.6

a. Do what you can to stay awake, such as drinking caffeine or sitting toward the front of the class.
b. Read the class material beforehand to make the lecture more interesting.
c. During the lecture, do some studying that is required for the course.
d. Make sure you are getting enough sleep every school night.
e. Skip the class if it is that dull and boring to you.
f. Sit at the back of the class so that the professor does not notice your eyes closing. (added)

Table 39. **Perseverance 4**

Weaker	Stronger
You have been working on a research paper **for two weeks**[1]. You initially had trouble settling on a satisfactory topic for this paper and then you had trouble finding sufficient reference material. Now you are having trouble organizing the material you do have. **The paper is due in two weeks—you would have time to start over**[1] on a new topic which sounds easier. **You did not have to inform the professor of your current topic so she would not know it if you decided to change topics**[2]**.** What approach would you take to complete this course requirement?	You have been working on a research paper **for a month**[1]. You initially had trouble settling on a satisfactory topic for this paper and then you had trouble finding sufficient reference material. Now you are having trouble organizing the material you do have. **The paper is due in one week—you might have time to start over**[1] on a new topic which sounds easier. **You had to submit your topic to the professor earlier in the semester for approval so she is expecting a paper on that topic**[2]. What approach would you take to complete this course requirement?
Note. 1 represents low *constraints*; 2 represents low *consequences**.	*Note*. 1 represents high *constraints*; 2 represents high *clarity**.
Words = 98	Words = 98
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 64.2	Flesch Reading Ease = 59.9
Flesch-Kincaid Grade Level = 8.3	Flesch-Kincaid Grade Level = 8.9

**Note*. The two situational strength dimensions are different.

a. Pick an easier topic and get on with it.
b. Ask the professor what she thinks you should do.
c. Stick with the topic you have, but ask the professor for some help with organizing the material.
d. Stick with the topic, and take your work over to the writing center to see if you can get some help.
e. Figure that a new topic that sounds easy may not be – you could have the same problem all over again, so just battle on with what you have already started.

Table 40. **Perseverance 5**

Weaker	Stronger
You have a professor you find "monstrously boring." He seems rude and longwinded and often reads material from the textbook. **He does not require attendance but gives a few extra points on exams for good attendance[1]. He posts general outlines of his lectures online[2]. You need to pass the course[2].** What would you do?	You have a professor you find "monstrously boring." He seems rude and longwinded and often reads material from the textbook. **He requires attendance and takes points away for absences[1]. He refuses to post any lecture notes online[2]. The course is required for your major, so you need to do well[2].** What would you do?
Note. 1 represents low *consequences*; 2 represents low *constraints*.	*Note*. 1 represents high *consequences*; 2 represents high *constraints*.
Words = 54	Words = 54
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 70.8	Flesch Reading Ease = 72.3
Flesch-Kincaid Grade Level = 5.6	Flesch-Kincaid Grade Level = 5.4

a. Find a seat in the back and focus on some other work during the class.
b. Continue to attend, working hard to pay attention. You know you can hang in there until the end of the semester.
c. Focus on taking as many notes as you can to make sure you completely understand the material.
d. Approach the professor and ask whether he could incorporate more interactive exercises into the lesson.
e. **Read something interesting or play games** to keep yourself occupied during the boring parts.
f. Make an effort to attend class but allow yourself to skip it on days when you are extra tired. (added)

Table 41. **Perseverance 6**

Weaker	**Stronger**
You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. You often want to raise your hand in class to ask the professor to explain something again**. Some professors are happy to clarify but others act annoyed when you ask too many questions[1]. You are not the only one struggling—other students also ask for further clarification. So you do not stand out[2].**What would you do?	You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. You often want to raise your hand in class to ask the professor to explain something again. However, **you have found that professors expect you to ask for extra help *outside of* class and to join available study groups[1]. Other students do not seem to be struggling. So you really stand out[2].** What would you do?
Note. 1 represents low *consistency**; 2 represents low *consequences*.	*Note*. 1 represents high *clarity**; 2 represents high *consequences*.
Words = 85	Words = 85
Passive Sentences = 14%	Passive Sentences = 12%
Flesch Reading Ease = 67.1	Flesch Reading Ease = 73.6
Flesch-Kincaid Grade Level = 6.9	Flesch-Kincaid Grade Level = 5.6

**Note*. The dimensions are different.

a. Talk with the professors and TAs to get help on how to study.
b. Find a study partner and work on homework and class material together.
c. **Get advice from** your parents **or close friends.**
d. Study hard, try your best, and don't worry about it.
e. Ask **your** advisor and professors if there are study groups or review sessions **you** can attend.
f. Hire a tutor for the difficult classes.

Table 42. **Perseverance 7**

Weaker	**Stronger**
You decided early in the term to do a paper on a topic you expected to be very interesting to you. **After you read more about your topic, you realized that it is less interesting than you expected**[1]. In addition, your job has taken more time than you wanted, and you have had more work in your other courses than you anticipated. The paper is **due in one week**[1] and it seems like you may have to engage in several "all-nighters" to complete it on time. **You want a good grade**[1]. What would you do?	You decided early in the term to do a paper on a topic you expected to be very interesting to you. **While you are still interested in your topic, you have found it very difficult to find information**[1]. In addition, your job has taken more time than you wanted, and you have had more work in your other courses than you anticipated. The paper is **due in three days**[1] and it seems like you may have to engage in several "all-nighters" to complete it on time. **You really need an "A"**[1]**.** What would you do?
Note. 1 represents low *constraints.*	*Note*. 1 represents high *constraints.*
Words = 95	Words = 95
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 72.3	Flesch Reading Ease = 70.7
Flesch-Kincaid Grade Level = 7.1	Flesch-Kincaid Grade Level = 7.3

a. Seek help from other students who may have had a similar experience.
b. Pick a topic that can be completed quicker. An "A" is an "A".
c. Set up a schedule on which you can complete all of the other work you need to do, spend as much time on the paper as possible, and meet with the instructor to discuss what you have so far and get suggestions.
d. Do whatever it takes to complete the paper, including "all-nighters".
e. Talk to the instructor about the situation and ask for advice.
f. Make the paper a priority, but take into account how much the paper is worth in the class.

Table 43. **Perseverance 8**

Weaker	**Stronger**
You are half way through a summer sports camp. You are finding the training difficult and tedious (e.g., long runs, difficult and lengthy exercises) even though **you are in very good shape**[1]. You have not learned anything new about the sport in which you are interested. **It is not clear to you why the coaches find the intensive training necessary**[2]. **You also don't know how open they are to feedback**[2]. How would you react?	You are half way through a summer sports camp. You are finding the training exceptionally difficult and tedious (e.g., long runs, difficult and lengthy exercises), particularly because **you aren't in great shape**[1]. You have not learned anything new about the sport in which you are interested. **The coaches say they want to get everyone into better shape—they believe this is key**[2]. **They have also mentioned being open to suggestions**[2]**.** How would you react?
Note. 1 represents low *constraints*; 2 represents low *clarity*.	*Note*. 1 represents high *constraints*; 2 represents high *clarity*.
Words = 74	Words = 74
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 73.1	Flesch Reading Ease = 62.8
Flesch-Kincaid Grade Level = 6.1	Flesch-Kincaid Grade Level = 7.5

a. Continue to train, because that is what is needed to perform.
b. See whether others at the camp feel the same way, and tell the coaches.
c. Consider not going to the camp again because it was a disappointing experience.
d. Be upset and discouraged about the camp experience, though others may be learning something.
e. Suggest new activities to the coaches to make the training more interesting.

Table 44. **Perseverance 9**

Weaker	**Stronger**
You are required to take an exam that is required of all calculus students, not specifically for your calculus course. You take and **fail this exam two times**[1]. **The exam is given once per month; you are allowed to take it up to six times per year**[1]. **You have already taken two classes toward your major but cannot continue** in **this major without the calculus credit**[2]. What would you do?	You are required to take an exam that is required of all calculus students, not specifically for your calculus course. You take and **fail this exam three times**[1]. **The exam is given every three months; you are allowed to take it up to three times per year**[1]. **You have already taken five classes toward your major but cannot continue in this major without the calculus credit**[2]. What would you do?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 70	Words = 70
Passive Sentences = 20%	Passive Sentences = 20%
Flesch Reading Ease = 72.9	Flesch Reading Ease = 70.5
Flesch-Kincaid Grade Level = 5.9	Flesch-Kincaid Grade Level = 6.3

a. Keep taking the exam in the hopes that you will eventually pass.
b. Begin to think of some other options.
c. Talk with the director of the program and explain that the exam covers material that is not part of any classes, find out where you can get help with this material.
d. Try taking the class from a different instructor and join a study group.
e. Find someone to tutor you with calculus specifically as it relates to the exam material.
f. Change your major.

Table 45. **Perseverance 10**

Weaker	Stronger
You are having problems in one of your classes. **The subject is kind of interesting and useful but some concepts are hard to wrap your mind around[1].** The second exam is **three weeks away[1]**, and you want to do well on it, especially considering the **"C" you received on the first exam[1]**. You have worked hard so far in college to maintain a good GPA and don't want to sacrifice it. **The professor makes himself available during office hours[1]**. What would you do?	You are having problems in one of your classes. **A lot of concepts are hard to wrap your mind around and you have zero interest in the subject[1].** The second exam is **next week[1]**, and you want to do well on it, especially considering the **"D" you received on the first exam[1]**. You have worked hard so far in college to maintain a good GPA and don't want to sacrifice it. **The professor's office hours conflict with your schedule[1]**. What would you do?
Note. 1 represents low *constraints*.	*Note*. 1 represents high *constraints*.
Words = 83	Words = 83
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 74.5	Flesch Reading Ease = 79.6
Flesch-Kincaid Grade Level = 6.2	Flesch-Kincaid Grade Level = 5.5

a. Continue to study and not give up even when you encounter difficult concepts.
b. Make an appointment and consult the professor for advice on studying and how to approach the material.
c. Set aside more time to study than you did studying for the previous exam.
d. Accept that you cannot always get great grades and focus your time and energy on other classes that you are doing well in.
e. Realistically assess how much your GPA will suffer if you don't do well in this course; it may not be a huge deal. (added)

Table 46. **Ethics 1**

Weaker	Stronger
You and your friends know that a mutual acquaintance has been dating someone for **six months**[1]. However, one of your friends tries her best to get a date with this individual. **She tells you about all her "schemes" but has not mentioned if she's expecting your help**[1]**. You have heard the couple talk of moving in together but your friend says they also argue a lot**[2]**.** How would you react?	You and your friends know that a mutual acquaintance has been dating someone for **over two years**[1]. However, one of your friends tries her best to get a date with this individual. **She wants you to play an active role in her "schemes" to split the couple up**[1]**. You have observed the couple get along very well and heard them talk about getting married after college**[2]**.** How would you react?
Note. 1 represents low *clarity*; 2 represents low *consistency*.	*Note*. 1 represents high *clarity*; 2 represents high *consistency*.
Words = 70	Words = 70
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 83.8	Flesch Reading Ease = 79.0
Flesch-Kincaid Grade Level = 5.0	Flesch-Kincaid Grade Level = 5.7

a. If the acquaintance is in a happy relationship, tell your friend to wait and rethink it. If not, it is OK to get a date.
b. Support my friend. Good for her.
c. Tell your friend to wait until the person is single or to just forget about the person.
d. Tell your friend that it is inappropriate to interfere in the relationship.
e. Just be annoyed at your friend. Do not get involved.

Table 47. **Ethics 2**

Weaker	**Stronger**
You know that a group of students in your class cheats on exams by putting formulas into scientific calculators or into cell phones. **The professor has warned against such activity but does not consistently use proctors to prevent cheating[1]. You expect that if the professor caught these students cheating, she would give them "Fs" on the exam[2]. You are not friends with the cheaters and sit far from them in class[3]**. What action would you take?	You know that a group of students in your class cheats on exams by putting formulas into scientific calculators or into cell phones. **The professor has clearly warned against such activity and consistently uses proctors to prevent cheating[1]. You expect that if the professor found out, she would give these students "Fs" for the entire course[2]. You are friends with two of the cheaters and sit close to them in class[3]**. What action would you take?
Note. 1 represents low *consistency*; 2 represents low *consequences*; 3 represents low *constraints*.	*Note*. 1 represents high *consistency*; 1 represents high *consequences*; 2 represents high *constraints*.
Words = 76	Words = 76
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 71.1	Flesch Reading Ease = 68.9
Flesch-Kincaid Grade Level = 7.1	Flesch-Kincaid Grade Level = 7.4

a. Try doing the same thing until people start getting caught.
b. Study the way you know best, don't cheat, but don't turn in the other students either.
c. Do nothing; it's none of your business.
d. Mention it to the professor so she can deal with the problems in the class.
e. Don't tell the professor, but make sure it is clear you are not involved in case they get caught.
f. Send the professor an anonymous message about what is going on.

Table 48. **Ethics 3**

Weaker	Stronger
In a class of about **200 students**[1], you discover that a group of your **classmates**[1] have worked out a scheme to share answers on an exam. **They might be open to your joining them**[2]. The professor has vision problems and will likely never notice. You are not doing very well in the class and **anticipate getting a "C" in the course**[3]**. One of the previous two exams was very hard**[4]**.**What would you do in these circumstances?	In a class of **40 students**[1], you discover that a group of your **close friends**[1] have worked out a scheme to share answers on an exam. **They are encouraging you to join them**[2].The professor has vision problems and will likely never notice. You are not doing very well in the class and **anticipate a "D" or "F" in the course**[3]**. Both of the previous two exams were very hard**[4]. What would you do in these circumstances?
Note. 1 represents low *constraints*; 2 represents low *clarity*; 3 represents low *consequences*; 4 represents low *consistency*.	*Note*. 1 represents high *constraints*; 2 represents high *clarity*; 3 represents high *consequences*; 4 represents high *consistency*.
Words = 76	Words = 76
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 72.2	Flesch Reading Ease = 78.9
Flesch-Kincaid Grade Level = 6.9	Flesch-Kincaid Grade Level = 6.0

a. Avoid being around these **people**.
b. It is not exactly honest but under the circumstances, the scheme is OK. You would join them.
c. Do your own work and not tell the professor about the scheme because it is not your problem.
d. Cheat **because you really have to** get a good grade.
e. Tell the professor about the scheme.
f. Study for the exam, but join the scheme as a backup strategy for the test.

Table 49. **Ethics 4**

Weaker	Stronger
One of your friends takes the same class you do, but at an earlier time. On the day of the midterm, he comes to your room and offers you a copy of the exam he just took. You know that the professor gives the same exam to all sections. **Some students say that this professor's tests are difficult; others say the tests are manageable[1]. You have spent five hours studying and have four hours left[2]. The exam is worth 20 percent of your final grade[3]**. What would you do?	One of your friends takes the same class you do, but at an earlier time. On the day of the midterm, he comes to your room and offers you a copy of the exam he just took. You know that the professor gives the same exam to all sections. **You have heard from many students that this professor's tests are very difficult[1]. You have spent only a couple of hours studying and have two hours left[2]. The exam is worth 45 percent of your grade[3].** What would you do?
Note. 1 represents low *consistency*; 2 represents low *constraints*; 3 represents low *consequences*.	*Note*. 1 represents high *consistency;* 2 represents high *constraints*; 3 represents high *consequences*.
Words = 89	Words = 89
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 88.1	Flesch Reading Ease = 87.4
Flesch-Kincaid Grade Level = 3.7	Flesch-Kincaid Grade Level = 4.2

a. Take the exam to study from, but not to memorize.
b. Tell your friend what you have studied and ask him whether that will be sufficient.
c. Look the exam over so that you can study the areas that you don't know well, using it as a study guide.
d. Thank your friend, and politely decline.
e. Decline the exam, and suggest to the professor that in the future she should have different forms of the exam.

Table 50. **Ethics 5**

Weaker	Stronger
As you walk out to pick up dinner[1], you find an expensive mobile phone on the edge of a sidewalk along the street. **There is no one around. You turn it on and find that it is not damaged and fully charged**[1]. There is no identifying information on the phone. **You own a mobile phone but yours is not as expensive**[1]. What would you do with the phone?	**As you are rushing to your next class with two friends**[1], you find an expensive mobile phone on the edge of a sidewalk along the street. **You try to turn it on and find that the battery is almost dead**[1]. There is no identifying information on the phone. **You recently ruined your own phone by accidentally dropping it in a pool**[1]. What would you do with the phone?
Note. 1 represents low *constraints.*	*Note*. 1 represents high *constraints.*
Words = 68	Words = 68
Passive Sentences = 16%	Passive Sentences = 0%
Flesch Reading Ease = 84.6	Flesch Reading Ease = 79.8
Flesch-Kincaid Grade Level = 4.2	Flesch-Kincaid Grade Level = 5.5

a. Leave it there **on the sidewalk**. The person who lost it might come back looking for it.
b. Since there is no identification, take it, and have the number changed.
c. Take it. Either wait until someone calls, or call a number on the phone. Figure out who the owner is and return it.
d. Put a 'lost and found' ad in the paper, or put up signs about the lost phone.
e. Make some effort to return it, but keep it if you do not find the owner.
f. Turn it in to the authorities or 'lost and found.' Alternatively, return it to a phone store/dealer.

Table 51. **Ethics 6**

Weaker	Stronger
You have been having trouble in a difficult course. On the day of the **first**[1] exam someone offers to sell you a copy of the exam. **It is possible that other students in the class will purchase this exam. If that is the case and many students do well, your poor grade will look even worse in comparison. The professor curves test grades, but if many students do well, you will not benefit from a curve**[2]. What would you do?	You have been having trouble in a difficult course. On the day of the **last**[1] exam someone offers to sell you a copy of the exam. **You have heard rumors that a large number of students in the class have purchased this exam. If many students suddenly do well, the professor will surely realize that students must have been cheating; if an investigation is launched and you are caught, you will fail the course or worse**[2]**.** What would you do?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 80	Words = 80
Passive Sentences = 0%	Passive Sentences = 20%
Flesch Reading Ease = 78.0	Flesch Reading Ease = 76.9
Flesch-Kincaid Grade Level = 5.6	Flesch-Kincaid Grade Level = 5.8

a. Purchase the exam so that you do not look worse than everybody else.
b. Take the exam with the knowledge that you have from studying.
c. Approach the instructor and explain what you have been offered.
d. Buy the exam, but try to miss some of the questions.
e. Inform the instructor anonymously in the hopes that he will change the exam.

Table 52. **Ethics 7**

Weaker	Stronger
One of your friends wants to copy the homework of another friend. **They have the same class but different instructors[1]. So it is highly unlikely that your friends would get caught[2].** The friend who has completed the work refuses to let your other friend copy the work. The second friend becomes very angry, complains to you, and **stares expectedly at you**. **You cannot tell what she expects from you[3]**. How would you react?	One of your friends wants to copy the homework of another friend. **They have the same instructor, but different TAs check their homework[1]. If TAs work together, your friends would get into big trouble for academic dishonesty[2]**. The friend who has done the work refuses to let your other friend copy it. The second friend becomes very angry, complains to you, and **asks you to intervene on her behalf[3].** How would you react?
Note. 1 represents low *constraints*; 2 represents low *consequences*; 3 represents low *clarity*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*; 3 represents high *clarity*.
Words = 73	Words = 73
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 79.2	Flesch Reading Ease = 72.8
Flesch-Kincaid Grade Level = 4.8	Flesch-Kincaid Grade Level = 6.1

a. Tell her to quit complaining and just do the assignment.
b. Explain that she's better off in the long run by doing the work.
c. Suggest to your friend that she ask for help with the homework instead.
d. Tell her that the friend who refused to cheat was under no obligation. She should respect that choice.
e. Not interfere in the conflict. If anything, you would support the friend refusing to allow copying.
f. Listen to each person and try to understand his/her point of view.

Table 53. **Ethics 8**

Weaker	Stronger
You are taking a particularly difficult **midterm**[1] and **finding that you cannot answer some of the multiple choice questions**[1]. You realize it is very easy to see the paper of a student who you know is doing well in the course. When you look at this student's paper, she notices and pulls her paper away. **She seems like a nice person so it is unlikely that she will tell the professor you were trying to cheat and get you into trouble**[2]. What action would you take?	You are taking a particularly difficult **final**[1] exam and **finding that you cannot answer most of the essay questions**[1]. You realize it is very easy to see the paper of a student who you know is doing well in the course. When you look at this student's paper, she notices and pulls her paper away. **Chances are high that she will get you into trouble by telling the professor you were attempting to cheat off her**[2]. **You would lose your scholarship**[2]. What action would you take?
Note. 1 represents low *constraints*; 2 represents low *consequences*.	*Note*. 1 represents high *constraints*; 2 represents high *consequences*.
Words = 86	Words = 86
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 73.2	Flesch Reading Ease = 74.2
Flesch-Kincaid Grade Level = 7.3	Flesch-Kincaid Grade Level = 6.4

a. Don't show that you are embarrassed, and see if you can look at another paper.
b. Just finish your exam as usual and then leave the room.
c. Pretend like it didn't happen, make an effort not to look there again, and finish the test.
d. Finish the exam, and approach **or email** the student after class to apologize.
e. Finish the exam, approach the student after class, and beg her not to tell on you. (added)

Table 54. **Ethics 9**

Weaker	Stronger
You look at a classmate's paper during a **midterm**[2] exam and the person turns to you telling you to quit. When the professor notices this exchange, she asks you both to leave. She says both of you are getting "Fs" on the exam. **Receiving the "F" will make it harder for you to pass the course.**[1] **You will need to make sure to do well on the final**[2]**.** What would you do now?	You look at a classmate's paper during a **final**[2] exam and the person turns to you telling you to quit. When the professor notices this exchange, she asks you both to leave. She says both of you are getting "Fs" on the exam. **Given your earlier grades in the class, this "F" means that you will fail the entire course.**[1] **This class is required for your major**[2]**.** What would you do now?
Note. 1 represents low *consequences*; 2 represents low *constraints*.	*Note*. 1 represents high *consequences*; 2 represents high *constraints*.
Words = 73	Words = 73
Passive Sentences = 0%	Passive Sentences = 0%
Flesch Reading Ease = 86.5	Flesch Reading Ease = 89.0
Flesch-Kincaid Grade Level = 4.8	Flesch-Kincaid Grade Level = 3.8

a. Make sure that you track down the other student and apologize to him/her.
b. Go to the professor and explain that your classmate had no part in your cheating, and was simply telling you to quit.
c. Go to the professor and apologize for trying to cheat—**perhaps she will let you retake the test**.
d. Go to the professor and try to convince her that the whole situation was a misunderstanding—you were not trying to cheat.
e. Resolve to not cheat again and try to put the whole incident behind you. (added)

Table 55. **Ethics 10**

Weaker	Stronger
The professor hands back the grade and feedback on a project your team submitted. In the comments is a pointed attack on the group for plagiarizing and **half a letter grade is deducted**. **The end result is a "B"[1]. You know which group member plagiarized[3]. But you do not know what you are expected to do in this situation**. **You do not know whether or not the professor would raise your individual project grade if you were to report the plagiarizer[2].** What would you do?	The professor hands back the grade and feedback on a project your team submitted. In the comments is a pointed attack on the group for plagiarizing and **a full letter grade is deducted. The end result is a "C"[1]. The professor expects students to follow an honor code and report anyone who has plagiarized[2]. If you report the plagiarizer, the professor will raise your grade to a "B"; that would really help your course grade[1]. However**, **the plagiarizer is your friend[3]**. What would you do?
Note. 1 represents low *consequences*; 2 represents low *clarity*; 3 represents low *constraints*.	*Note*. 1 represents high *consequences*; 2 represents high *clarity*; 3 represents high *constraints*.
Words = 85	Words = 85
Passive Sentences = 28%	Passive Sentences = 14%
Flesch Reading Ease = 75.3	Flesch Reading Ease = 72.9
Flesch-Kincaid Grade Level = 5.8	Flesch-Kincaid Grade Level = 5.7

a. Discuss the situation with the group and come up with a decision about what to do, or not to do, together.
b. **Convince** the person responsible for the plagiarism to **talk to the professor, assume responsibility, and get the other team members' grades raised**.
c. Tell the professor who was responsible for the plagiarism **in hopes that other team members' grades will be raised**.
d. Contact the professor anonymously with the name of the individual responsible for the plagiarism. (added)
e. Accept the grade, learning from the experience and vowing never to let it happen again.

APPENDIX C: SME Instructions for SJT Stem Ratings (Weak and Strong Stems Rated Separately)

Your task will be to rate situations from a situational judgment test on their situational strength. Thank you again for agreeing to help with this task!

The table below provides:
1. A definition of situational strength
2. Definitions of the four facets comprising situational strength

Table 56. **Information for SMEs**

Situational Strength – The extent to which the appropriate course of action in the situation is clear.			
Clarity	**Consistency**	**Constraints**	**Consequences**
Degree to which cues regarding one's responsibilities are available and comprehensible.	Degree to which cues regarding one's responsibilities are consistent with each other.	Degree to which one's decision-making and actions are constrained by external forces one cannot control.	Degree to which one's decisions or actions have considerable consequences (positive, negative) for the self or relevant others.
Clear situations… • Provide specific/easy to understand info about responsibilities • Provide straightforward info about what someone should do to succeed • Make clear what someone should expect • Provide specific info about which actions to take • Make clear what behavior is expected • Provide specific info about how others would react	Situations with high consistency… • Provide sources of info that are consistent with each other • Provide sources of info that are consistent over time • Describe responsibilities/requirements that are compatible • Provide consistent info about what to do to succeed • Provide consistent info about others' expectations/desires	Situations with high constraints… • Prevent one from making his/her own decision • Prevent one from doing things his/her own way • Have other people/outside forces limit one's freedom to make decisions • Have personal circumstances limit what one can do	Situations with high consequences… • Have important consequences of one's behavior for self/others • Have serious outcomes when one does the wrong thing • Have others put at risk when one does the wrong thing

Please review the above table carefully. Use the descriptions of the situational strength facets and the overall strength definition when rating the situations that follow.

Please note that a situation does *not* need to have *all* the qualities of a highly clear situation to be a highly clear situation. The same goes for the other situational strength facets.

Please consider each situation carefully. Double click on the box next to your desired response rating and select "checked".

1.
After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. The Red Cross will probably get enough donors to help all the people who need blood. How would you react in this situation?[1]

Please rate this situation on:

Table 57. **Rating scale for SMEs**

Clarity	Consistency	Constraints	Consequences	Situational Strength
☐ Extremely Low ☐ Very Low ☐ Somewhat Low ☐ Neither Low nor High ☐ Somewhat High ☐ Very High ☐ Extremely High	☐ Extremely Low ☐ Very Low ☐ Somewhat Low ☐ Neither Low nor High ☐ Somewhat High ☐ Very High ☐ Extremely High	☐ Extremely Low ☐ Very Low ☐ Somewhat Low ☐ Neither Low nor High ☐ Somewhat High ☐ Very High ☐ Extremely High	☐ Extremely Low ☐ Very Low ☐ Somewhat Low ☐ Neither Low nor High ☐ Somewhat High ☐ Very High ☐ Extremely High	☐ Extremely Low ☐ Very Low ☐ Somewhat Low ☐ Neither Low nor High ☐ Somewhat High ☐ Very High ☐ Extremely High

[1]*Note*. This is the weak version of the situation.

APPENDIX D: SME Instructions for SJT Stem Ratings (Comparison of Weak and Strong Stems)

Your task will be to compare situations from a situational judgment test (for college students) on their situational strength. Thank you for agreeing to help with this task!

The table below provides:
1. A definition of situational strength
2. Definitions of the four facets comprising situational strength

Table 58. **Information for SMEs**

Situational Strength – The extent to which the appropriate course of action in the situation is clear.			
Clarity	**Consistency**	**Constraints**	**Consequences**
Degree to which cues regarding one's responsibilities are available and comprehensible.	Degree to which cues regarding one's responsibilities are consistent with each other.	Degree to which one's decision-making and actions are constrained by external forces one cannot control.	Degree to which one's decisions or actions have considerable consequences (positive, negative) for the self or relevant others.
Clear situations… • Provide specific/easy to understand info about responsibilities • Provide straightforward info about what someone should do to succeed • Make clear what someone should expect • Provide specific info about which actions to take • Make clear what behavior is expected • Provide specific info about how others would react	Situations with high consistency… • Provide sources of info that are consistent with each other • Provide sources of info that are consistent over time • Describe responsibilities/requirements that are compatible • Provide consistent info about what to do to succeed • Provide consistent info about others' expectations/desires	Situations with high constraints… • Prevent one from making his/her own decision • Prevent one from doing things his/her own way • Have other people/outside forces limit one's freedom to make decisions • Have personal circumstances limit what one can do	Situations with high consequences… • Have important consequences of one's behavior for self/others • Have serious outcomes when one does the wrong thing • Have others put at risk when one does the wrong thing

Please review the above table carefully. Use the descriptions of the situational strength facets when comparing the situations that follow.

Please consider each situation carefully. Double click on the box next to your desired response rating and select "checked" (or highlight the desired option if that's easier).

Please note that you do not need to select responses to the situations; the response options are provided for context.

A. After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. The Red Cross will probably get enough donors to help all the people who need blood. How would you react in this situation?

B. After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. The Red Cross clearly has a shortage of blood donors. Several of your close relatives need blood. How would you react in this situation?

a. Encourage others to donate blood.
b. Donate money to the Red Cross instead.
c. Volunteer your time to generate money for the Red Cross.
d. Volunteer to give out cookies and help at the blood drives.
e. Ask the Red Cross if you could help them in any other way.
f. Convince your best friend to donate blood since you cannot.

Please compare these two versions of the situation on:

Table 59. **Rating scale for SMEs**

Clarity	Consistency	Constraints	Consequences	Situational Strength
☐ A Extremely Higher ☐ A Much Higher ☐ A Somewhat Higher ☐ A and B Equal ☐ B Somewhat Higher ☐ B Much Higher ☐ B Extremely Higher	☐ A Extremely Higher ☐ A Much Higher ☐ A Somewhat Higher ☐ A and B Equal ☐ B Somewhat Higher ☐ B Much Higher ☐ B Extremely Higher	☐ A Extremely Higher ☐ A Much Higher ☐ A Somewhat Higher ☐ A and B Equal ☐ B Somewhat Higher ☐ B Much Higher ☐ B Extremely Higher	☐ A Extremely Higher ☐ A Much Higher ☐ A Somewhat Higher ☐ A and B Equal ☐ B Somewhat Higher ☐ B Much Higher ☐ B Extremely Higher	☐ A Extremely Higher ☐ A Much Higher ☐ A Somewhat Higher ☐ A and B Equal ☐ B Somewhat Higher ☐ B Much Higher ☐ B Extremely Higher

APPENDIX E: SJT Item Keying Procedure

Table 60. **SJT item keying procedure**

1. Identify the "best" alternative(s)[1] for **each item**.	1. Temporarily consider the alternative with the highest mean effectiveness rating to be the "best". 2. Also consider "best" any alternatives with mean effectiveness ratings within the critical difference of the temporary "best" one. Compute the critical difference between mean ratings for two alternatives necessary for significance at $p < .10$. The size of the critical difference will depend on the standard deviation of effectiveness ratings and the number of raters. 3. If at least 10% of raters judged one of the "best" alternatives to be "worst", exclude that alternative from the "best" set. 4. Sum the percentages of raters who judged an alternative in the remaining set of "best" ones to be best. a. If the total of these percentages is at least 51%, those alternatives will be keyed "best". b. If at least one alternative was dropped from the "best" set during step 3, and the endorsement percentages for the remaining "best" alternatives sum to less than 51%, none of the alternatives will be keyed "best." c. If no alternatives were dropped during step 3, and the endorsement percentages for the "best" alternatives sum to less than 51%, include as "best" the alternative with the next highest mean effectiveness rating (unless it violates the rule in step 3). Next, if the endorsement percentages for the "best" alternatives sum to at least 51% or greater, key all those as "best" responses and stop. If the total percentages sum to less than 51%, include as "best" the alternative with the next highest mean effectiveness rating (unless it violates the decision rule in step 3). If the total percentages sum to at least 51%, key all those as "best" responses and stop. If these rules cannot be used to add alternatives to the set of "best" ones and get a total percentage of "best" responses of 51% or more, do not key any alternative as "best."
2. Identify the "worst" alternative(s)[1] for **each item**.	Steps will be analogous to those above.
3. Identify the "neither best nor worst" alternative(s) for **each item**.	Alternatives that are not keyed as "best" or "worst", will be keyed "neither best nor worst."

Note. [1]The number of "best"/"worst" alternatives for an item can technically range from zero to the total number of alternatives. The source of the described procedure is Friede et al. (2003).

APPENDIX F: Biodata Items by Dimension

Below are questions about your background history and life experiences. Please select the answer that best describes you.

Citizenship

1. How many times in the past year have you volunteered in social service or charity organizations?
   a. Never
   b. Once
   c. Twice
   d. Three
   e. Four times or more

2. During the past two years, how many times did you work with not-for-profit groups?
   a. 0
   b. 1
   c. 2
   d. 3 or 4
   e. 5 or more

3. How many hours of volunteer work did you do while in high school?
   a. 0
   b. Between 1 and 10
   c. Between 11 and 30
   d. Between 31 and 75
   e. More than 75

4. During the last year, how many times have you given money, food, or clothes to a charity or a poor person in need?
   a. 0
   b. 1
   c. 2
   d. 3
   e. More than 3

5. In the past year, how many hours were you engaged in community service or volunteer activities?
   a. None
   b. Less than 10 hours
   c. 11-40 hours
   d. 41-80 hours
   e. More than 80 hours

6. How important has it been in the past for you to be involved in community or volunteer work?

a. Extremely important
b. Very important
c. Important
d. Not very important
e. Not at all important

7. In the past, how likely were you to help a stranger in need (e.g., giving directions to a lost person)?
a. Extremely Likely
b. Very Likely
c. Somewhat Likely
d. Not very Likely
e. Not at all Likely

8. In the past year, in how many fundraisers have you participated?
a. None
b. 1
c. 2
d. 3
e. 4 or more

9. During the past year, how often have you recycled?

a. Never
b. Not very often
c. Sometimes
d. Often
e. Always

10. Of the town mayor, the state governor, and your state senators, how many of them can you name?
a. None
b. 1
c. 2
d. 3
e. 4

Adaptability

1. In the past, how difficult have you found it to adjust to major changes in your life (e.g., moving, a new school, a new job)?
a. Extremely Difficult
b. Very Difficult
c. Difficult
d. Not Very Difficult
e. Not at all Difficult

2. Compared with others, how long does it take you to feel comfortable in new situations or places?

a. A very long time
b. A long time
c. Neither a short nor a long time
d. A short time
e. A very short time

3. How difficult has it been for you to deal with situations that forced you to make adjustments to your daily life (e.g., a broken leg, illness, or family crisis)?

a. Very difficult
b. Difficult
c. Not easy but not difficult
d. Easy
e. Very easy

4. How often have you failed to meet responsibilities because you had taken on too much?

a. Very often
b. Often
c. Sometimes
d. Rarely
e. Never

5. How difficult has it been for you to continue with something after being interrupted and having to take care of something else?

a. Very easy
b. Easy
c. Not easy but not difficult
d. Difficult
e. Very difficult

6. How often do you plan ahead and make a specific schedule of things you need or want to do?

a. Very often
b. Often
c. Sometimes
d. Rarely
e. Never

7. How effective would others say you are at handling multiple projects simultaneously?

a. Much more effective than most people
b. Somewhat more effective than most people
c. About as effective as most people
d. Somewhat less effective than most people
e. Much less effective than most people

8. In the past, how difficult has it been for you to change your study habits to improve on a skill or to do better in a class?
   a. Very difficult
   b. Difficult
   c. Not easy but not difficult
   d. Easy
   e. Very easy

9. To what extent have you been bothered by sudden changes in your schedule?
   a. To a great extent
   b. To a large extent
   c. To a moderate extent
   d. To a slight extent
   e. Not at all

10. When you are working on a serious and relatively difficult task and a phone call interrupts you, how do you usually react?
   a. With a great deal of annoyance - it is hard to get back to the original task
   b. You are irritated - it's hard to stay on task when you are interrupted
   c. It bothers you just a little - you'd really prefer not to be interrupted
   d. It doesn't bother you - you feel one of the challenges of any job is the ability to "juggle" several things at a time

Perseverance

1. To what extent would your friends describe you as someone who goes after what you want?
   a. Not at all
   b. A slight extent
   c. A moderate extent
   d. A large extent
   e. A great extent

2. How frequently do you fail to get what you want because you did not put in enough effort?
   a. Very often
   b. Often
   c. Sometimes
   d. Rarely
   e. Never

3. To what extent has it been important to you to do your very best whenever you take on a project?
   a. Extremely important
   b. Very important
   c. Important
   d. Not very important
   e. Not at all important

4. How often have you accomplished something you initially thought was very difficult or almost impossible?
   a. Very often
   b. Often
   c. Sometimes
   d. Rarely
   e. Never

5. How often have you finished a project when faced with difficult circumstances?
   a. Very often
   b. Often
   c. Sometimes
   d. Rarely
   e. Never

6. How often do others tend to compliment you on your determination to continue with a project under difficult circumstances?
   a. Very often
   b. Often
   c. Sometimes
   d. Rarely
   e. Never

7. How often have you achieved a personal goal that seemed unattainable at first?
   a. Very often
   b. Often
   c. Sometimes
   d. Rarely
   e. Never

8. How often do you tend to give up on a task after being told that you were not doing well?
   a. Almost all the time
   b. Most of the time
   c. Sometimes
   d. Rarely
   e. Never

9. How important is it to you to succeed in whatever task you are engaged in?
   a. Extremely important
   b. Very important
   c. Important
   d. Not very important
   e. Not at all important

10. When encountering problems that take a long time to solve, how impatient do you tend to become?

a. Extremely impatient
b. Very impatient
c. Somewhat impatient
d. Slightly impatient
e. Not at all impatient

Ethics

1. In the past, how likely were you to return money that you received by accident (e.g., received extra change after buying something)?

a. Much more likely than most people
b. Somewhat more likely than most people
c. About as likely as others
d. Somewhat less likely than most people
e. Much less likely than most people

2. During high school, how many times have you expressed disapproval or anger at a friend for behaving in a manner that you considered to be unethical or wrong?

a. Never
b. Once
c. Twice
d. Three or Four
e. Five or More

3. In the past year, how many times have you copied someone else's work and submitted it as your own (at school or at work)?

a. Never
b. Once
c. Twice
d. Three or four times
e. More than five times

4. In high school, how many times have you cheated on a school project, assignment, or test?

a. 0
b. 1
c. Two or three
d. Four to ten
e. More than ten

5. When you have found someone else's belongings, how often have you attempted to return them?

a. Always
b. Most of the time
c. Half of the time

d. Less than half of the time
e. I have never found someone else's belongings

6. In high school, how many times were you given detention (or a similar punishment)?
a. Never
b. Once
c. Twice
d. Three of four times
e. Five times or more

7. In your first three years of high school, how often did you skip classes without a legitimate reason?
a. Most of the time
b. A lot
c. Sometimes
d. Once or twice
e. Never

8. If a fellow student offered you a copy of upcoming exam questions that he had retrieved from the teacher's recycling bin, how likely would you be to accept a copy?

a. Extremely likely
b. Quite likely
c. Somewhat unlikely
d. Not at all likely

9. If you were struggling with a school assignment, and a fellow student with more expertise offered to finish it for you, how likely is it that you would accept the offer?
a. Extremely likely
b. Quite likely
c. Somewhat unlikely
d. Not at all likely

10. How many times have you been accused of acting unethically?
a. Very often
b. Often
c. Sometimes
d. Rarely
e. Never

APPENDIX G: Personality Items

Below are phrases describing people's behaviors. Please use the provided rating scale to indicate how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself in relation to other people you know of the same sex as you are and roughly your same age.

1. Make people feel at ease. (A+)
2. Have a rich vocabulary. (O+)
3. Don't talk a lot. (E-)
4. Have difficulty understanding abstract ideas. (O-)
5. Am interested in people. (A+)
6. Feel comfortable around people. (E+)
7. Follow a schedule. (C+)
8. Insult people. (A-)
9. Get chores done right away. (C+)
10. Make a mess of things. (C-)
11. Sympathize with others' feelings. (A+)
12. Don't mind being the center of attention. (E+)
13. Keep in the background. (E-)
14. Leave my belongings around. (C-)
15. Feel little concern for others. (A-)
16. Change my mood a lot. (ES-)
17. Often forget to put things back in their proper place. (C-)
18. Am full of ideas. (O+)
19. Feel others' emotions. (A+)
20. Have a soft heart. (A+)
21. Pay attention to details. (C+)
22. Shirk (i.e. skip out on) my duties. (C-)
23. Am not interested in other people's problems. (A-)
24. Am the life of the party. (E+)
25. Am always prepared. (C+)
26. Get irritated easily. (ES-)
27. Have excellent ideas. (O+)
28. Use difficult words. (O+)
29. Get stressed out easily. (ES-)
30. Start conversations. (E+)
31. Get upset easily. (ES-)
32. Do not have a good imagination. (O-)
33. Am relaxed most of the time. (ES+)
34. Often feel blue. (ES-)
35. Talk to a lot of different people at parties. (E+)
36. Have frequent mood swings. (ES-)
37. Take time out for others. (A+)
38. Spend time reflecting on things. (O+)
39. Have a vivid imagination. (O+)

40. Am not interested in abstract ideas. (O-)
41. Don't like to draw attention to myself. (E-)
42. Like order. (C+)
43. Seldom feel blue. (ES+)
44. Have little to say. (E-)
45. Worry about things. (ES-)
46. Am exacting in my work. (C+)
47. Am quick to understand things. (O+)
48. Am easily disturbed. (ES-)
49. Am quiet around strangers. (E-)
50. Am not really interested in others. (A-)

Note. E = Extraversion; A = Agreeableness; C = Conscientiousness; ES = Emotional Stability; O = Openness to Experience. Positively scored items indicated with '+'; negatively scored items indicated with '-'.

APPENDIX H: Behaviorally Anchored Rating Scales (BARS)

Below, you will be asked to rate yourself on several dimensions of college performance. Rate yourself according to how well you have done in college so far and believe that you will do in the future.

Please read the definition of each performance dimension and the provided examples **before you make your rating**. The two examples may not be exactly related to your behavior on a given dimension, but they give you ideas of excellent, acceptable, and poor levels of performance.

Remember that everyone has both strengths and weaknesses. Your ratings should therefore reflect your best qualities as well as things you can develop or improve on.

Knowledge

Definition: Gaining knowledge and mastering facts, ideas and theories and how they interrelate, and the relevant contexts in which knowledge is developed and applied. Grades or GPA can indicate, but not guarantee, success on this dimension.

Before you make your rating, please read these two examples:

Example 1

You have never been very good at writing essays or papers and finds that many of your classes in college require written assignments. You get failing grades on your first two essays even though you spent a great deal of time preparing these papers, you realize that your classes require three more papers this term. How do you expect you would deal with this situation?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You continue with existing skill level, and hopes to get better at writing by the end of the course.	You keep practicing writing essays alone, and make progress on future assignments.	You go to talk to the professors and commit to submitting extra work so that you can receive extra feedback. You make use of the writing center to learn how to write better essays.

Example 2

The professor has asked each member of the class to write a paper on foreign relations policy. Students are free to select different countries as the focus of their papers. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You find someone else who is going to cover the same country, and split the work.	You choose the country about which you already have some background knowledge, and build on that.	You select a country that you know little or nothing about, and do extensive research so that you can learn from the experience.

Now, record the number corresponding to your level on this dimension.

Continuous Learning

Definition: Being intellectually curious and actively interested in continuous learning. Actively seeking new ideas and new skills, both in core areas of study as well as in peripheral or novel areas.

Before you make your rating, please read these two examples:

Example 1

You have been given a course project on a topic that is new to you. All other students are using computers to access data, information, and files to accomplish the task. You do not have the background to work on this project, and you have no knowledge of the software required to do this project. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You lose interest in the project and ask a friend for help in doing it.	You ask the professor what resources you should use to get the skills that you need.	You look for a way to learn the skills needed, perhaps buying a book about the software. You take pleasure in learning the new skills, and put them to use to complete this and future projects.

Example 2

A professor offers students a variety of special project opportunities. All these projects would definitely enhance students' understanding of the material, but the professor's grading policy is such that doing these projects will have little impact on the students' grades. Under the circumstances would you expect that you do one or more of these projects?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You do not take on any of the projects.	You take on one of the projects to get in good standing with the instructor.	You take on more than one of the projects. Your goal as a student is to gain knowledge, regardless of the grades you receive for the knowledge gained.

Now, record the number corresponding to your level on this dimension.

Multicultural Appreciation

Definition: Showing openness, tolerance, and interest in a diversity of individuals (e.g., by culture, ethnicity, or gender). Actively participating in, contributing to, or influencing a multicultural environment.

Before you make your rating, please read these two examples:

Example 1

You have had very little exposure to Native American culture, and a professor has invited you to attend an evening presentation on campus that will address Native American customs. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You don't go on the outing.	You go to the talk, because you might be interested in Native American culture and didn't know it.	You go to the talk excited to learn about a culture to which you have had very limited exposure, and ask lots of pertinent questions.

Example 2

You live in a dorm that is planning a multi-cultural potluck, where residents are encouraged to bring a dish that is typical in the country of their ancestors. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You do not go to the event. (It involves too much time, effort, and money.)	You join in the event, bring a dish, and enjoy the food.	You call your family and get a recipe to cook. You do some research on the significance of the food in that culture, and engage in conversations about the cultures of other countries to expand your and others' understanding.

Now, record the number corresponding to your level on this dimension.

Artistic Appreciation

Definition: Appreciating art and culture, either at an expert level or simply at the level of one who is interested.

Before you make your rating, please read these two examples:

Example 1

Your roommate's class is going on an outing to an art gallery. You are not registered for the class, but think the trip sounds interesting. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You don't go on the outing.	You discuss the outing after your roommate went, and then go to the art gallery on your own time.	You make a special arrangement to go with the class, participate fully in the trip, and have fun at the art gallery.

Example 2

You have set ideas about what music is pleasing to the ear, and friends are pushing you to join them at a concert that they think you would enjoy. The band would be playing music that you prefer to avoid. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You would not go, but would decline as politely as possible.	You go to the concert planning to enjoy only a few pieces.	You go to the concert with an open mind, looking forward to trying something new, especially since it came with a recommendation.

Now, record the number corresponding to your level on this dimension.

Leadership

Definition: Demonstrating skills in a group, such as motivating others, coordinating groups and tasks, serving as a representative for the group, or otherwise performing a managing role in a group.

Before you make your rating, please read these two examples:

Example 1

You are part of a three-person group working on a class project with a quickly approaching deadline. One member of the team is not pulling her weight. The team member avoids assignments, complains about the amount of work that has to be done, and says the project doesn't really matter anyway. While you are all classmates, you seem to be the group leader. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You see if the person could be removed from the group.	You try to get the team member motivated to do her work. If that doesn't help the situation, you just put more effort into the project yourself in order to complete it.	You speak with her in private and offer her encouragement to complete her portion of the project. If the group member still does not pull her own weight, you bring the problem up with the instructor.

Example 2

An intramural sport group of which you are an active member holds an annual event to raise funds for some of its activities. The person selected to lead this event has left school. At this point it is not certain whether any fundraising event will take place, even though this event is important to the group. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You wait for someone else to solve the problem.	You tell everyone to work together and help out since this event benefits the group.	You assume the leadership position, and take charge to ensure that the event goes on.

Now, record the number corresponding to your level on this dimension.

Interpersonal Skills

Definition: Communicating and dealing well with others, whether in informal social situations or more formal school-related situations. Being aware of the social dynamics of a situation and responding appropriately.

Before you make your rating, please read these two examples:

Example 1

Your roommate, usually a tidy person, has recently experienced some personal difficulties. As a result, the roommate has become quite distracted and has left much of the household responsibilities to you. Your have talked to the roommate about your concerns, and empathetically requested that the roommate resume his/her share of the responsibilities as soon as possible. A month passes and you are still doing too much of the roommate's work. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You ask to change rooms.	You do his/her share of the work, and put anything of the roommate's that affects you in the roommate's area of the room.	You talk with the roommate again and explain that you are suffering as a result of the roommate's behavior. You attempt to come up with a mutually acceptable plan of action.

Example 2

You have been standing in line for the restroom for some time after a campus event, and someone cuts into the line ahead of you. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You comment loudly to someone nearby how rude it is that people cut in line.	You tell the person that there is a line.	You calmly and politely inform the person that there is a line and ask that they move to the back.

Now, record the number corresponding to your level on this dimension.

Citizenship

Definition: Being responsible to society and the community, and demonstrating good citizenship. Being actively involved in the events in one's surrounding community, which can be at the neighborhood, town/city, state, national, or college/university level. Activities may include volunteer work for the community, attending city council meetings, and voting.

Before you make your rating, please read these two examples:

Example 1

A friend on your floor is always organizing "social" activities including trips to local bars. Aside from the fact that this person is underage and failing some classes, you realize that the individual is drinking half a dozen or more drinks at least three or four times a week. No one else seems to know or be concerned about the person. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You ask him/her once about this behavior and see where the discussion leads, then leave the individual to his/her own course of action.	You use humor to broach the topic and offer alternatives to the individual's usual "social" activities.	You talk with him/her about easing up on the alcohol, explaining what patterns you have observed, and express concern. You offer to help, and encourage the individual as often as possible.

Example 2

After a local disaster, the Red Cross asked for volunteer blood donors. Because of a medical condition, you cannot donate blood. How do you expect you would react in this situation?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You do nothing – you're off the hook.	You encourage others to donate blood.	You ask the Red Cross if you could help in any other way, and follow through with their suggestions.

Now, record the number corresponding to your level on this dimension.

Health

Definition: Possessing the physical and psychological health required to engage actively in a scholastic environment. This would include participating in healthy behaviors, such as eating properly, exercising regularly, and maintaining healthy personal and academic relations with others, as well as avoiding unhealthy behaviors, such as alcohol/drug abuse, unprotected sex, and ineffective or counterproductive coping behaviors.

Before you make your rating, please read these two examples:

Example 1

You find that you are eating more fattening and greasy food than normal and that you have not been getting sufficient exercise. You have gained 15 pounds, but find it difficult to change your eating and exercising habits. How do you expect you would deal with this situation?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You don't worry about it. You only live once, so eat what you want.	You try to establish a regular exercise routine and focus on eating healthy foods.	You get help from someone with experience in this area, such as a health professional or nutritionist and change your eating habits. You get some friends together to exercise together. There is power in numbers.

Example 2

All of the people who live near you seem to party, drink and use drugs on weekend nights. You like most of these people, but do not want to engage in some of the behavior in which they engage. You have no one else to hang out with. How do you expect you would deal with this situation?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You continue to go along with the group and their activities.	You continue to be friends and hang out with them, but do not engage in their activities.	You join a club and find other friends, and new, healthy behavior to engage in.

Now, record the number corresponding to your level on this dimension.

Career Orientation

Definition: Having a clear sense of career one aspires to enter into, which may happen before entry into college, or at any time while in college. Establishing, prioritizing, and following a set of general and specific career-related goals.

Before you make your rating, please read these two examples:

Example 1

You find that when you started school you planned to major in an area that you now find very uninteresting and your grades are not as good as you would like. You know that you do not want to major in this subject. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You stick with it, as you're already on a career path, even if it isn't a great match.	You explore other options and try to change your major to something you like.	You visit a career counselor to consider possibilities, and meet with your advisor to arrange changing majors to something more interesting and thus useful in the

Example 2

You know what kind of job you want and that getting some experience in that field will be helpful in getting a job after graduation. You have no idea how to get this experience or an internship. How do you expect you would proceed?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You wait until graduation to deal with the idea of what may be helpful in getting a job.	You attend all the campus career fairs and use the available campus resources.	You research jobs on the internet, and take what you have found to a career counselor to help create an action plan. You write a resume, and then contact a group of identified potential employers for internship possibilities.

Now, record the number corresponding to your level on this dimension.

Adaptability

Definition: Adapting successfully to a changing environment (at school or home), dealing well with gradual or sudden and unexpected changes. Being effective in planning one's everyday activities and dealing with novel problems and challenges in life.

Before you make your rating, please read these two examples:

Example 1

Two of your classes have regular projects with assigned deadlines. Some way into the semester, you begin experiencing unforeseen complications with more than one of your projects. Despite your best efforts, it does not look like you will be able to complete all of your assignments adequately and on time. How do you expect that you would proceed?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You simply do what you can to complete both projects, but know that the resulting work will be unsatisfactory.	You talk to others in each class to see if there is any way that you could help each other out and finish the projects on time.	You contact your professors as soon as you realize that you are having difficulty completing the assignments on time. You solicit their advice and try to work out alternative solutions that enable you to complete both projects.

Example 2

You are going through an especially busy period at school. It is the end of the semester, you have papers due, need to prepare for exams, and coworkers at your part-time job are asking that you work more shifts to help out. You find that you are beginning to lose track of details and are feeling overwhelmed. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You take on the shifts to help the coworkers.	You decline the extra shifts and begin to prioritize your school tasks.	You apologize to the coworkers and decline the extra shifts, and tell the coworkers that school is your first priority. You prioritize the school work and create a new schedule to get it all done.

Now, record the number corresponding to your level on this dimension.

Perseverance

Definition: Committing oneself to goals and priorities set, regardless of the difficulties that stand in the way. Goals range from long-term goals (e.g., graduating from college) to short-term goals (e.g., showing up for class every day even when the class isn't interesting).

Before you make your rating, please read these two examples:

Example 1

You are finding your freshman year very difficult. The courses are hard, and you feel your grades are not satisfactory. Material in class seems to be covered very quickly. What do you expect you would do?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You resign yourself to mediocrity.	You study hard, try your best, and try not to worry about it.	You talk to your advisor and teachers. You see if there are study groups or review sessions that you can attend. You are determined to find ways to keep up, and improve, as graduating with good grades is important to you.

Example 2

You moved from a different state to go to school. You have had trouble making friends and your courses are taking much more time than you thought they would. Your family continues to encourage you to do well in school and you know they want you to continue, but you discover that one of your parents is seriously ill. What do you expect you would do in this situation?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You take the semester off and go home.	You try to get most of your work done and make time to go home often.	You stay at school and focus on doing well for your parents. You are determined to persevere and maintain good grades, and go home on breaks.

Now, record the number corresponding to your level on this dimension.

Ethics

Definition: Having a well-developed set of values, and behaving in ways consistent with those values. In everyday life, this probably means being honest, not cheating (on exams or in committed relationships), and having respect for others.

Before you make your rating, please read these two examples:

Example 1

You are taking a particularly difficult exam when you realize how easy it is to see the paper of another student who you know is doing well in the course. When you look at this other student's paper, she notices, and pulls her paper away. You are deeply embarrassed. What action do you expect you would take?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You try not to show that you're embarrassed, and see if you can look at another paper.	You just finish your exam as usual and then leave the room.	You finish the exam without looking at the work of any other students. You seek out the other student after class and apologize to her.

Example 2

You find a very expensive cell phone on the edge of a sidewalk along the street. There is no identifying information on the phone. What do you expect you would do with the phone?

Unsatisfactory (1)	Fulfills Expectations (4)	Exceptional (7)
You keep the phone.	You turn the phone in to the authorities.	After posting flyers and making several unsuccessful attempts to find the owner of the phone, you hand the phone in.

Now, record the number corresponding to your level on this dimension.

APPENDIX I: Organizational Citizenship Behavior

Indicate how frequently you have performed each of the behaviors listed below SINCE ENTERING THIS UNIVERSITY.

1. Gone out of your way to help other students from your school with social problems.
2. Gone out of your way to make new students feel welcome at school.
3. Shown genuine concern and courtesy towards other students.
4. Helped tutor other students struggling with their assignments.
5. Helped students who have been absent from class.
6. Defended your school when others tried to criticize it.
7. Encouraged friends and family to support your school.
8. Shown school spirit (e.g., worn a school t-shirt or put a school decal on your car).
9. Did things to improve your school.
10. Participated in student government or other clubs that try to make your school a better place.

APPENDIX J: Impression Management Scale

Please indicate the extent to which the following statements are true of you.

1. I sometimes tell lies if I have to. (R)
2. I never cover up my mistakes.
3. There have been occasions when I have taken advantage of someone. (R)
4. I never swear.
5. I sometimes try to get even rather than forgive and forget. (R)
6. I always obey laws, even if I'm unlikely to get caught.
7. I have said something bad about a friend behind his or her back. (R)
8. When I hear people talking privately, I avoid listening.
9. I have received too much change from a salesperson without telling him or her. (R)
10. I always declare everything at customs.
11. When I was young I sometimes stole things. (R)
12. I have never dropped litter on the street.
13. I sometimes drive faster than the speed limit. (R)
14. I have done things that I don't tell other people about. (R)
15. I never take things that don't belong to me.
16. I have taken sick-leave from work or school even though I wasn't really sick. (R)
17. I have never damaged a library book or store merchandise without reporting it.
18. I have some pretty awful habits. (R)
19. I don't gossip about other people's business.

Note. R indicates that the item is reverse-scored.

APPENDIX K: Manipulation Check Items

The situations in the questions you responded to above…

1. Provided specific information about college-related responsibilities.
2. Provided easy-to-understand information about college-related requirements.
3. Provided straightforward information about what someone in the situation needs to do to succeed.
4. Made clear exactly what someone in the situation should expect.
5. Provided specific information about which actions to take.
6. Made clear exactly what behavior is expected from someone in the situation.
7. Provided specific information about how others would react to someone in the situation. **(added)**

8. Described sources of information (e.g., other students, parents) that were always consistent with each other.
9. Described responsibilities that were compatible with each other.
10. Described requirements that were highly compatible with each other.
11. Described sources of information that were highly consistent over time.
12. Provided consistent information about what someone in the situation needs to do to succeed. **(added)**
13. Provided consistent information about others' expectations or desires. **(added)**
14. Provided consistent information about what is expected from someone in the situation. **(added)**

15. Described situations in which an individual is prevented from making his/her own decision.
16. Described situations in which constraints prevent an individual from doing things in his/her own way.
17. Described situations in which an individual is prevented from choosing how to do things.
18. Described situations in which other people limit an individual's freedom to make decisions.
19. Described situations in which outside forces limit an individual's freedom to make decisions.
20. Described situations in which other people limit what an individual can do.
21. Described situations in which an individual's personal circumstances (e.g., commitments) limit what he/she can do. **(added)**

22. Described situations in which someone's decisions have very important consequences for other people.
23. Described situations where serious consequences occur when an individual does the wrong thing.
24. Described situations where important outcomes are influenced by someone's actions.
25. Described situations where other people are put at risk when an individual does the wrong thing.
26. Described situations where there are consequences if an individual deviates from what is expected.

27. Described situations in which an individual's behavior has very important consequences for him/her. **(added)**

Note. Clarity (1-7); Consistency (8-14); Constraints (15-21); Consequences (22-27).

APPENDIX L: Demographic Questions

1. What is your name? _____

2. What is your PID? _____

3. What is your age? _____

4. What is your gender?
   a. Male
   b. Female

5. Is your ethnicity Hispanic/Latino?
   a. Yes
   b. No

6. If your ethnicity is not Hispanic/Latino please select your race below. Select more than 1 response if applicable.
   a. American Indian or Alaska Native
   b. Asian
   c. Black or African American
   d. Native Hawaiian or Other Pacific Islander
   e. White
   f. Other

7. If other, please specify. _____

8. What was your High School GPA? _____

9. Did you take the ACT or SAT?
   a. ACT
   b. SAT
   c. Both

10. What was your ACT composite score? (Highest if took test more than once) _____

11. What was your SAT critical reading score? (Highest if took test more than once) _____

12. What was your SAT math score? (Highest if took test more than once) _____

13. What is your class standing?
   a. Freshman
   b. Sophomore
   c. Junior
   d. Senior
   e. Other/non-degree

14. What is your cumulative college GPA? (For scoring key development sample)
    a. less than 2.00
    b. 2.00 to 2.29
    c. 2.30 to 2.59
    d. 2.60 to 2.89
    e. 2.90 to 3.19
    f. 3.20 to 3.49
    g. 3.50 to 3.79
    h. 3.80 to 3.99
    i. 4.00

15. Which of the following best describes your intended or current major?
    a. None, I have not yet decided on a major
    b. Business
    c. Engineering
    d. Fine Arts or Humanities
    e. Social Science (i.e., Psychology, Sociology, Political Science, etc)
    f. Natural Science (i.e., Biology, Chemistry, Physics, etc)
    g. Other

APPENDIX M: SAT to ACT Conversion Chart

Table 61. **SAT to ACT conversion chart**

SAT Composite Score	ACT Composite Score
2400	36
2340-2390	35
2280-2330	34
2220-2270	33
2160-2210	32
2100-2150	31
2040-2090	30
1980-2030	29
1920-1970	28
1860-1910	27
1800-1850	26
1740-1790	25
1680-1730	24
1620-1670	23
1560-1610	22
1500-1550	21
1440-1490	20
1380-1430	19
1320-1370	18
1260-1310	17
1200-1250	16
1140-1190	15
1080-1130	14
1020-1070	13
960-1010	12
900-950	11

REFERENCES

REFERENCES

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471-499.

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.

Bauer, T. N., & Truxillo, D. M. (2006). Applicant reactions to situational judgment tests: Research and related practical issues. In J.A. Weekley & R.E. Ployhart (Eds.), *Situational Judgment Tests: Theory, measurement, and application* (pp. 233–252). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

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

Beaty, J. C., Cleveland, J. N., & Murphy, K. R. (2001). The relation between personality and contextual performance in "strong" versus "weak" situations. *Human Performance, 14*, 125-148.

Billington, A. Q. (2012). Helping undergraduates improve their grades: A comparison of verbal and imagery rehearsal for enhancing training transfer. Unpublished doctoral dissertation, Michigan State University.

Bowers, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review, 80*, 307-336.

Breland, H., Maxey, J., Gernand, R., Cumming, T., & Trapani, C. (2002, March). *Trends in use college admission 2000: A report of a survey of undergraduate admissions policies, practices, and procedures*.

Brooks, M. E. & Highhouse, S. (2006). Can good judgment be measured? In J.A. Weekley & R. E. Ployhart (Eds.), *Situational Judgment Tests: Theory, measurement, and application* (pp. 39 - 56). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Buyse, T. (2011). The use of situational judgment tests in admission to higher education: Validity and coaching effects. Unpublished doctoral dissertation, Ghent University.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Campion, M. C., & Ployhart, R. E. (April, 2012). Situational judgment tests as interactionist psychology. Symposium to be presented at the annual meeting of the Society for Industrial and Organization Psychology, San Diego, CA.

Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*, 143-159.

Chan, D., & Schmitt, N. (2002). Situational judgment and job performance. *Human Performance, 15*, 233−254.

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83-117.

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.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112, 155-159.*

Cooper, O. D., Day, E. A., Connelly, S., Arsenault, M. L., Hardy, J. H., Mracek, D. L. (2013, April). *Development of a construct-oriented situational judgment test of sensemaking skills.* Paper presented at the 28th annual conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Cooper, W. H., & Withey, M. J. (2009). The strong situation hypothesis. *Personality and Social Psychology Review, 13*, 62-72.

Cullen, M. J., Sackett, P. R., & Lievens, F. (2006). Threats to the operational use of situational judgment tests in the college admission process. *International Journal of Selection and Assessment, 14*, 142-155.

Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 191–233). Greenwich, CT: JAI Press.

Dalessio, A.T. (1994). Predicting insurance agent turnover using a video-based situational judgment test. *Journal of Business and Psychology, 9*, 23-32.

Daugherty, T. K., & Lane, E. J. (1999). A longitudinal study of academic and social predictors of college attrition. *Social Behavior and Personality, 27*, 355-362.

Drzakowski, S., Friede, A., Imus, A., Kim, B., Oswald, F. L., Schmitt, N., & Shivpuri, S. (2004). Further development of the *Assessment of Life Experiences Questionnaire (ALEQ) Biodata Measure and Life Events Assessment and Development (LEAD) Situational Judgment Inventory: Investigating differences across gender, ethnic and institutional subgroups.* Research Report submitted to College Board, New York.

Eckes, T. (1995). Features of situations: A two-mode clustering study of situation prototypes. *Personality and Social Psychology Bulletin, 21*, 366–374.

Edwards, B. D., & Arthur, W. (2007). An examination of factors contributing to a reduction in subgroup differences on a constructed-response paper-and-pencil test of scholastic achievement. *Journal of Applied Psychology, 92*, 794-801.

Elicker, J. D., Foust, M. S., O'Malley, A. L., & Levy, P. E. (2008). Employee lateness behavior: The role of lateness climate and individual lateness attitude. *Human Performance, 21*, 427-441.

Endler, N. S., & Magnusson, D. (1976). Toward an interactional psychology of personality. *Psychological Bulletin, 83*, 956-974.

Evans, M. G. (1985). A Monte Carlo study of the effects of correlated method variance in moderated multiple regression analysis. *Organizational Behavior and Human Decision Processes, 36*, 305-323.

Forgas, J. P. (1976). The perception of social episodes: Categorical and dimensional representations in two different social milieus. *Journal of Personality and Social Psychology, 34*, 199–209.

Friede, A., Gillespie, M., Kim, B., Oswald, F., Ramsay, L., & Schmitt, N. (2003). *Summary of the interim and final reports: Biodata and Situational Judgment Inventories as measures of college success*. Research report submitted to College Board, New York.

Fuller, J. B., Hester, K., & Cox, S. S. (2010). Proactive personality and job performance: Exploring job autonomy as a moderator. *Journal of Managerial Issues, 22*, 35-51.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*, 773-794.

Gellatly, I. R., & Irving, P. G. (2001). Personality, autonomy, and contextual performance of managers. *Human Performance, 14*, 231-245.

Gessner, T. L., & Klimoski, R. J. (2006). Making sense of situations. In J.A. Weekley & R. E. Ployhart (Eds.), *Situational Judgment Tests: Theory, measurement, and application* (pp. 13 - 38). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Glaze, R. M., Jarrett, S., Schurig, I., Arthur, W., & Taylor, J. E. (2011). *The efficacy of three SJT response formats.* Paper presented at the 26th annual conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Goldberg, L.R. (2011). Big five factor markers. Retrieved online from http://ipip.ori.org/newBigFive5broadKey.htm

Grand, J. A., Golubovich, J., Ryan, A. M., & Schmitt, N. (2013). The detection and influence of problematic item content in ability tests: An examination of sensitivity review practices for personnel selection test development. *Organizational Behavior and Human Decision Processes, 121*, 158-173.

Grant, A. M., & Rothbard, N. P. (2013, April 29). When in Doubt, Seize the Day? Security Values, Prosocial Values, and Proactivity Under Ambiguity. *Journal of Applied Psychology*. Advance online publication. doi: 10.1037/a0032873

Graziano, W. G., & Eisenberg, N. H. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnston, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–824). San Diego, CA: Academic Press.

Hattrup, K., & Jackson, S. E. (1996). Learning about individual differences by taking situations seriously. In K. R. Murphy (Ed.), Individual differences and behavior in organizations (pp. 507-547). San Francisco: Jossey-Bass.

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology, 57*, 639-683.

Helmreich, R. L., Sawin, L. L., & Carsrud, A. L. (1986). The honeymoon effect in job performance: Temporal increases in the predictive power of achievement motivation. *Journal of Applied Psychology, 71,* 185–188.

Hezlett, S. A., Kuncel, N. R., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., & Camara, W. (2001, April). The predictive validity of the SAT: A meta-analysis. Paper presented in D. Ones & S. Hezlett (Chairs), *Predicting performance: The interface of I-O psychology and educational research.* Symposium presented at the 16th Annual Convention of the Society for Industrial and Organizational Psychology, San Diego, CA.

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: Erlbaum.

Hough, L. M. (1998). Effects of intentional distortion in personality measurement and evaluation of suggested palliatives. *Human Performance, 11,* 209–244.

Humphreys, L. G. (1968). The fleeting nature of the prediction of college academic success. *Journal of Educational Psychology, 59,* 375–380.

Humphreys, L. G., & Taber, T. (1973). Postdiction study of the Graduate Record Examination and eight semesters of college grades. *Journal of Educational Measurement, 10,* 179–184.

Jansen, A., Melchers, K. G., Lievens, F., Kleinmann, M., Brändli, M., Fraefel, L., & König, C. J. (2013). Situation assessment as an ignored factor in the behavioral consistency paradigm underlying the validity of personnel selection procedures. *Journal of Applied Psychology, 98*, 326-341.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*, 386-408.

Jones, C., & Decotiis, T. A. (1986). Video-assisted selection of hospitality employees. *The Cornell H.R.A. Quarterly, 27,* 68-73.

Kobrin, J. L., Patterson, B.F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average.* New York: College Board Research Report No. 2008-5.

Krokos, K. J. (1999). A test of an integrated model of correctional officer job performance. Unpublished master's thesis. North Carolina State University.

Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*, 148–161.

Latham, G. P., Saari, L. M., Pursell, E. D., & Campion, M. A. (1980). The situational interview. *Journal of Applied Psychology, 65*, 422-427.

Li, A., & Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment, 14*, 131-141.

Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology, 90*, 442-452.

Lievens, F., De Corte, W., & Brysse, K. (2003). Applicant perceptions of selection procedures: The role of selection information, belief in tests, and comparative anxiety. *International Journal of Selection and Assessment, 11*, 67-77.

Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology, 93*, 268-279.

Lievens, F., Ones, D. S., & Dilchert, S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology, 94,* 1514-1535.

Lievens, F., & Sackett, P. R. (2011, October 3). The Validity of Interpersonal Skills Assessment Via Situational Judgment Tests for Predicting Academic Success and Job Performance. *Journal of Applied Psychology*. Advance online publication. doi:10.1037/a0025741

Lin, P.-C., & Humphreys, L. G. (1977). Predictions of academic performance in graduate and professional school. *Applied Psychological Measurement, 1,* 249–257.

MacKenzie, W. I., Ployhart, R. E., Weekley, J. A., & Ehlers, C. (2010). Contextual effects on SJT responses: An examination of construct validity and mean differences across applicant and incumbent contexts. *Human Performance, 23*, 1-21.

Magnusson, D. (1971). An analysis of situational dimensions. *Perceptual and Motor Skills, 32*, 851–867.

Mael, F. A., Connerley, M., & Morath, R. A. (1996). None of your business: Parameters of biodata invasiveness. *Personnel Psychology, 49*, 613–650.

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

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

McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology, 36*, 979-1016.

Meyer, R. D., Dalal, R. S., & Bonaccio, S. (2009). A meta-analytic investigation into the moderating effects of situational strength on the conscientiousness–performance relationship. *Journal of Organizational Behavior, 30*, 1077-1102.

Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management, 36*, 121-140.

Meyer, R. D., Dalal, R. S., José, I. J., Hermida, R., Chen, T. R., Vega, R. P., Brooks, C. K., Khare, V. P. (in press). Measuring job-related situational strength and assessing its interactive effects with personality on voluntary work behavior. *Journal of Management, X*, xx-xx.

Mischel, W. 1977. The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 333-352). Hillsdale, NJ: Lawrence Erlbaum.

Monson, T. C., Hesley, J. W., & Chernick, L. (1982). Specifying when personality traits can and cannot predict behavior: An alternative to abandoning the attempt to predict single-act criteria. *Journal of Personality and Social Psychology, 43*, 285-399.

Mumford, T. V., Van Iddekinge, C. H., Morgeson, F. P., & Campion, M. A. (2008). The team role test: Development and validation of a team role knowledge situational judgment test. *Journal of Applied Psychology, 93*, 250-267.

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

Murphy, K. R. (1989). Is the relationship between cognitive ability and performance stable over time? *Human Performance, 2*, 183–200.

Organ, D. W. (1997). Organizational citizenship behavior: It's construct cleanup time. *Human Performance, 10*, 85–97.

Oswald, F.L., Schmitt, N., Kim, B.H., Ramsay, L.J. & Gillespie, M.A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*, 187–208.

Paulhus, D. L. (1991). Measurement and control of response bias. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Paulhus, D. L. (1994). Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6. Unpublished manuscript. University of British Columbia, Vancouver, Canada.

Pereira, G., & Harvey, V. S. (1999). Situational judgment tests: Do they measure ability, personality, or both? In J. P. Clevenger (chair), *The Construct Validity of the Situational Judgment Inventory.* Symposium presented at the 14th annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Pervin, L. A. (1985). Personality: Current controversies, issues, and directions. *Annual Review of Psychology, 36,* 83-114.

Ployhart, R.E. (1999). An interactionist approach to assessing personality in work contexts: Construct validation of a predictor of customer service performance. Unpublished doctoral dissertation, Michigan State University.

Ployhart, R. E. (2006). The predictor response process model. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests*. Mahwah, NJ: Lawrence Erlbaum & Associates.

Ployhart, R. E., & Ehrhart, M. G. (2002). Modeling the practical effects of applicant reactions: Subgroup differences in test-taking motivation, test performance, and selection rates. *International Journal of Selection and Assessment, 10,* 258-270.

Ployhart, R. E., & Ehrhart, M. G. (2003). Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection & Assessment,* 11, 1–16.

Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61,* 153-172.

Reynolds, D. H., Sydell, E. J., Scott, D. R., & Winter, J. L. (2000, April). *Factors affecting situational judgment test characteristics.* Paper presented at the 15th annual conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Reynolds, D. H., Winter, J. L., & Scott, D. R. (1999, May). *Development, validation, and translation of a professional-level situational judgment inventory.* Paper presented at the 14th annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.

Ramsay, L. J., Schmitt, N., Oswald, F. L., Kim, B. H., & Gillespie, M. A. (2006). The impact of situational context variables on responses to biodata and situational judgment inventory items. *Psychology Science, 48,* 268-287.

Richman-Hirsch, W. L., Olson-Buchanan, J. B., & Drasgow, F. (2000). Examining the impact of administration medium on examinee perceptions and attitudes. *Journal of Applied Psychology, 85,* 880-887.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*, 261-288.

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*, 103-139.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54*, 297-330.

Ryan, A. M. (2001). Explaining the Black-White test score gap: The role of test perceptions. *Human Performance, 14*, 45-75.

Ryan, A. M., & Huth, M. (2008). Not much more than platitudes? A critical look at the utility of applicant reactions research. *Human Resource Management Review, 18*, 119-132.

Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology, 83*, 298-307.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology, 50*, 707-721.

Sackett, P.R., Kuncel, N.R., Arneson, J.J., Cooper, S.R., & Waters, S.D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin, 135*, 1-22.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 56*, 302–318.

Salter, N. P. (2009). A multitrait-multimethod approach to isolating situational judgment from situational judgment tests. Unpublished doctoral dissertation, Bowling Green State University.

Santelices, M. V., Ugarte, J. J., Flotts, P., Radovic, D., & Kyllonen, P. (2011). *Measurement of new attributes for Chile's admissions system to higher education* (ETS Research Report No. RR-11-18). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RR-11-18.pdf

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

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1*, 199-223.

Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, S. L. (1995). Frame-of-reference effects on personality scores and criterion-related validity. *Journal of Applied Psychology, 80,* 607–620.

Schmitt, N., Billington, A., Keeney, J., Reeder, M., Pleskac, T., Sinha, R., & Zorzie, M. (2010). *Development and validation of measures of noncognitive college student potential.* Research report submitted to College Board, New York.

Schmitt, N., Billington, A. Q., Keeney, J., Oswald, F. L., Pleskac, T., Sinha, R., & Zorzie, M. (2009). Prediction of four-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, *94*, 1479-1498.

Schmitt, N., & Chan, D. (2006). Situational judgment tests: Method or construct? In J.A. Weekley & R. E. Ployhart (Eds.), *Situational Judgment Tests: Theory, measurement, and application* (pp. 135 - 155). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Schmitt, N., Clause, C. S., & Pulakos, E. D. (1996). Subgroup differences associated with different measures of some common job-relevant constructs. International *Review of Industrial and Organizational Psychology, 11*, 115-139.

Schmitt, N., & Golubovich, J. (2013). Biographical information. In K. F. Geisinger (Ed.), *APA Handbook of Testing and Assessment in Psychology* (pp. 437-455). Washington, DC, US: American Psychological Association.

Schmitt, N., Oswald, F. L., Kim, B. H., Imus, A., Merritt, S., Friede, A., & Shivpuri, S. (2007). The use of background and ability profiles to predict college student outcomes. *Journal of Applied Psychology, 92*, 165-179.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology, 82*, 719-730.

Schneider, R. J.,& Hough, L. M. (1995). Personality and industrial/organizational psychology. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 75–129). Chichester, England: Wiley.

Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*, 325-343.

Shultz, M .M., & Zedeck, S. (2011). Predicting lawyer effectiveness: Broadening the basis for law school admission decisions. *Law & Social Inquiry, 36*, 620-661.

Siemsen, E., Roth, A., & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods, 13,* 456-476.

Sinha, R., Oswald, F., Imus, A., & Schmitt, N. (2011). Criterion-focused approach to reducing adverse impact in college admissions. *Applied Measurement in Education, 24, 137-161.*

Smith, B. M., & Nelson, L. D. (1975). Personality correlates of helping behavior. *Psychological Reports,* 37, 307–310.

Smithikrai, C. (2008). Moderating effect of situational strength on the relationship between personality traits and counterproductive work behaviour. *Asian Journal of Social Psychology, 11*, 253−263.

Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology, 81*, 134-141.

Stemler, S. E. (2012). What should university admissions tests predict? *Educational Psychologist, 47*, 5-17.

Stemler, S. E. & Sternberg, R. J. (2006). Using situational judgment tests to measure practical intelligence. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational Judgment Tests: Theory, measurement, and application* (pp. 107 - 134). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WISC: A model for college and university admissions. *Educational Psychologist, 47*, 30-41.

Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*, 207-228.

Taber, T. D., & Hackman, J. D. (1976). Dimensions of undergraduate college performance. *Journal of Applied Psychology, 61*, 546-558.

Ten Berge, M. A., & De Raad, B. (2002). The structure of situations from a personality perspective. *European Journal of Personality, 16*, 81–102.

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88,* 500–517.

Titus, M. A. (2004). An examination of the influence of institutional context on student persistence at 4-year colleges and universities: A multilevel approach. *Research in Higher Education, 45*, 673-699.

Trippe, D. M. (2002). An evaluation of the construct validity of situational judgment tests. Unpublished master's thesis, Virginia Polytechnic Institute and State University.

Van Iddekinge, C. H., McFarland, L. A., & Raymark, P. H. (2007). Antecedents of impression management use and effectiveness in a structured interview? *Journal of Management, 33*, 752-773.

Wagerman, S. A. & Funder, D. C. (2006, January). *The Riverside Situational Q-Sort: A Tool for Understanding the Psychological Properties of Situations*. Poster presented at the annual meeting of the Society for Personality and Social Psychology, Palm Springs, CA.

Wagerman, S.A., & Funder, D.C. (2009). Situations. In P.J. Corr and G. Matthews (Eds.), *Cambridge Handbook of Personality* (pp. 27-42). Cambridge: Cambridge University Press.

Weekley, J. A., & Jones, C. (1997). Video-based situational testing. *Personnel Psychology, 50,* 25 – 49.

Weekley, J. A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology, 52,* 679 – 700.

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

Wish, M., Deutsch, M., & Kaplan, S. J. (1976). Perceived dimensions of interpersonal relations. *Journal of Personality and Social Psychology, 33*, 409-420.

Withey, M. J., Gellatly, I. R., & Annett, M. (2005). The moderating effect of situation strength on the relationship between personality and provision of effort. *Journal of Applied Social Psychology, 35*, 1587-1608.

Zorzie, M. (2012). Individual adaptability: Testing a model of its development and outcomes. Unpublished doctoral dissertation, Michigan State University.

Zwick, R. (2007a). *College admissions testing*. Alexandria, VA: National Association for College Admission Counseling.

Zwick, R. (2007b). College admissions in twenty-first-century America: The role of grades, tests, and games of chance. *Harvard Educational Review*, *77*, 419–429.