



Minnesota State University, Mankato
Cornerstone: A Collection of Scholarly
and Creative Works for Minnesota
State University, Mankato

All Graduate Theses, Dissertations, and Other
Capstone Projects

Graduate Theses, Dissertations, and Other
Capstone Projects

2012

Perceptions of a Text-Based SJT versus an Animated SJT

Amanda Helen Halabi
Minnesota State University - Mankato

Follow this and additional works at: <https://cornerstone.lib.mnsu.edu/etds>



Part of the [Industrial and Organizational Psychology Commons](#)

Recommended Citation

Halabi, A. H. (2013). Perceptions of a Text-Based SJT versus an Animated SJT [Master's thesis, Minnesota State University, Mankato]. Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. <https://cornerstone.lib.mnsu.edu/etds/156/>

This Thesis is brought to you for free and open access by the Graduate Theses, Dissertations, and Other Capstone Projects at Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato. It has been accepted for inclusion in All Graduate Theses, Dissertations, and Other Capstone Projects by an authorized administrator of Cornerstone: A Collection of Scholarly and Creative Works for Minnesota State University, Mankato.

Perceptions of a Text-Based SJT versus an Animated SJT

By

Amanda Halabi

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

In

Industrial/ Organizational Psychology

Minnesota State University, Mankato

Mankato, Minnesota

December, 2012

Perceptions of a Text-Based SJT versus an Animated SJT

Amanda Halabi

This thesis has been examined and approved by the following members of the thesis committee.

Kristie Campana, Ph.D., Advisor

Lisa Perez, Ph.D.

Paul Schumann, Ph.D.

Abstract

Halabi, Amanda H., M.A. Minnesota State University, Mankato 2012

Research has indicated that Multimedia SJTs hold higher face validity than text-based SJTs while also reducing the adverse impact that has been experienced with the use of cognitive ability tests. (Chan & Schmitt, 1997; Motowidlo et al., 1990; Weekley & Jones, 1997, 1999). The animated SJT, a new and emerging form of testing, holds many similarities to the MMSJT. However, no research has been conducted to support this claim. The present study examined the perceptions of an animated SJT versus a text based SJT to provide additional information to the field of selection tests. The research was conducted to provide insight into applicant reactions as well as information about what drives the preference for each type of Situational Judgment Test. Overall, results indicated that the animated SJT was perceived more positively than the text based SJT, with significantly more participants preferring the animated to the text based SJT.

TABLE OF CONTENTS

I.	INTRODUCTION	4
II.	METHOD	15
III.	RESULTS	20
IV.	DISCUSSION.....	26
V.	REFERENCES.....	32
VI.	APPENDIX.....	35

LIST OF TABLES

Table 1. Sample Demographics.....	14
Table 2. Descriptive Statistics for the SJT Measures.....	19
Table 3. Intercorrelations between all Measures.....	20

Perceptions of a Text-Based SJT versus and Animated SJT

Situational Judgment Tests

A situational test is a form of evaluation in which an applicant is presented with a scenario and is asked to react to that specific situation. By this definition, “assessment centers, work samples, situational interviews, and SJTs are all examples of situational tests” (Oostrom, Born, Serlie, & van, 2011).

In this same vein, a situational judgment test, or an SJT, is “designed to measure judgment in work settings” (McDaniel, Morgenson, Finnegan, Campion, & Braverman, 2001). In its typical form, an SJT presents an applicant with a specific, future-oriented scenario, also known as the question stem. The applicant is then asked to choose the most appropriate response to the scenario out of a list of possible outcomes, or response options. Consider the following example of an SJT item (this is a sample item from the retail SJT used in this study):

Scenario/ Question Stem

A customer was just in the store yesterday and purchased a phone and a 12-month service contract. If a customer is not satisfied, he is able to cancel the contract within a week and his money would be refunded. Today he has returned with the phone. The customer walks into the store and sets his phone on the counter.

Dialogue

Customer: “Hi – I want to return this phone. The display doesn’t seem to be working – I get all these strange symbols on it.”

Salesperson: “OK. Let’s take a look and see what might be going on here.”

Customer: (shakes his head) “No. I just want to return the phone and cancel my contract – if it’s already having issues now, I don’t want to invest time trying to fix it.”

Question

Choose the Most Effective and Least Effective course of action from the options below.

Response Options

- State that it is likely a problem that is specific to that brand of phone and ask if he would be interested in trying a different brand.
- Agree to return the phone for the customer, but while you process the return, engage in small talk to build rapport and potentially get him to look at other phones.
- Return the phone for the customer and apologize for any inconvenience.
- Ask him whether he would consider taking home a new phone (same model) and see if he continues to have the same issue.

SJTs tend to operate under the assumptions of the behavioral consistency theory.

This theory postulates that the current behavior of an individual can be used to predict their future behaviors. An SJT typically asks an applicant to respond to a job situation, as with the sample above. The response of the applicant is therefore an indication of their behavior, and can be used to predict future job performance (Motowidlo, Dunnette, & Carter, 1990).

Note that SJTs can have various response formats. The one above asks the respondent to choose the best and worst answer out of a list of possible outcomes. Other SJTs ask the respondent to rate their level of agreement with each statement. Still other forms may ask the respondent to rank order each answer from most effective outcome to least effective outcome. Typically, SJTs are said to have two different types of response formats, those that measure the knowledge of the applicant and those that measure behavioral tendencies. Knowledge response formats ask respondents to choose the best possible response from the response outcomes. Knowledge response formats may also

involve rating the effectiveness of each response outcome. Conversely, behavioral tendency response formats typically ask the respondent to rate what they would be most likely to do in a certain situation. They may also ask the respondent to rate the likelihood that they may perform any given response outcome. The SJT used in this study will mirror the example provided above, therefore utilizing the knowledge response format. Although research has indicated that the behavioral tendency response format correlates more highly with the Big 5 personality traits, knowledge response formats typically have a higher correlation with cognitive ability, making this response format very useful for selection (Diggman, 1990; McDaniel, Hartman, Whetzel, & Grubb, 2007).

SJTs are capable of assessing multiple constructs within one session, as they are used as a measurement method rather than a specific construct (Chan & Schmitt, 1997; McDaniel & Nguyen, 2001; McDaniel et al, 2001; Weekley & Jones, 1999). This allows SJTs to be used in the way that an interview or an assessment center would be used to screen applicants (McDaniel et al, 2001).

History of SJTs

As mentioned above, one of the main focuses of an SJT is employee selection. Employee selection is a vital part of every organization. Ensuring that an applicant is a “good fit” allows companies to save time and money in the long run. An employee that will perform well on the job will provide a higher return on investment for an organization, thereby reducing their long term costs. In order to ensure that a company is hiring the applicants who will provide the highest return on investment, Industrial-Organizational psychologists began constructing tests that would measure various

qualities of an applicant, such as cognitive ability, personality, and integrity. Another selection test that was developed to assess applicants was the Situational Judgment Test (SJT). This selection tool was designed to assess how an applicant would respond to various situations they encounter at work (Weekley & Ployhart, 2006).

Although the first SJT can be dated back to the early 1900s, the first widely used SJT emerged in 1926 (McDaniel et al., 2001). The George Washington Social Intelligence Test contained a piece titled Judgment in Social Situations which required a “keen judgment, and a deep appreciation of human motives, to answer correctly” (Moss, 1926, p. 26). As time progressed, the army began using a form of SJTs to assess soldiers during World War II, where a soldier was judged on their ability to use “common sense, experience, and general knowledge” (Whetzel & McDaniel, 2009). Soon, SJTs were created and adapted to the workplace, used primarily to assess the supervisory potential of applicants using the Practical Judgment Test and the Supervisory Practices Test. By the 1950s and 1960s, SJTs were even used to measure the success of managers in the workplace (Whetzel & McDaniel, 2009).

Recently, Situational Judgment Tests have found their way back into the workplace. Research has indicated that SJTs are an effective selection tool used both in the United States and in a variety of European countries (McDaniel et al., 2001). Reasons for this renewed interest in SJTs come in a variety of forms.

According to Whetzel & McDaniel (2009), four main reasons exist for the popularity SJTs have acquired in the past 20 years. First, an article published by Motowidlo et al. (1990) drew attention to the use of SJTs in the workplace as selection

tools. Second, multiple meta-analyses have documented the high validity of SJTs as predictors of future job performance.

The third reason involves the nature of other predictors used today to predict job performance. As demonstrated throughout the literature, cognitive ability tends to be the highest predictor of future job performance with a validity coefficient of approximately .32 (Schmidt, 2002). Essentially, general cognitive ability is the ability to learn (Schmidt, 2002). Unfortunately, numerous studies have also demonstrated that measures of cognitive ability result in significant subgroup differences and tend to have low face-validity (Clearly, 1968). Due to these negative results of using cognitive ability to predict performance, many organizations hesitate to use cognitive ability tests as their primary selection procedure. Research on SJTs, however, indicated that SJTs demonstrate much lower levels of race-based adverse impact, making them accessible in the workplace and more resistant to potential lawsuits than tests of cognitive ability (Chan & Schmitt, 1997; Motowidlo et al., 1990; Weekley & Jones, 1997, 1999). Additionally, SJTs generally have a correlation of .31 with tests of cognitive ability, indicating that they measure the cognitive ability of applicants to some degree, providing organizations some of the benefits of a cognitive ability test without the potential downfalls (McDaniel et al., 2001; Whetzel & McDaniel, 2009).

The focus of the current study is on the fourth and final reason for the renewed interest in SJTs. Literature has shown that SJTs tend to have high face validity, meaning that they are rated as job relevant by the applicants who are taking them. This study

examines the perceptions of SJTs presented in alternate formats, namely a text based SJT versus a new approach, the animated SJT.

The Advantages of Using SJTs for Selection

As mentioned above, SJTs have demonstrated moderate criterion-related validity in predicting job performance (Funke & Schuler, 1997; McDaniel et al., 2001). In a meta-analysis conducted by McDaniel et al. (2001), 102 correlation coefficients were gathered based on data from 10,640 individuals. Results showed a moderate level of predictive validity at .34. The results of this study were considered to generalize to all types of work samples, as studies involving students were excluded from the validity calculations. The relationship between a SJT and performance, however, is moderated by whether the SJT was created using a thorough job analysis versus a group of experts. As is consistent with data collected on personality tests and interviews, the predictive capabilities of SJTs rise if they are created using a job analysis (McDaniel et al, 2001; Tett, Jackson, & Rothstein, 1991; Wiesner & Cronshaw, 1988).

Predictive validity is critical to SJTs. This is because the situations presented in SJTs can be so job specific that the applicant must have previously performed the job in order to choose the best outcome. Research has demonstrated that SJTs have moderate predictive validity (Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001). A recent meta-analysis conducted on approximately 118 coefficients from research using job incumbents found the overall validity of SJTs to be approximately .26 (McDaniel et al, 2007).

Additionally, scores on SJTs tend to have a positive correlation with scores on general mental ability tests, emotional stability, conscientiousness, and agreeableness (McDaniel & Nguyen, 2001). It has been found, however, that SJTs have higher incremental validity when paired with tests of cognitive ability than tests of general mental ability administered alone (Clevenger, 2001; McDaniel et al, 2001). SJTs also have high correlations with age and length of experience (McDaniel & Nguyen, 2001; Smith & McDaniel, 1998). Specifically, the older and more experienced an applicant is, the better they perform on the SJT. Based on these results it can be inferred that SJTs measure the knowledge and skills of an applicant. These knowledge and skills were said to have been gathered from relevant life and work experiences (Smith & McDaniel, 1998). These results lead to the first hypothesis in this study:

Hypothesis 1: Participants with experience in the retail industry will perform better than those who have no prior retail oriented work experience on the animated SJT.

Since the discovery of this effective measure of job performance, different forms of the Situational Judgment Tests have emerged. These types of SJTs are presented and further explored below.

Types of Situational Judgment Tests

In the past, SJTs have typically been administered in a paper and pencil format (McDaniel, et al., 2001; Weekley & Jones, 1999). This form of the SJT is considered low fidelity. Fidelity refers to how similar the situation being presented will be to what is

truly experienced while on the job (Funke & Schuler, 1997; Livens, Peeters, & Schollaert, 2008).

In an effort to create a test that had high fidelity, psychologists have designed the video based SJT, also known as the Multimedia Situational Judgment Test. In their typical form, the MMSJTs provide an interactive view of a scenario, allowing the applicant to watch a scene acted out by human beings and then responding with the appropriate action to take based on the scenario. MMSJTs are considered high fidelity due to the likeness of the situation to what could be experienced on the job. Using video clips allows the applicant to see individuals play a scene out rather than having to read and comprehend what is occurring subjectively. This allows the MMSJT to more highly resemble the criterion, resulting in a higher criterion-related validity (Lievens & Sackett, 2006; Motowidlo et al., 1990). Specifically, Lievens & Sackett (2006) conducted a study to test the predictive validity of high-fidelity SJTs. Results indicated that the MMSJTs had a greater ability to predict future performance on interpersonally oriented courses than their identical low-fidelity text-based counterparts. According to Olson-Buchanan and Drasgow (year), these results may be due to the point that:

[Video-based SJT assesses] see and hear people interacting, perceive or, importantly, fail to perceive their emotions and stress, and confront dilemmas about one's choice of action or inaction With this format, we may be able to better understand how the assessee will interpret verbal and nonverbal

behaviors of others in the workplace and choose to respond.

(p. 253)

It should be noted that not all researchers are in agreement about the high-fidelity relationship between the criterion and the MMSJT. Some have argued that the video-based SJT introduces contextual cues as well as extraneous variables by which the applicant may be affected (Weekley & Jones, 1997). This viewpoint argues that by having actors read the SJT prompts, unnecessary error is introduced into the measurement method. For example, in an MMSJT applicants will be exposed to contextual cues as well as body language cues that are not present in a text-based SJT. This may cause the applicant to choose a particular response based on these cues rather than the scenario that is being acted out. Advocates of this viewpoint argue that these extraneous variables add unnecessary error into the SJT, potentially resulting in a lower predictive validity of future job performance. To date, however, there has been no empirical evidence that supports this claim.

In addition to higher criterion-related validity, the video-based SJT resulted in a lower correlation with cognitive ability than the text-based SJT, potentially leading to less adverse impact (Chan & Schmitt, 1997; Livens & Sackett, 2006). In a meta-analysis conducted by McDaniel et al (2001), results indicated that there was much variability in results measuring the correlation between cognitive ability and SJTs. According to Livens and Sackett (2006), this variability could be a result of the differences between a video-based and a text-based SJT. Specifically, “the greater reading component inherent in a written administration might lead to a stronger correlation between cognitive ability

and the (text) SJT” (Lievens & Sackett, 2006, p. 1182). Chan and Schmitt (1997) held all variables (including content) constant except for the test method. Results indicated that paper-and-pencil tests result in high-subgroup differences because they require the ability to read and comprehend material.

Applicant reactions to SJTs

In general, paper-and-pencil measures of performance tend to have lower applicant reactions than tests involving simulations (Macan, Avedon, Paese, & Smith, 1994; Schmidt, Greenthal, Hunter, Berner, & Seaton, 1977; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993). These positive attitudes towards testing simulations such as work samples and assessment centers stem from the ability of the applicant to see the job-relatedness of the task they are completing.

Research has demonstrated that both text and multimedia based SJTs are perceived as favorable by applicants, with the multimedia SJTs resulting in an increase in endorsement from applicants. Additionally, research shows that multimedia SJTs are more enjoyable and are perceived as more modern than text-based tests (Olson-Buchanan & Drasgow, 2006).

Another area where the video-based SJT has shown significant benefits over the text-based SJT is in face validity. High face validity indicates that participants perceive the test they are taking as being job-related and relevant to the position for which they are applying. Studies have consistently demonstrated higher face validity for MMSJTs (Chan & Schmitt, 1997; Olson-Buchanan & Drasgow, 2006). Specifically, Olson-

Buchanan and Drasgow (2006) conducted a study that compared a multimedia conflict resolution skills SJT to the same test in text format. Results showed that the students in the study reacted more favorably towards the MMSJT.

The results presented thus far were derived from MMSJTs, which typically have human subjects act out the scenario and its potential solutions. A new form of the MMSJT is the animated approach. In this type of SJT, a scenario is presented to the applicant by animated characters that act out the plot. Answers are then presented to the applicant, who is asked to choose the most appropriate response from the list of responses. An area in which the literature is lacking is animated SJTs. Although similar to multimedia-based SJTs, these tests have not been studied thoroughly enough to draw any conclusions. The results of previous studies as well as the lack of research performed on animated SJTs form the basis for hypotheses 2, 3, and 4 of this study:

Hypothesis 2: The animated SJT will be preferred to the text based SJT.

Hypothesis 3: The animated SJT will be perceived as conveying the best impression of an organization to future job candidates.

Hypothesis 4: The animated SJT will be deemed more “modern,” more “fun,” and more “engaging” than the text-based SJT.

Additionally, an area in which research is lacking is whether the preference for a text-based SJT over a multimedia SJT is more highly correlated with scores on a cognitive ability test. Chan and Schmitt (1997) suggest that positive applicant reactions

may lead to increased test performance, such that “differential subgroup test reactions may explain some of the job performance and behavior variance or test performance associated with race.” Research on cognitive ability tests and work samples has shown no sub-group differences in applicant reactions (Schmidt et al, 1997). Chan and Schmitt (1997), however, conducted a study that was specific to SJTs. Results indicated that white participants tend to rate paper and pencil SJTs more favorably than black participants. These results support the notion that applicants who prefer paper-and-pencil SJTs may perform better on a test of cognitive ability than those who prefer the MMSJT. These results lead to the final exploratory analysis that will be tested in this study:

Exploratory analysis: Those who prefer the text-based SJT will attain higher cognitive ability scores than those who prefer the animated SJT.

Proposed Research

This research addresses the lack of literature on Animated Situational Judgment Tests in order to provide additional information to the field. Additionally, perceptions of text-based SJTs versus animated SJTs will be examined. This research can be beneficial to organizations looking to implement an animated SJT into their selection procedures, as it will provide some insight into applicant reactions as well as information about what drives the preference of each.

Methods

Participants

All data collection took place at a small, Midwestern university. The sample consisted of 148 undergraduate students enrolled in an Introduction to Psychology course at the university. Each participant who completed the battery of online tests was awarded extra credit in a psychology course for his or her participation. Table 1 outlines the demographic properties of the sample utilized. Note that participants were not required to respond to the demographic questions.

Table 1

Sample Demographics

Variable	N
Sex	
Men	32*
Women	54*
Race	
Hispanic or Latino/a	2
White	76
Black or African-American	5
Asian	1
Hawaiian or Pacific Islander	0
Native American or Alaskan	0
Native	0
Two or more races	0
Other	2
Age	
0-17	0
18-20	67
21-29	17
30-39	0
40-49	0
50-59	1
60 and older	0

* Participants do not add up to 148 due to the use of less data to run Hypotheses 1-4.

Measures

A small consulting firm located in Minneapolis provided all of the following measures. The firm uses SJTs in their selection procedures, and allowed the current study to use all their measures free of charge. The following measures were utilized:

Situational Judgment Test. The first measure used in this study was the Retail Sales and Services Simulation US – ENG. This measure was an animated SJT comprised of eight items. All of the items within this measure were based in a retail store environment, with scenarios specific to issues encountered in this type of work setting. The test took approximately 15 minutes to complete. A text-based SJT that is a parallel form to the animated SJT was used in the study. The text-based SJT was identical to the animated SJT, containing eight items that took approximately 15 minutes to complete. A sample item from the text-based SJT is provided in Appendix A.

Applicant Reactions. A survey was administered after both the animated and text based SJT that measured applicant reactions to the various forms. The reactions measure was comprised of 32 items and took about four minutes to complete. Applicants were provided a list of adjectives and were asked to select all of the adjectives that applied to the SJT they had completed most recently. The measure was presented in the form of a checklist. Examples of adjectives used were “modern,” “amateur,” and “complex.”

Once participants completed both the animated SJT and the text-based SJT, applicants were asked to complete a survey comparing the text-based SJT to the animated

SJT. This survey was comprised of four items and took approximately two minutes to complete. An example item was “Which test would provide the most useful information about a candidate?” Respondents were then asked to indicate their preference between the written version of the test and the animated version.

Cognitive Ability. Participant’s cognitive ability was assessed using the G Short Dynamic Contrast Test (GSDCT) and the Following Instructions – US R2 test (FIUSR2). Both are measures of cognitive ability. Because the GSDCT is a computer-adaptive test, the number of items each participant answered varied depending upon their performance. In general, most participants took approximately 10 minutes to complete the test. The test questions from the GSDCT asked respondents to rationalize through various questions such as word problems, identifying patterns, and other problem solving related questions. The second cognitive ability test, FIUSR2, consisted of 40 items and took eight minutes to complete. Participants were provided with a set of instructions and were asked to solve problems based on the information provided. This test was timed and shut down after eight minutes. Higher scores on this test indicated higher levels of cognitive ability. Samples of the cognitive ability questions are provided in Appendix B.

Procedures

The surveys were developed and administered using online survey software. The administration of this study occurred in two parts. After a participant signed up for the study, the following steps were taken:

Participants were sent an e-mail with instructions and links for the surveys. Participants first accessed a testing platform where they were asked to create a username and password as dictated by the instructions. After logging in, participants were presented with a consent form that they signed before continuing the study. Participants then completed the battery of questionnaires and surveys. Each testing session began with either the animated or text based SJT. To counter-balance the order in which the situational judgment tests were distributed, researchers alternated the form sent to participants for each group of 15. Specifically, the first 15 participants received Form A on the testing platform and therefore took the text-based SJT first. The next 15 received Form B on the testing platform and therefore took the animated SJT first and so forth. Following the first SJT, participants went through the series of reactions measures. They then participated in the biographical questionnaires, followed by the SJT type they had not yet completed (i.e. Form A received the animated SJT, Form B received the text-based SJT). The SJT was followed by the reactions measures questionnaire, wrapping up with the four text versus animated SJT comparison questions and a demographics questionnaire.

Ten days after completing the battery of tests for part 1 of the study, the participant was sent an e-mail reminding them about part two of the study. Exactly two weeks after the completion of part one of the study, the link to part two was given to each participant. The participant had one week to complete part two of the study. A reminder e-mail was sent to those who did not immediately complete part two two, four, and six days after the part two link was initially sent. All part two testing sessions were identical,

beginning with the animated SJT, followed by the Following Instructions cognitive ability measure. Participants then took the biographical data questionnaire and ended with the GSDCT measure. Upon completing the battery of tests at time two, all participants were awarded extra credit.

Time two allowed us to gather cognitive data without overloading the participant at time one with multiple mentally taxing tasks. Participants were only eligible to complete part two of the study if they had previously completed part one of the study. Only data from individuals who completed both parts of the study was used.

Results

Cleaning of the data was performed on each of two data sets. First, the sample was split into two files, one with the responses to all reaction measures collected from SurveyMonkey, and one with the responses from the testing platform used to collect the SJT responses, biographical data, and cognitive measure responses. Only cases that had completed both Time 1 and Time 2 were used for the analyses.

Specific steps were taken to clean the reactions measures data. First, if identical duplicate cases were found, all but one of the cases was deleted. It was assumed that the duplicate cases were due to the respondent clicking the “Submit” button multiple times, and therefore submitting the same answers from the same session more than once. Therefore, the very first answer set submitted was the one retained for this study. Additionally, if multiple cases were found for one participant but the cases were not identical, all the cases for that individual were deleted. Additionally, deletions were made for individuals who randomly filled in answers (e.g. letters instead of numbers, invalid ID

numbers, no answers on the entire test) as well as any test runs performed by the researchers. This resulted in a sample size of 101 participants for the reactions measures portion of the survey, as opposed to the 148 total sample size.

In order to clean the SJT data, first the data was split into Time 1 and Time 2 sessions. Any multiple session ID numbers were deleted along with any individual without a Time 1 completion date. This resulted in 148 participants for the SJT portion of the survey.

After the data cleaning was complete, demographic variables were examined for any trends in the data. The average age of participants was 18, with approximately 90% being of Caucasian-American descent. Almost half of the participants had previous retail experience, with the average number of years working in the industry being around two years. Additionally, over 75% of participants stated that they felt very comfortable using the Internet. Table 2 has been provided below outlining the descriptive statistics of the data collected from the SJTs.

Descriptive Statistics for the SJT Measures

Test Type	N	Mean	Std. Deviation
Retail MMSJT Time 1	148	5.77	1.75
Retail Text-SJT Time 1	148	6.03	1.00
Following Instructions Cog. Ability Test	148	12.21	11.98
G-Short Dynamic Cog. Ability Test	148	1.71	1.14
Retail MMSJT Time 2	148	4.42	2.77

It should also be noted that the intercorrelations of the measures was examined.

Table 3 has been provided below with a list of the results.

Table 3

Intercorrelations between all s2p measures

		T1 Biographical Data A	T1 Biographical Data B	T1 Retail MMSJT	T1 Retail Text-SJT	T2 Cognitive Measure A	T2 Biographical Data A	T2 Retail MMSJT	T2 Cognitive Measure B
T1 Biographical Data A	Pearson r	1							
	Sig.								
	N	148							
T1 Biographical Data B	Pearson r	.90**	1						
	Sig.	0							
	N	148	148						
T1 Retail MMSJT	Pearson r	.73**	.69**	1					
	Sig.	0	0						
	N	148	148	148					
T1 Retail Text-SJT	Pearson r	.26**	.28**	.35**	1				
	Sig.	0.002	0	0					
	N	148	148	148	148				
T2 Cognitive Ability A	Pearson r	.24**	.24**	.19*	0.04	1			
	Sig.	0.003	0.004	0.021	0.642				
	N	148	148	148	148	148			
T2 Biographical Data A	Pearson r	.44**	.42**	.23**	0	.59**	1		
	Sig.	0	0	0.006	0.96	0			
	N	148	148	148	148	148	148		
T2 Retail MMSJT	Pearson r	.31**	.29**	.26**	0.05	.66**	.89**	1	
	Sig.	0	0	0.001	0.568	0	0		
	N	148	148	148	148	148	148	148	
T2 Cognitive Measure B	Pearson r	.31**	.30**	.26**	0.05	.73**	.84**	.93**	1
	Sig.	0	0	0.001	0.538	0	0	0	
	N	148	148	148	148	148	148	148	148

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Note: Actual test names have not been provided above. The following key has been provided: "T1" is equivalent to the Time 1 testing session, "T2" is equivalent to the Time 2 testing session, "Biographical Data A" is equivalent to the CoreBio2, "Biographical Data B" is equivalent to the CoreBio3, "Cognitive Measure A" is equivalent to the Following Instructions Cognitive Ability Measure, and "Cognitive Measure B" is equivalent to the GShortDynamic Cognitive Ability Measure.

In order to test hypothesis 1, an independent samples t-test indicated that individuals with previous experience in the retail industry do not differ significantly from those with no prior retail experience in terms of performance on an animated SJT ($t(53) = .367, p = n.s.$). Although those with no prior retail experience ($M = 6.30, SD = 0.54$) performed better on average than those with prior retail experience ($M = 6.23, SD = 0.78$), this difference was too small to be significant.

For hypothesis 2, a chi-square goodness of fit test was performed to determine whether participants would prefer the animated SJT to the text-based SJT. The results indicate that the obtained frequencies in preferences differ from those expected under the null hypothesis. In other words, the hypothesis that the two distributions, being the preference for the animated SJT preference and the preference for text-based SJ, would be equally preferred in a 50/50 split was rejected, $\chi^2(1, N = 86) = 70.00, p < .001$. Specifically, most participants preferred the MMSJT (81%) to the text-based SJT (19%).

For hypothesis 3, participants were asked which test, the MMSJT or the text-based SJT, would give the best impression to job candidates of an organization. Specifically, participants were asked whether the MMSJT would convey the best impression of an organization to future job candidates or if the text-based SJT would convey the best impression of an organization to future job candidates. A chi-square goodness of fit test was performed to determine whether participants would perceive the MMSJT as conveying a better impression to job candidates about the organization than the text-based SJT. The results indicate that the obtained frequencies differ from those expected under the null. Specifically, there was not a 50/50 split in the distribution, with

significantly more participants stating that the MMSJT would convey a better impression of an organization to future job candidates (83%) than the text-based SJT (17%), $\chi^2(1, N=86)=71.00, p < .001$.

In order to look at hypothesis 4, a chi-square goodness of fit test was performed to determine whether participants would perceive the MMSJT as being more engaging than the text-based SJT. After each SJT, participants were asked to view a checklist of adjectives and check all that apply to that particular form of the SJT. Engaging was one of the answer choices provided, and the number of individuals who checked “engaging” after the text-based SJT was compared to the number of individuals who checked “engaging” after the animated SJT. The results indicate that the obtained frequencies differ from those expected under the null, $\chi^2(1, N=68)=18.00, p < .001$. Specifically, there was not a 50/50 split in the distribution, with the number of participants stating that the MMSJT was engaging (76%) being significantly higher than the number stating that the text-based SJT was engaging (26%).

To further examine hypothesis 4, a chi-square goodness of fit test was performed to determine whether participants would perceive the MMSJT as being more fun than the text-based SJT. The same procedure was used as the one described above. The results indicate that the obtained frequencies differ from those expected under the null, $\chi^2(1, N=28)=5.00, p < .01$. Specifically, there was not a 50/50 split in the distribution, with the number of participants stating that the MMSJT was fun (82%) being significantly higher than the number stating that the text-based SJT was fun (18%).

To conclude hypothesis 4, a chi-square goodness of fit test was performed to determine whether participants would perceive the MMSJT as being more modern than the text-based SJT. The results indicate that the obtained frequencies differ from those expected under the null, $\chi^2(1, N= 51)= 14.00, p < .01$. Specifically, there was not a 50/50 split in the distribution, with the number of participants stating that the MMSJT was modern (73%) being significantly higher than the number stating that the text-based SJT was modern (27%).

To conclude the study, an exploratory analysis was conducted in order to examine the relationship between preference for a particular SJT and cognitive ability. Specifically, the study examined whether those who preferred the text-based SJT would have higher cognitive ability scores than those who preferred the animated SJT. An independent samples t-test indicated that individuals who prefer the text-based SJT do not differ significantly from those who prefer the animated SJT in terms of cognitive ability ($t(74) = -1.575, p = n.s.$). Although those who preferred the MMSJT ($M = 2.01, SD = 1.07$) had slightly higher cognitive ability scores than those who preferred the text-based SJT ($M = 1.51, SD = 1.14$), this difference was too small to be significant.

Discussion

The purpose of this study was to address the lack of literature on Animated Situational Judgment Tests in order to provide additional information to the field. Additionally, this research was conducted to shed light on the perceptions of text-based SJTs versus animated SJTs. The research sought to provide insight into applicant reactions as well as information about what drives the preference for each type of

Situational Judgment Test. The results of this research indicated that animated SJTs are perceived more positively than text-based SJTs. According to Chan and Schmitt (1997), paper-and-pencil tests resulted in high-subgroup differences due to their requirement for literacy and the comprehension of written material. The use of animated SJTs may alleviate the reliance on this component of cognitive ability, potentially leading to less adverse impact (Livens & Sackett, 2006; Chan & Schmitt, 1997). Each of the findings of this study will be discussed in turn below.

First, this study hypothesized that individuals with prior retail work experience would perform better than those who had no prior experience in the retail industry on the animated SJT (Hypothesis 1). These results were surprising, as previous research has found positive correlations with length of experience. Specifically, the older and more experienced an applicant is, the better they perform on an SJT (McDaniel & Nguyen, 2001; Smith & McDaniel, 1998). These results could be due to the homogeneous sample used in this study. Specifically, the mean age for participants was 18-20, providing very little time for experience in the retail industry. It would be beneficial to re-examine this question with an older and more experienced sample in the future.

Next, we looked at the preferences of participants for each situational judgment test. Specifically, I hypothesized that the animated SJT would be preferred to the text-based SJT (Hypothesis 2). Past research has supported this claim, with results indicating that paper-and-pencil measures of performance result in lower applicant reactions than tests involving simulations (Macan, Avedon, Paese, & Smith, 1994; Schmidt, Greenthal,

Hunter, Berner, & Seaton, 1977; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993). In line with the research, this hypothesis was supported, with 75.9% of participants who responded preferring the animatedSJT. These results may stem from the ability of the applicant to see the job-relatedness of tests involving simulations, and therefore preferring to engage in these types of measures.

This research also hypothesized that participants would state that the animated SJT conveyed the best impression of an organization to applicants (Hypothesis 3). Results indicated that participants said the MMSJT would convey a better job impression to applicants than a text-based SJT. Research on the face validity of SJTs indicate that MMSJTs consistently demonstrate higher face validity than text- based SJTs (Chan & Schmitt, 1997; Olson-Buchanan & Drasgow, 2006).

Additionally, this research hypothesized that the animated SJT would be deemed more modern, more fun, and more engaging than the text-based SJT (Hypothesis 4). Support for all three of these claims was found in this research. These results replicate previous findings. Specifically, Olson-Buchanan and Drasgow (2006) found support for this same claim, with applicants rating MMSJTs as more enjoyable and modern than text-based SJTs. The support for this hypothesis as well as hypotheses 2 and 3 all speak to the benefits that the use of animated SJTs in organizations can bring.

As previously mentioned, high face validity indicates that participants perceive the test they are taking as being job-related and relevant to the position for which they are applying. Studies have consistently demonstrated higher face validity for MMSJTs (Chan

& Schmitt, 1997; Olson-Buchanan & Drasgow, 2006). In addition to demonstrating participants' preference for the animated situational judgment test, this research indicated that participants had generally positive reactions to the animated SJT as compared to the text-based SJT. This information could be beneficial to organizations, as the use of an animated SJT provides similar benefits as the multimedia SJT while being more practical in terms of cost and time spent on hiring actors and filming the multimedia aspect of the test. Therefore, organizations may obtain higher face validity as well as better perceptions of being modern and relevant. Using an animated SJT may provide organizations with an edge to their competition in terms of retaining job applicants while simultaneously allowing the organization to gain all of the benefits of an SJT (see Chan & Schmitt, 1997; Funke & Schuler, 1997; McDaniel et al., 2001; McDaniel & Nguyen, 2001; McDaniel et al., 2001; Weekley & Jones, 1999).

The last question addressed in this research focused on cognitive ability. Specifically, the study looked at scores on a cognitive ability test and hypothesized that those who prefer the text-based SJT will show higher cognitive ability scores than those who prefer the animated SJT (Hypothesis 5). Past research has suggested that text-based SJTs result in higher correlations with cognitive ability than MMSJTs (Chan & Schmitt, 1977). This study took an extra step to look into the claim that those who prefer the comprehension component of text-based SJTs may in turn also have higher cognitive ability scores. Support for this claim was not found in this study, with results indicating the opposite trend. In other words, individuals who preferred the animated SJT performed slightly better, although not significantly better, than those who preferred the text-based

SJT on a measure of cognitive ability. Again, as with hypothesis 1, these results may be³⁰ due to the homogeneity of the sample, as well as the high number of participants who preferred the animated SJT to the text based SJT. Of the 74 participants used in the sample, 62 preferred the MMSJT while only 14 preferred the text-based SJT. Because so few participants preferred the test-based SJT, these results should be viewed with caution, and more data should be gathered to further investigate this question.

The current study has some limitations that should be noted. The first is related to the participants in this study. Because participants all came from a Midwestern university, the sample was quite homogenous with respect to age and race. The purpose of this study was to help organizations better understand an emerging selection technique, however the participants in the study were mostly university freshman, making the generalizability of the results not ideal. Additionally, although participants earned extra credit for their participation, they were not in a high-stakes environment. Therefore, the motivation a job applicant may have experienced while taking an SJT may not have translated to the participants of this study.

The current study also poses many possibilities for future research in the area of animated SJTs. This study has provided preliminary information about the nature of the animated SJT while also providing information about the perceptions of the test to job applicants. Participants stated that the animated SJT was more modern, fun and engaging, and preferred the animated SJT to the text-based SJT. In the future, research that provides the criterion-related validity and incremental validity for the animated SJT is needed. This research suggests that it may be similar to that of the multimedia SJT; however further research is necessary in order to support this claim. Additionally, although

research has supported that adverse impact will be reduced when using a multimedia approach to a situational judgment test, at this time there is no empirical support for this claim in relation to an animated SJT. Last, future research that utilizes a sample composed of job applicants will allow for more generalizable results and would be beneficial to the business community as a whole.

References

- 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.
- Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.
- Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology, 86*(3), 410-417.
- Digman, J.M. (1990). Personality structure: Emergence of the five factor model. *Annual Review of Psychology, 41*, 417–440
- Funke, U., & Schuler, H. (1998). Validity of stimulus and response components in a video test of social competence. *International Journal of Selection and Assessment, 6*(2), 115-123.
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology, 91*, 1181–1188.
- Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review, 37*, 426–441.
- Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology, 47*, 715- 738.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III (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*(4), 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.
- Moss, F. A. (1926). Do you know how to get along with people? Why some people get ahead in the world while others do not. *Scientific American*, 135, 26–27.
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, 75, 640–647.
- Olson-Buchanan, J. B., & Drasgow, F. (2006). Multimedia situational judgment tests: The medium creates the message. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests* (pp. 253–278). San Francisco: Jossey-Bass.
- Oostrom, J. K., Born, M. P., Serlie, A. W., & van, d. M. (2011). A multimedia situational test with a constructed-response format: Its relationship with personality, cognitive ability, job experience, and academic performance. *Journal of Personnel Psychology*, 10(2), 78-88.
- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15(1-2), 187-211.
- Schmidt, F. L., Greenthal, A. L., Hunter, J. E., Bemer, J. G., & Seaton, F. W. (1977). Job samples vs. paper-and-pencil trade and technical tests: Adverse impact and examinee attitudes. *Personnel Psychology*, 30, 187-197.
- Smith, K. C., & McDaniel, M. A. (1998, April). *Criterion and construct validity evidence for a situational judgment measure*. Paper presented at the 13th Annual Convention of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Smither, I. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49-76.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.
- 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.

Weekley, J. A., & Ployhart, R. E. (2006). An introduction to situational judgment testing. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests* Mahwah, NJ: Lawrence Erlbaum Associates.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*(3), 188-202

Wiesner, W. H., & Cronshaw, S. F. (1988). The moderating impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61*, 275-290.

Appendix A

Scenario

A customer was just in the store yesterday and purchased a phone and a 12-month service contract. If a customer is not satisfied, he is able to cancel the contract within a week and his money would be refunded. Today he has returned with the phone. The customer walks into the store and sets his phone on the counter.

Dialogue

Customer: “Hi – I want to return this phone. The display doesn’t seem to be working – I get all these strange symbols on it.”

Salesperson: “OK. Let’s take a look and see what might be going on here.”

Customer: (shakes his head) “No. I just want to return the phone and cancel my contract – if it’s already having issues now, I don’t want to invest time trying to fix it.”

Question

Choose the Most Effective and Least Effective course of action from the options below.

Responses

- State that it is likely a problem that is specific to that brand of phone and ask if he would be interested in trying a different brand.
- Agree to return the phone for the customer, but while you process the return, engage in small talk to build rapport and potentially get him to look at other phones.
- Return the phone for the customer and apologize for any inconvenience.
- Ask him whether he would consider taking home a new phone (same model) and see if he continues to have the same issue.

Appendix B

- A. Review the facts below:
- Jennifer has a college degree.
 - Jason works in Tom's department.
 - All of Tom's employees have college degrees.
 - Tom is one of three sales department managers.

Which of these statements must be true?

- a) Tom has a college degree.
 - b) Jennifer works in the sales department.
 - c) Jennifer works in Tom's department.
 - d) Jason has a college degree.
- B. Purchase orders are filed according to amount and company name using the system below.
All closed purchased orders have a \$0 balance. Company names that begin with a number are filed in drawer 2 or 4 according to the purchase order amount.

Amount	Name	Drawer
\$0	A-Z	1
\$0.01 - \$5000	A-M	2
\$0.01 - \$5000	N-Z	3
\$5000.01 - up	A-M	4
\$5000.01 - up	N-Z	5