

Fall 2016

Game-Framing Cognitive Assessments to Improve Applicant Perceptions

Andrew Burnett Collmus
Old Dominion University

Follow this and additional works at: http://digitalcommons.odu.edu/psychology_etds

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

Recommended Citation

Collmus, Andrew Burnett, "Game-Framing Cognitive Assessments to Improve Applicant Perceptions" (2016). *Psychology Theses & Dissertations*. 41.
http://digitalcommons.odu.edu/psychology_etds/41

This Thesis is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

**GAME-FRAMING COGNITIVE ASSESSMENTS TO IMPROVE APPLICANT
PERCEPTIONS**

by

Andrew Burnett Collmus
B.S. May 2014, Colorado State University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfilment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
December 2016

Approved by:

Richard N. Landers (Director)

Xiaoxiao Hu (Member)

Bryan Porter (Member)

ABSTRACT

GAME-FRAMING COGNITIVE ASSESSMENTS TO IMPROVE APPLICANT PERCEPTIONS

Andrew Burnett Collmus
Old Dominion University, 2016
Director: Dr. Richard N. Landers

Research has shown that although cognitive testing is key to quality hiring, applicants often react poorly to cognitive ability tests. Applicant reactions theory indicates that time-length judgments of a selection procedure can affect applicant perceptions. It was thus hypothesized that game-framing, the act of labeling something a game without changing the content, would cause participants to perceive that time was moving faster while completing a battery of cognitive ability tests. Similarly, it was expected that game-framing would increase test motivation and decrease test anxiety. Perceived length was tested as a mediator for the effects of game-framing on test anxiety and on test motivation. Structural equation modeling was used to evaluate the hypothesized relationships. In the observed dataset, game-framing caused decreases in perceived length, perceived length was positively related to test motivation, and perceived length mediated the relationship between game-framing and test motivation. The results of this study demonstrate that game-framing affects time perceptions. This finding has implications for gamification researchers, namely, that game-framing effects should be measured and accounted for in future studies. Furthermore, applicant reactions theorists have suggested that perceived time length is a key variable in the overall applicant reactions model, and this study is the first to empirically investigate perceived time length of a selection procedure in this context. Results

indicate that perceived length may not relate to other applicant reaction variables as predicted by applicant reactions theory.

I dedicate this paper to my wife, Mary. Without you, I would not be where I am today.

ACKNOWLEDGMENTS

First off, I would like to thank my advisor, Richard Landers, for his guidance on this and many other projects. I feel lucky to have an advisor who quickly and thoughtfully attends to his students, providing advice and revisions at all hours of the day. My development as a writer and as a scholar has greatly progressed under your guidance.

To my committee members, Xiaoxiao Hu and Bryan Porter, thank you so much for being flexible and available during this process. You have both been accommodating and gone out of your way to ensure the thesis got done on time. Through your advice and review of my work, I am confident that this project will become my first peer-reviewed journal submission, and I am grateful.

To Peggy, Mary, and Linda, thank you all for helping me along this process. I would not be on track without your help in several instances.

To all the ODU Psychology students, thank you for the friendship and support over these past few years. It has been great.

To my wife, Mary, thank you for dealing with the long hours and crankiness. I am not sure how you do that, and I'm certain most people would not. I love and appreciate everything you do for me.

To my dog, Tracer. Stop making me feel guilty every morning when I leave the house! But thanks for being ecstatic every night when I come home.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	viii
Chapter	
I. INTRODUCTION	1
APPLICANT REACTIONS THEORY	3
GENERAL COGNITIVE ABILITY	5
GAMIFICATION	7
FRAMING.....	10
HYPOTHESES	11
II. METHOD	16
PARTICIPANTS.....	16
MATERIALS AND MEASURES.....	17
PROCEDURE	18
III. RESULTS.....	20
HANDLING OF CASES EXHIBITING CARELESS RESPONDING	20
MODEL FIT	21
HYPOTHESES	25
IV. DISCUSSION	29
LIMITATIONS	32
FUTURE DIRECTIONS.....	34
CONCLUSION	35
REFERENCES	36
APPENDICES	42
A. 3D ROTATION SAMPLE ITEM	42
B. LETTER AND NUMBER SERIES SAMPLE ITEM	43
C. MATRIX REASONING SAMPLE ITEM.....	44
D. ROGRESSIVE MATRICES SAMPLE ITEM	45
E. VERBAL REASONING SAMPLE ITEM.....	46
F. TEST ATTITUDES SURVEY	47
VITA	49

LIST OF TABLES

Table	Page
1. Latent Variable CFAs	22
2. Direct and Indirect Effects	23
3. Descriptive Statistics and Correlations	24

LIST OF FIGURES

Figure	Page
1. Theoretical Model.....	15
2. Conceptual Model with Path Estimates	28

CHAPTER 1

INTRODUCTION

General cognitive ability has been found to be a consistent predictor of job performance across jobs (Schmidt, Oh & Shaffer, 2013); however, applicants often react poorly to cognitive ability tests used in employee selection (Gilliland, 1993). This has been the general state of the literature for decades. General cognitive ability has been conclusively demonstrated to be the best individual predictor of job performance (Schmidt & Hunter, 1998), and its predictive validity increases as the criterion job's task complexity increases (Hunter, 1986). Yet research has also consistently shown that cognitive ability tests are rated less favorably than resumes, work samples, and interviews (Gilliland, 1993; Hausknecht, Day, & Thomas, 2004; Anderson, Salgado, & Hülshager, 2010). The same body of research has shown that poor reactions can lead to applicant withdrawal (Ryan, Sacco, McFarland, & Kriska, 2000), which reduces the utility of a selection system. Other outcomes of poor applicant reactions include decreased organizational attraction and offer acceptance intentions, and an increase in litigation intentions.

According to applicant reactions (AR) theory, such poor reactions to cognitive ability tests can be explained by poor perceived procedure characteristics and applicant perceptions, including increased perceived test length, decreased motivation to test, and increased test anxiety (Hausknecht et al., 2004). Perceived procedure characteristics are constructs that represent how applicants perceived various aspects of the selection procedure. Perceived length is one such characteristic. Such characteristic theoretically cause changes in broader applicant perceptions, including test anxiety and test motivation. Thus, the existing AR literature suggests that changes in perceived length of process should correspond to changes in test anxiety and test motivation.

Given this theoretical framework, presenting cognitive ability assessments to test-takers as if they were games, a test modification called “game-framing,” can be theoretically linked to improved reactions. Game-framing is a type of framing, which refers to a change in the way information is presented to impact the manner in which it is perceived. Game-framing in particular is intended to modify perceptions of a task or process to make it seem more game-like, making it a type of gamification (Armstrong, Landers, & Collmus, 2015). In the terms of the framing literature, game-framing is a type of attribute framing, which refers to a change intended to modify the attractiveness of a task or item (Levin, Schneider, & Gaeth, 1998). When a cognitive ability test is game-framed, it is in part intended to encourage people to get lost in the game so that they perceive time is moving faster. Thus, the general goal of this paper is to demonstrate that when cognitive ability tests are framed as games, test-takers perceive less time passes during testing, causing increases in motivation to complete the assessment and decreases in anxiety.

Although AR theory provides a general picture as to which individual perceptions and attitudes might change as a result of game-framing, the specific mechanisms are currently unknown and therefore represent a gap in the current literature. That is, AR literature describes the nomological net of reactions to selection procedures but does not provide many practical recommendations on how to modify selection practices to improve actual reactions (Ryan & Huth, 2008), especially in the context of game-framing. To fill this gap, the goal of this research theoretically is to extend AR theory by proposing and empirically supporting the direct effects and mediational processes for the impact of game-framing on applicant perceptions as described above. Thus, this study will build upon AR theory and also further the study of gamification.

Applicant Reactions Theory

AR theory describes antecedents, moderators, and outcomes related to applicant perceptions of selection procedures. Antecedents of applicant perceptions include person characteristics, such as personality and race, job characteristics, perceived procedure characteristics, and organizational context (Hausknecht et al., 2004). Outcomes of applicant reactions include such things as organizational attractiveness, which affects job pursuit or offer acceptance intentions, likelihood and success of litigation against the hiring organization, and validity and utility (Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993). Stage in the selection process, job desirability, and available alternatives, among other processes, moderate the relationship between applicant perceptions and outcomes (Hausknecht et al., 2004).

Organizational justice theory is the foundation of AR theory, although it has been extended to include other theoretical mechanisms. In two seminal AR papers, Gilliland (1993) and Smither et al. (1993), both proposed that organizational justice perceptions are the link between selection procedures and various outcomes. That is, perceptions of fairness explain the link between a given selection procedure and a candidate's likeliness to withdrawal from the application process, recommend the organization to friends and colleagues, reapply to the organization, accept an offer, or other outcomes in AR literature. In recent years, as AR theory has expanded to include new antecedents and moderators beyond those that are traditionally found in the organizational justice literature (e.g., Hausknecht et al., 2004; Ryan & Ployhart, 2000). Such additional variables now include, stage in the selection process, industry norms, job norms, market conditions (e.g., unemployment and job availability), individual characteristics (e.g., work experience, demographics, personality), and organizational context.

According to AR theory, applicant perceptions are the focal processes involved in AR. Applicant perceptions are a class of variables that are the core mediator between antecedents and outcomes in the AR model and include *procedural justice*, *distributive justice*, *test anxiety*, *test motivation*, *attitudes toward tests*, and *attitudes toward selection*, among others (Hausknecht et al., 2004). These constructs describe internal thoughts or feelings that buffer the effects of applicant reactions and perceptions on outcomes. Of particular note in AR theory in relation to game-framing are two particular perceptions: *test anxiety* and *test motivation*. Test anxiety is a situational state that is in part related to fear of negative evaluation (Hembree, 1988). Test motivation is an applicant perception that describes the level of effort one is willing to put forth on a given test. According to Campbell, McCloy, Oppler, and Sager (1993) three factors comprise task motivation: (1) the choice to exert effort, (2) the intensity and (3) duration effort to put forth. In a similar vein, test motivation can be conceptualized as the combination of choice, intensity, and duration of effort one is willing to put forth in a test as well as the level of desire one has to achieve a maximal outcome.

Perceived procedure characteristics are theorized to be direct antecedents to applicant perceptions, and of particular interest to those considering game-framing is perceived length of process. Ryan and Ployhart (2000) suggested that if selection processes are perceived as too lengthy, it may be demotivating to applicants; however, I could identify no empirical evidence supporting this relationship. The relationship between perceived length and test anxiety appears to be similarly untested. Research on time perceptions more broadly demonstrates several different ways that individuals perceive time and time passage. In general, there are two high-level subjective time perceptions: interval length estimation and passage of time judgments. The former is typically assessed by interval estimation (e.g., “estimate in minutes and seconds the

length of the experience”), whereas the latter is typically assessed by speed perceptions on a likert type scale (e.g., from 1 = *time dragged* to 5 = *time flew*; Sucala, Scheckner, and David, 2011). Time estimations come in two varieties, prospective and retrospective (Wearden, 2005). In prospective estimation, participants know ahead of time that they will be estimating time passage. In retrospective estimation, it is only after a task is completed that participants become aware that they are to estimate the subjective passage of time. This distinction has methodological implications and is therefore explained further in a later section.

General Cognitive Ability

Human cognitive ability is a hierarchical construct consisting of one general factor (g) at the highest level, operationalized as the shared variance between a variety of cognitively-loaded tasks (Kuncel, Hezlett, & Ones, 2004). Contemporary intelligence theorists generally ascribe to a 3-stratum theory of intelligence. The second-order dimensions of g are typically specified as fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval ability, broad cognitive speediness, and processing speed (Carroll, 1993, p. 626). Each of those constructs can also be conceptualized as consisting of even more specific dimensions, such as reading speed, spelling ability, or quantitative reasoning (Carroll, 1993). At each level of the g hierarchy, variables exhibit positive manifold (i.e., they correlate positively). The property of positive manifold means that different cognitive ability assessments, using different assessment tasks, can be used to infer g (e.g., Wonderlic, WAIS, Shipley’s, Stanford-Binet); Spearman referred to this as *indifference of the indicator*.

Cognitive ability is well-established as a consistent predictor of job performance in a variety of contexts. Additionally, g predicts proclivity to learn new tasks. Thus, g is the single

best differentiator of complex task performance – those high in *g* will learn new tasks more quickly and be able to perform mastered tasks at a higher level than those who are lower in *g*. In addition to theoretical soundness, these assertions are backed by nearly a century of empirical data collected in organizations and academic settings (Kuncel et al., 2004; Schmidt & Hunter, 2004).

Because applicant reactions to cognitive ability tests are generally poorer than to resumes, work samples, and interviews (Anderson et al., 2010; Gilliland 1993) in the United States, such tests are a prime target for an intervention to improve reactions. More specifically, since *g* is the apparent single best predictor of job performance, it would benefit organizations most if it was also the one most preferred by applicants. Since this is not the case, there is room to improve applicant reactions to tests of *g*.

Cognitive ability tests themselves are prime candidates for game-framing. Specific definitions of *g* abound, but researchers generally agree that it consists of the ability to reason, think abstractly, solve problems, and acquire new knowledge (Snyderman & Rothman, 1988, p. 56). As a reflection of that, many cognitive assessment tasks require mental object rotation (similar to a puzzle), or applied logic (similar to riddles). Thus, since many cognitive assessments are similar to existing leisure games (e.g., puzzles and riddles), the presentation of these assessments as games should be highly believable. In other words, the questions commonly seen on cognitive assessments naturally lend themselves to game-framing.

Cognitive ability tests have previously been presented as games or puzzles by researchers to affect perceptions of the testing situation, although the effects of this presentation have not been studied in the context of personnel selection. For example, Buford and O’Leary (2015) used performance in *Portal 2*, a puzzle and problem solving video game, to predict scores on several

measures of fluid intelligence. They found strong positive correlations between Portal 2 performance and Raven's Progressive Matrices ($r = .49, p < .001$) and Shipley-2 Block Pattern ($r = .49, p < .001$). Rather than using off-the-shelf video games, such as Portal 2, other researchers have created video games with the express purpose of assessing intelligence. For example, Delgado, Uribe, Alonso and Diaz (2016) created a battery of assessment games, which tended to correlate strongly with subtests of the Wechsler Intelligence Scale for Children (WISC-III) between $r = .43$ and $r = .60$ (these correlations represent those between a given game in the assessment battery and its intended corresponding WISC-III subscale). However, since assessment games are a method, rather than a construct, there is variability in their efficacy as measurement instruments. As another example, McPherson and Burns (2008, Table 7) reported lower correlations for two games designed to assess g and working memory, between $r = .44$ and $r = .22$ (this range includes significant correlations between an assessment game and another intelligence instrument, such as Raven's Progressive Matrices). These examples demonstrate that in some cases, the line between game and cognitive ability test is unclear. Given this, I contend that some cognitive ability tests, especially those already containing game-like features, could be readily perceived as games if presented as such.

Gamification

Gamification involves the use of game elements to shape behaviors, perceptions, or psychological states. Game elements are attributes of games that differentiate them from normal tasks. Garris, Ahlers, and Driskell (2002) list six game elements that differentiate games from simulations, fantasy, rules/goals, sensory stimuli, challenge, mystery, and control. More recently, Wilson et al. (2008) listed 19 different game elements found in learning games. Game elements can exist explicitly within the game's programming or implicitly within the mind of the player

(Collmus, Armstrong, and Landers, 2016). A standalone game (e.g., a console video game), contains many game elements. Gamification represents the spectrum between a standalone game and a mundane or serious task; as more game elements are added, the task becomes more game-like. A ubiquitous example of gamification within organizations is the use of leaderboards to create an atmosphere of competition related to job performance (Landers, Bauer, & Callan, 2015). In this context, leaderboards are intended to affect motivation, which is in turn expected to increase the actual desired behavior, job performance.

Previous psychological research has in fact demonstrated that one type of gamification, game-framing, can affect performance on intelligence tests, although this was not the intended purpose of that research. Steele and Aronson (study 1, 1995) sought to investigate potential causes of racial group differences on intelligence test scores. They administered a 30-item test, composed of difficult GRE verbal items, to three different groups. Group 1 was told that the test was diagnostic of verbal ability, that the purpose of the research was to investigate individual differences related to reading and verbal abilities, and that they would be provided their results to better understand their own strengths and weaknesses related to verbal problem solving; these participants were told to put forth a strong effort in order to help the researchers analyze their verbal ability. Group 2 was told that the research was intended to better understand the psychological factors of verbal problem solving; the study description did not mention verbal ability, and the participants were told that performance feedback would be provided to better familiarize them with the type of questions they will be seeing on future tests (presumably in the next few years of undergraduate education). Group 3 had similar treatment to Group 2, except they were told that the exercise was a mental challenge, and the purpose of the research was to investigate the problem solving process; again, verbal ability was not mentioned to the

participants in this condition. The authors reported a significant main effect, $F(2, 107) = 4.74, p < .02$ in which participants in the challenge condition performed higher than participants in the diagnostic and non-diagnostic conditions (i.e., Group 3 outperformed Groups 1 and 2). Thus, although these researchers were not explicitly measuring game-framing, the results from study 1 showed that higher scores on a cognitive ability test (GRE Verbal) were obtained by presenting the test as a challenge, rather than as a test of ability.

Research in other areas of social science also suggests that game-framing should have an effect on perceptions. For example, it has been shown that test-takers rated a gamified survey as more enjoyable than the non-gamified version. Downes-Le Guin, Baker, Mechling, and Ruyle (2012) found that as more game elements were added to a survey, the survey was rated increasingly enjoyable by respondents. Mavletova (2014) found similar effects in an adolescent sample when her gamified survey was rated as more enjoyable than the non-gamified version. Thus, empirical evidence indicates that game elements can improve the test-taking experience from the respondent's perspective.

Previous research in gamification demonstrates motivational effects when describing tasks as games to research participants. Lieberoth (2015) randomly assigned participants to three conditions: core task, game-framed task, and game mechanics task. The core task in this study was a facilitated small group discussion of written cues in which the participants rated each response from one to five stars, five stars being the best rating. The participants in each group took turns reading the prompt and facilitating discussion. The game-framed task utilized a game board and pawns, in which the prompts were read from cardboard game cards. The participants in each group would take turns reading the prompts, as before, and then would move their pawn one space forward on the game board. Although this condition had the initial appearance of a

game, many game elements necessary to classify the task as a “game” were missing.

Specifically, each turn the facilitator moved their pawn forward once space, with no regard to the discussion content or its rating. Thus, performance was not contingent on any actions under the players’ control, a defining aspect of games. In the full game condition, participants took turns reading prompts in their groups as before but would also move their pawn forward a number of spaces that corresponded to the group’s rating of the facilitated discussion. They were also given a game objective to reach the end of the board first. Thus, participants in this condition had a game with rules that they could win by facilitating better discussions when it was their turn. Lieberoth reported significantly higher interest/enjoyment ratings in the game-framing ($M = 3.247$, $SD = 0.854$, $p < .007$) and full game ($M = 3.366$, $SD = 0.596$, $p < .001$) conditions than in the non-game condition ($M = 2.602$, $SD = 0.614$). Although this lends some initial support to the existence of a game framing effect, because framing was accompanied by the addition of a game board and game pieces, the game framing effect measured is confounded with any effect of those game elements, limiting the interpretability of this study in regards to framing in particular.

Framing

More broadly, it has been demonstrated that framing can modify human thought and decision-making processes in significant ways. Tversky and Kahneman (1981) provide an example of the classic risky choice paradigm. In their study, the choice-makers (participants) were asked to decide which course of action to pursue in the face of a hypothetical disease outbreak that is expected to kill 600 people. In the first condition, participants could choose between two options. In Program A, 200 people would be saved, whereas in Program B, there was a 1 in 3 chance that 600 will be saved and a 2 in 3 chance that no people will be saved. In the second condition, the choice was reframed as Program A, in which 400 people would die, or

Program B, in which there is a 1 in 3 possibility that no one will die and a 2 in 3 possibility that 600 people will die. Although worded differently, the choices presented to the two groups were objectively identical in result probabilistically. However, the change in wording corresponded to a change in the choice that participants were likely to make. More specifically, the change in wording corresponds with a change in majority preference from the risk-averse option to the risk-taking option (in this case, risk-averse defines the choice that guarantees lives will be saved, while risk-taking defines the choice in which there is a possibility that no lives will be saved). When the choice is framed as an opportunity to guarantee 200 lives will be saved, the majority of participants chose that option. When the choice is framed as a guarantee that 400 lives will be lost, the majority of participants made the alternate choice.

One particular type of framing, attribute framing, should be of interest to gamification researchers due to its relative simplicity in implementation. Attribute framing involves framing object or event characteristics to modify its attractiveness (Levin, Schneider, Gaeth, 1998). As an example, Levin and Gaeth (1988) found that when beef was labeled “75% lean” it was rated as better tasting than when the same beef was labeled “25% fat”, thus demonstrating the effects of framing on human perception; specifically, the way something is presented can make it seem more appealing. This type of framing is a bit simpler than examples demonstrated via the risky choice paradigm and thus may be applicable in a wider variety of situations. However, much like risky choice framing, the meaning or content is identical but the presentation is different.

Hypotheses

Since theory and empirical evidence demonstrate that framing can affect perceptions of a choice, task, or item, I contend that game-framing cognitive assessments should make those assessments seem more gamelike to test-takers. In other words, because framing can be used to

influence perceptions in other contexts, game-framing should influence perceptions in the present context. Specifically, the goal of game-framing cognitive ability assessments is to encourage participants to accept the idea that they are completing puzzles and riddles rather than a cognitive ability assessment.

When individuals are highly engaged in an activity, their time perceptions are altered (Gable & Poole, 2012). During fun experiences, such as games, time seems to move more quickly (Sackett, Meyvis, Nelson, Converse, and Sackett, 2010). Previous studies have shown that gamified surveys are perceived as more engaging, evidenced by perceptions of test length and enjoyment in comparison to nongamified tests or surveys (Downes-Le Guin et al., 2012; Lieberoth, 2015; Mavletova, 2014). Therefore, since game-elements can cause a survey to seem more enjoyable, and enjoyable activities seem to take less time, the first hypothesis follows.

Hypothesis 1. Game-framed cognitive tests will have lower perceived length than those same tests when assessment-framed.

Prior research (e.g., Downes-Le Guin et al., 2012; Lieberoth 2015; Mavletova, 2014) demonstrates that gamification may improve motivational factors related to test-taking, and also that reactions to cognitive ability tests, of which motivation is a key component (Hausknecht et al., 2004), are poor relative to their standing as an excellent predictor of performance (Anderson et al., 2010). Therefore, the use of game-framing with cognitive ability tests is likely to improve the motivational component of applicant reactions.

Hypothesis 2. Participants will report higher test motivation for game-framed cognitive ability tests than those same tests when assessment-framed.

Because test anxiety is in part due to fear of negative evaluation (Hembree, 1988), assessments framed as a puzzle-game rather than as an evaluation of ability should additionally

invoke fewer internal thoughts and emotions that result from comparing oneself to others or worrying about test results. The results of research by Steele and Aronson (1995) suggest that game-framed cognitive tests are perceived as less evaluative. In their diagnostic condition, in which the tests were presented as ability assessment tools, 65% of participants believed that the purpose of the study was to evaluate their abilities. In the game-framed condition, in which these tests were presented as a challenge, only 11% of participants believed the purpose of the study was to evaluate their abilities.

Hypothesis 3. Participants will report lower test anxiety (the measured aspect of test anxiety) for game-framed cognitive ability tests than those same tests when assessment-framed.

Theory and research suggest a link between time perceptions and motivation. Campbell et al, (1993) describe duration of effort as a key part of motivation. Thus, as time seemingly moves faster or slower, the duration of effort choice should, respectively, last longer or shorter in actual time. As would be expected from Campbell and colleagues' theory of task motivation, metacognition researchers have found a relationship between motivation and faster time perception (Gable & Poole, 2012). Furthermore, meta-analytically derived estimates of the relationship between perceived procedure characteristic variables and applicant perception variables are moderately positive; 21 meta-analyses of the various relationships between these variable classes provided average effects ranging from .14 to .54 (Hausknecht et al., 2004).

Hypothesis 4. Perceived length will be negatively correlated with test motivation

A component of test anxiety is the tendency to ruminate on one's standing relative to others' performance (Lievens et al., 2003). Consequentially, test-takers who perceive that time is moving quickly will also perceive less time to ruminate, resulting in lowered perceptions of anxiety. This relationship is also suggested by applicant reactions theorists (e.g., Hausknecht et

al., 2004; Ryan & Ployhart, 2000). Meta-analytic results between perceived procedure characteristics and applicant perceptions range from .14 to .54 (Hausknecht et al., 2004); however, since anxiety is not a positive outcome it is expected that the relationship will be negative, but of similar size. It is furthermore worth noting that since time perceptions are implied by theory but not empirically tested, this range of meta-analytic estimates (i.e., from Hausknecht et al., 2004) is the best existing data from which to calibrate hypotheses 4 and 5.

Hypothesis 5. Perceived length will be positively correlated with test anxiety

Finally, tests of the indirect effects of game-framing on reactions (test motivation and test anxiety) via time perceptions will be tested. Specifically, applicant reactions models (e.g., Hausknecht et al., 2004; Ryan & Ployhart, 2000) dictate that the selection procedure characteristics organizations can control, such as game-framing, affect applicant perceptions as a function of their intermediary effect on perceived procedure characteristics. In the AR model, perceived procedure characteristics represent an antecedent class of variable in the AR model, whereas applicant perceptions are the central class of variable. In contextualizing gamification in general and game-framing specifically to the AR context, the present study therefore proposes that the relationship between game-framing and the two applicant perceptions described earlier is mediated by the perceived procedure characteristic perceived length.

Hypothesis 6. Perceived length will mediate the relationship between game-framing and test motivation.

Hypothesis 7. Perceived length will mediate the relationship between game-framing and test anxiety.

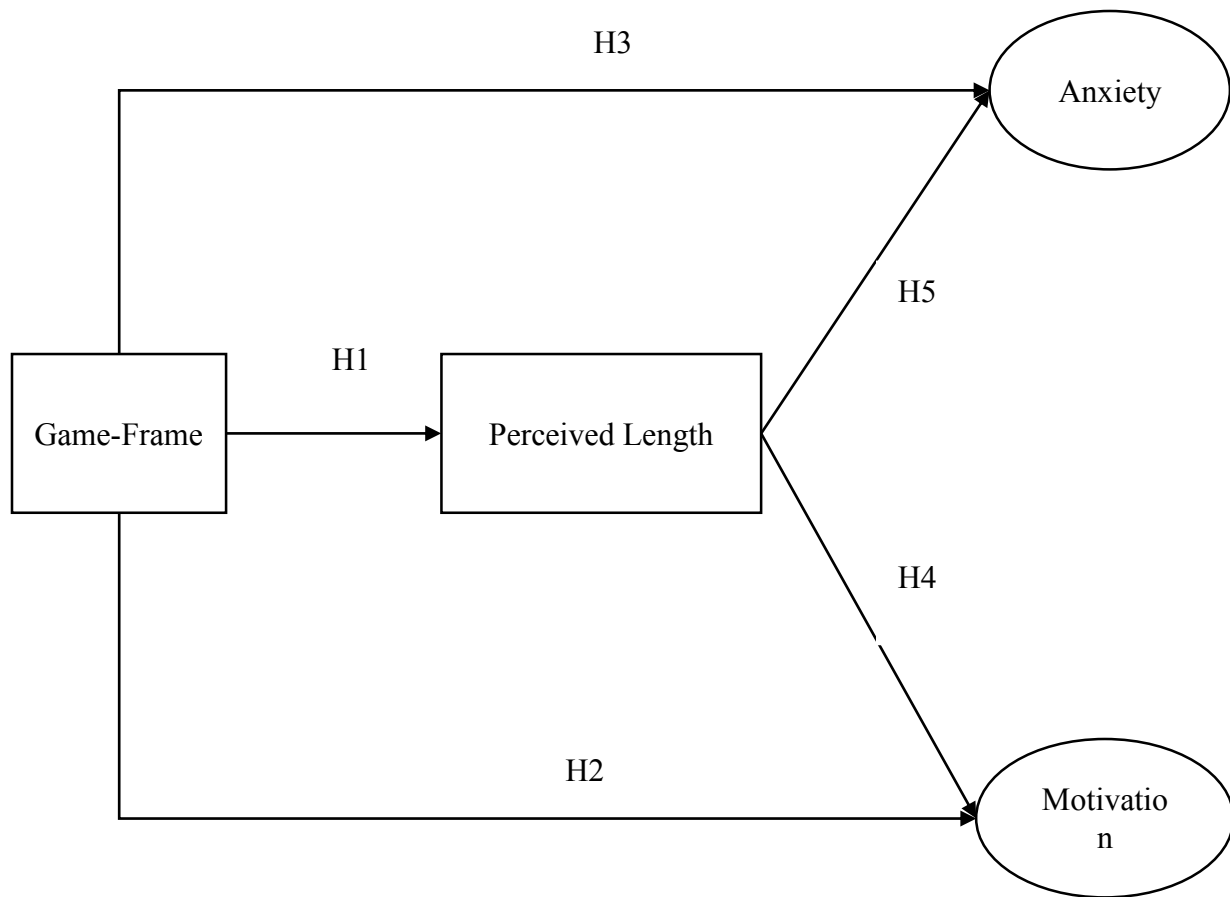


Figure 1. Theoretical model

CHAPTER 2

METHOD

Participants

In total, 358 participants were recruited for the study. A power analysis was conducted based upon a previous game-framing effect study (Lieberoth, 2015; $\eta^2 = 0.197$), which indicated that 206 participants would be required to detect the effect using a Pearson's correlation. Due to expected incompletions or other bad data, around 350 participants were recruited in order to achieve the 206 number of usable data points indicated by power analysis. Additional participants for the present study were also required in order to test the structural theoretical model with structural equation modeling (SEM). Participants were recruited via MTurk and paid \$1.75 to participate.

After data cleaning, the final sample (see Careless Respondent section in Ch. 3) consisted of 336 individuals, of whom 30 (8.9%) self-identified as African American or Black, 2 (0.6%) as Arab American, 17 (5.1%) as Asian American, 257 (76.5%) as European American or White, 2 (.6%) as Native American or Native Alaskan, 1 (.3%) as Pacific Islander or Native Hawaiian, 10 (3.0%) as Other (single race), and 17 (5.1%) self-identified as Two or more races. One-Hundred and Sixty (47.6%) self-identified as male, while the remainder self-identified as female. Seven participants (2.1%) were ages 65 and older, while 329 (97.9%) were in 18-64 age range. Two-Hundred and Eighty participants (83.3%) were employed either full ($N = 225$; 67.0%) or part-time ($N = 55$; 16.4%), with an average workweek of 38.62 hours and average job tenure of 7.47 years.

Materials and Measures

Cognitive ability tests. The following five cognitive ability tests were obtained from the International Cognitive Ability Resource Team (ICAR, 2014): Three Dimensional Rotation, Letter and Number Series, Matrix Reasoning, Progressive Matrices, and Verbal Reasoning. Items from these tests can be found in Appendices A through E, respectively. Condon and Revelle (2014) conducted a series of validity tests regarding the ICAR items. Specifically, they investigated the factor structure of each test and also of the overall battery and verified the presence of an emergent *g* factor. In addition, Condon and Revelle also sought validation of the ICAR items against a commercially available intelligence assessment, finding corrected correlations between the 16-item ICAR sample test and Shipley-2 composites A and B of 0.82 and 0.81, respectively.

Game-framing. Game-framing, the experimental IV, was dummy coded such that 0 represents the assessment frame and 1 represents game frame.

Time length. Time estimations come in two varieties, prospective and retrospective (Wearden, 2005). In prospective estimation, participants know ahead of time that they will be estimating time passage. In retrospective estimation, it is only after a task is completed that participants become aware that they are to estimate the subjective passage of time. This distinction has important bearings for research methodology, namely, that retrospective time estimation can only occur once in a given study. Once individuals are given a time estimation prompt, they are unable to avoid thinking about the passage of time (Wearden, 2005). This discrepancy in thought creates a bifurcation between the first and any subsequent time estimations. Thus, to obtain multiple time estimations from the same individuals in a given study, one must stick to prospective estimation or adopt the methodology recommended by Hicks

(1992) in which participants are provided with cues and asked to make retrospective judgments of multiple events all at once. That is, after the study completed, participants were asked to provide subjective time estimations for several different events that occurred during the study.

Time length estimation was assessed using the interval-length estimation procedure described by Sucala et al. (2011). Participants were asked to “estimate, in minutes and seconds, the length of the experience”. An example answer was provided that read “e.g., 6:20 equals six minutes and twenty seconds”.

Applicant perceptions. Applicant perceptions were measured with two subscales of the Test Attitude Survey described by Arvey, Strickland, Drauden, and Martin (1990). The Test Attitude Survey consists of 9 dimensions, including a 10-item test motivation scale ($\alpha = .85$) and a 10-item comparative anxiety scale ($\alpha = .80$). These two scales were used in the study. The full 45-item TAS scale can be found in Appendix F.

Demographics. A brief demographic questionnaire asked participants to report their age, gender identity, racial identity, and employment status. If a participant indicated current employment, it was requested that the participant indicate their industry, tenure, and current hours worked per week.

Procedure

Participants were recruited from MTurk’s online platform. Within the MTurk interface, the study was titled *Problem Solving*. The study description read “this is part of a university study on human problem solving.” After agreeing to participate in the study, participants were randomly placed into either the game-frame or the assessment-frame condition via a random number generator in the Qualtrics survey platform. In the assessment-frame condition, participants were asked to complete the cognitive ability measures with the following

instructions: “This research is part of a larger effort to understand adult cognitive functioning. You will complete a series of intelligence tests, similar to those used in IQ assessment. Please follow the instructions on each task, and make your best effort to select the correct response”. Participants in the game-frame conditions were given different instructions but otherwise experience identical stimuli: “This research is part of a larger effort to understand how humans solve puzzles, logic games, and riddles. You will be completing a series of puzzles and logic games. Please follow the instructions on each task, and make your best effort to select the correct response.” Participants then completed all cognitive tests, followed by time-length estimation, applicant reactions scales, and a demographic survey.

CHAPTER 3

RESULTS

Handling of Cases Exhibiting Careless Responding

Careless responding has an adverse effect on the interpretation of survey results because those observed responses do not reflect the constructs intended by the researcher (Meade & Craig, 2012). Recently, methodologists (e.g., DeSimone, Harms, & DeSimone, 2015; Meade & Craig, 2012) have recommended the use of Mahalanobis distance (MD) to identify multivariate outliers, because extreme MD values, which reflects multivariate outlyingness, can signal careless survey response patterns. When using this technique, an MD value for each scale within each case is calculated by regressing the scale mean onto its item scores and saving the MD values resulting from that analysis. These MD values follow a chi-square distribution, in which the degrees of freedom are equal to the number of items in that scale (McLaughlan, 1999). Given this, a p -value can be calculated for each MD value and therefore each scale. Thus, for the cognitive ability and test attitude scales, an MD and accompanying p -value was calculated for each participant on each scale based on their response patterns within the scale. As an exception, the MD values for 3D Rotation were not used in screening calculations because that scale proved exceptionally difficult for participants. Forty-six percent of respondents ($N = 164$) did not answer a single 3D Rotation item correctly. Thus, higher MD values on this scale may have indicated astute participants rather than careless responders. Individuals who had 2 or more significant MD values ($N = 22$) were removed from further analyses. After the cleaning procedure, hypothesis testing in Mplus then proceeded with the cleaned sample of 336 participant response sets.

Model Fit

The model depicted in Figure 1, including all direct and indirect effects, was tested with structural equation modeling using Mplus 7 (Muthén & Muthén, 2015) with a slight modification. After revisiting Arvey et al. (1990), it was determined that the anxiety and motivation scales should be allowed to covary, and this lack of covariance was leading to misfit in the proposed model. With this change, the model was tested using maximum likelihood estimation. The initial results indicated poor fit to the data, ($\chi^2(205) = 755.268, p < .001$; RMSEA = 0.089; CFI = 0.83; SRMR = .063); however, this initial model yielded several Modification Index (M.I.) values greater than 10. During the model respecification process, several within scale error variances were allowed to correlate. Specifically, the within-scale error variances with the largest M.I. value was allowed to correlate, and then the model was re-analyzed in Mplus. This process was repeated until the Mplus software no longer indicated within-scale error variance M.I. values greater than 10. Eight paths were freed for the anxiety scale and eight for the motivation scale. After respecification, the model had adequate fit to the data ($\chi^2(189) = 344.228, p < .001$; RMSEA = .058; CFI = .952; SRMR = .051), supporting the proposed structural and measurement models. Factor pattern loadings for the two latent variables, test anxiety ($\alpha = .899$) and test motivation ($\alpha = .929$), can be found in Table 1. Table 2 shows path model estimates, and Figure 2 shows the updated model with path estimates. Table 3 shows variable means, standard deviations, and correlations.

Table 1

Latent Variable CFAs (standardized factor pattern loadings)

Item	Estimate	SE	<i>p</i>
Test Anxiety ($\alpha = .899$)			
anx1	0.796	0.031	<.001
anx2	0.783	0.036	<.001
anx3	0.730	0.034	<.001
anx4	0.658	0.042	<.001
anx5	0.699	0.041	<.001
anx6	0.666	0.041	<.001
anx7	0.585	0.046	<.001
anx8	0.693	0.039	<.001
anx9	0.498	0.051	<.001
anx10	0.585	0.046	<.001
Test Motivation ($\alpha = .929$)			
moti1	0.640	0.043	<.001
moti2	0.817	0.030	<.001
moti3	0.776	0.033	<.001
moti4	0.791	0.025	<.001
moti5	0.828	0.029	<.001
moti6	0.744	0.034	<.001
moti7	0.764	0.031	<.001
moti8	0.676	0.035	<.001
moti9	0.741	0.034	<.001
moti10	0.712	0.039	<.001

Table 2

Direct and Indirect Effects (Maximum Likelihood Estimation)

Hypothesis	Path	Estimate	SE	<i>p</i>	95% CI*
Unstandardized Direct Effects					
1	Cond to PL	-3.003	1.492	0.044	[-5.930, -0.106]
2	Cond to Moti	-0.022	0.065	0.731	[-0.146, 0.110]
3	Cond to Anx	-0.146	0.101	0.147	[-0.340, 0.057]
4	PL to Moti	0.005	0.003	0.041	[0.000, 0.011]
5	PL to Anx	-0.146	0.101	0.147	[-0.005, 0.012]
Standardized Direct Effects†					
1	Cond to PL	-0.221	0.106	0.038	[-0.420, -0.004]
2	Cond to Moti	-0.039	0.112	0.731	[-0.256, 0.185]
3	Cond to Anx	-0.171	0.119	0.152	[-0.398, 0.068]
4	PL to Moti	0.128	0.058	0.027	[0.009, 0.236]
5	PL to Anx	0.055	0.068	0.419	[-0.076, 0.189]
Indirect Effects					
6	Moti thru PL	-0.028	0.018		[-0.078, -0.002]
7	Anx thru PL	-0.012	0.018		[-0.064, 0.012]

Note. STDY standardization; * = Bias corrected bootstrapped CIs; Cond = condition; PL = Perceived Length; Moti = Test motivation; Anx = Test Anxiety.

Table 3

Descriptive Statistics and Correlations (N = 336)

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5
1. Condition	0.52	0.50	-				
2. Perceived length	23.93	13.63	-.11	-			
3. Actual length	34.05	17.08	-.03	.66	-		
4. Test anxiety	2.79	0.56	-.032	.07	.00	(.90)	
5. Test motivation	3.62	0.41	-.005	.14	.17	-.08	(.93)

Note. Scale coefficient alphas on diagonal; Perceived Length and Actual Length in minutes.

Hypotheses

All hypotheses were tested using maximum likelihood estimation and bias-corrected bootstrapped confidence intervals (see Table 2). Unstandardized effects are reported for Hypothesis 1 because the effect size is most interpretable without standardization: game-frame as a dummy coded experimental manipulation and perceived length as a time in minutes. Hypotheses 2, 3, 6, and 7 are criterion-standardized, because it is inappropriate to standardize a binary independent variable (i.e., condition), yet standardizing the dependent variables (i.e., continuous latent variables) is necessary for interpretation. Hypotheses 4 and 5 were also criterion-standardized because of the regression of outcomes on the perceived length variable (time in minutes).

Hypothesis 1 predicted a negative relationship between game-framing and perceived length. The data supported this prediction ($b = -3.003, p = .044$). Specifically, as condition changes from assessment-frame to game-frame, participants' perceptions of time passage is lessened by three minutes. Post-hoc analysis further supported this hypothesis by providing evidence that the difference in time-length perception were not due to actual time differences between conditions. The average completion time in the assessment frame ($M = 34.56, SD = 18.22$) and the average completion time in the game-frame ($M = 33.58, SD = 16.01$) were not significantly different $t(334) = 0.526, p = .599$. To further investigate these results, a follow up SEM analysis in Mplus was run, specifying completion time as a mediator between game-framing (condition) and perceived length. Actual and perceived time were related ($b = .527, p = .000$); however, even controlling for that effect, there was still an observed relationship between perceived time and condition ($b = -0.485, p = .025$). Thus, Hypothesis 1 appears robust to alternative explanations related to differences in actual time spent across conditions.

Hypothesis 2 predicted a positive relationship between game-framing and test motivation. The data did not support this prediction ($\beta = -0.039, p = 0.731$). In fact, there was a slight negative relationship, indicating that participants in the game-frame condition reported .039 SDs lower on the test motivation scale.

Hypothesis 3 predicted a negative relationship between game-framing and test anxiety. The results did not support this prediction; however, this may be because the effect was smaller than expected ($\beta = -0.171, p = 0.152$). The data indicate that participants in the game-frame condition reported .171 SDs lower on the test anxiety scale.

Hypothesis 4 predicted a negative relationship between perceived length and test motivation. The results did not support this prediction ($\beta = 0.128, p = 0.027$). Indeed, the data indicated a strong positive relationship between the two variables, meaning that as perceived length increased by one minute, reported test motivation increased by 0.128 SDs.

Hypothesis 5 predicted a positive relationship between perceived length and test anxiety. The data did not support this ($\beta = 0.055, p = 0.419$). For each one-minute increase in perceived length, participants only reported a .055 increase in test anxiety, below the standards for either statistical or practical significance.

Hypothesis 6 predicted that perceived length would partially mediate the relationship between game-framing and test motivation. The data supported this proposition ($\beta = -0.028$), 95% CI [-0.078, -0.002], supporting a small indirect effect of game-framing on test motivation via its intermediary effect on perceived length.

Hypothesis 7 predicted that perceived length would partially mediate the relationship between game-framing and test anxiety. The results did not support this hypothesis ($\beta = -0.012$),

95% CI [-0.064, 0.012], suggesting that the indirect effect of game-framing through anxiety is negligible in the observed sample.

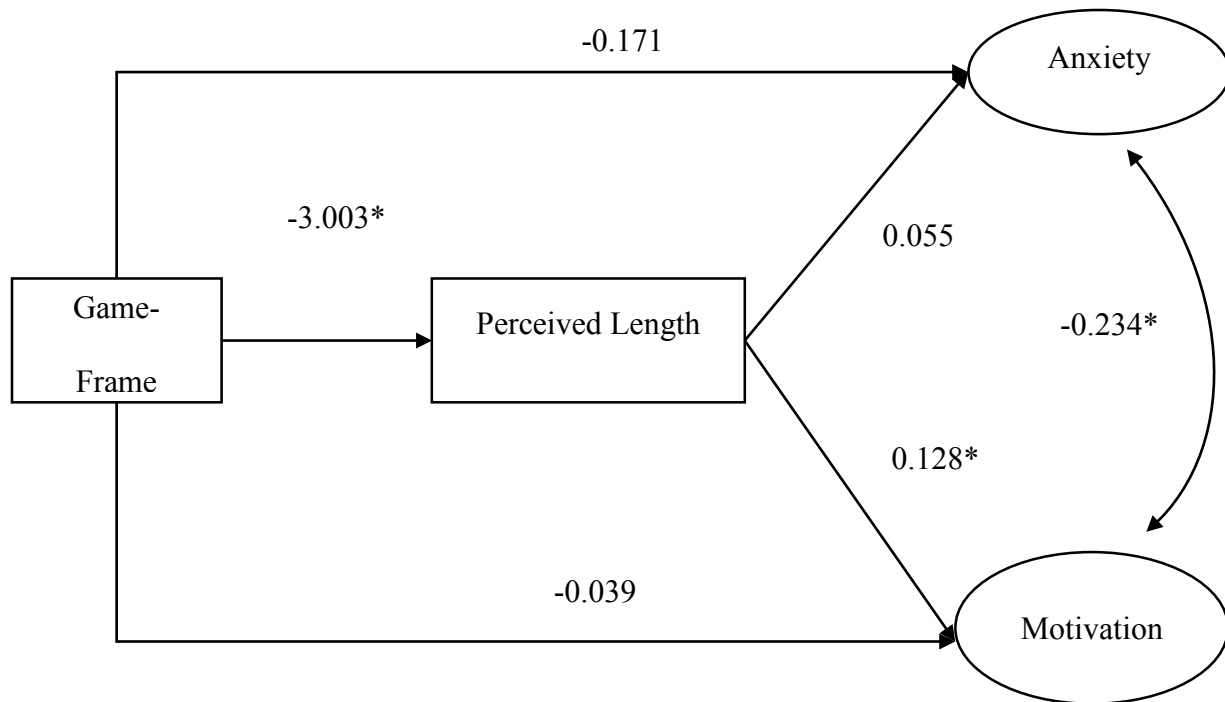


Figure 2. Conceptual model with path estimates. * $p < .05$, The covariance between Anxiety and Motivation is XY standardized, the effect of game-frame on perceived-length is unstandardized, and all other effects are criterion-standardized.

CHAPTER 4

DISCUSSION

This study makes six primary contributions. First, it theorizes and tests a direct effect of game-framing on time-length estimations of a cognitive battery, which has key implications for gamification research and applicant reactions theorists. Isolating the effects of game-framing sets a foundation for future researchers to parse out its individual causal effects, in response to current recommendations to improve the gamification literature (Seaborn and Fels, 2014). In comparison to the Lieberoth (2015) study—in which the framing intervention also involved the inclusion of game pieces and a game board—the effects observed here are modest yet meaningful. Lieberoth found a large effect for game-framing on the interest/enjoyment facet of the IMI but reported no significant differences on the other scale dimensions or the overall scale. In contrast, the present study found a relatively small effect ($r = -.11$) for framing on time perceptions. Whereas the effects in the Lieberoth study could have been caused by a game board or game pieces effect combined with the framing effect, the framing effect in the present study has been much more clearly isolated. The validity of the framing effect was further supported with a post-hoc analysis that did not find significant differences in actual completion times between groups; in short, although time perceptions changed, actual time spent did not. Additionally, insofar as applicant reactions researchers wish perceived length to remain in the applicant reactions model, there is now a demonstrated intervention that can affect time-length perceptions of cognitive ability tests.

Second, this study provides evidence contrary to the theoretically suggested direction of the relationship between time perceptions and motivation. Perceived length had a small positive relationship with test motivation, meaning that those who perceived the tests as taking more time

also reported higher levels of motivation. In contrast, Gable and Poole (2012) reported increases in approach motivation along with decreases in time perception in three different experiments. From that study, it was expected that individuals who started off highly motivated to complete the survey would consequently experience time faster. However, the data revealed the opposite relationship.

Third, this study provides the first empirical test of perceived length (PL) as a variable in the applicant reactions model, and its relationship was in the opposite direction expected. The present study represents, to my knowledge, the first empirical test of PL in the applicant reactions literature. In a meta-analytic test of the theoretical applicant reactions model, Hausknecht et al. (2004, Figure 1) list length of process as an example of a perceived procedure characteristic; however, PL was never measured or reported in any meta-analyses they identified (Hausknecht et al., 2004, Table 2). Ryan and Ployhart (2000) mentioned that applicants who perceive a selection procedure as overly time-consuming may be demotivated to perform during that procedure. Beyond these two papers, PL does not appear to have been tested in AR theory. Thus, a key contribution of the present study is the test of empirical relationships between perceived length and two applicant perceptions, test motivation and test anxiety.

Fourth, several relationships and paths theoretically indicated by applicant reactions theory were not found, demonstrating a need for further research. The relationship between PL and test anxiety ($\beta = 0.055, p = 0.419$) was statistically and practically insignificant. Simply put, the data do not indicate a meaningful relationship between perceived length of cognitive ability tests (or puzzles) and test anxiety. The relationship between PL and test motivation ($\beta = 0.128, p = 0.027$), although significant and in the range of other relationships between perceived procedure characteristics and applicant perceptions (cf. Hausknecht et al., 2004), is in the

opposite direction than that suggested by AR theory. If, as suggested by Ryan and Ployhart (2004), increases in perceived length cause decreases in motivation, this relationship should be negative. Perceived length did not mediate the relationship between game-framing and test anxiety ($\beta = -0.012$), 95% CI [-0.064, 0.012], although game-framing did have an effect on PL ($b = -3.003$, $p = .044$). Thus, several relationships contained within the AR theoretical model were not observed in the present study.

There are two potential theoretical explanations for this. First, perceived length may in fact relate to motivation and anxiety differently than expected or may have a more complex relationship with outcomes than simple linearity. For example, this relationship could be moderated by the test length, such that a longer battery with more questions could be perceived differently than the relatively short battery used in the present study. If so, the effects of actual test length on perceived temporal length and test motivation should be considered for inclusion in the AR model. Second, changes in computerized testing may have changed expectations about length in the sixteen years since Ryan and Ployhart suggested the relationship. Specifically, the use of computers for employment testing has become a norm. In contrast, the studies included in Hausknecht et al. (2004) are often from the early and mid-1990's. Some of those studies were almost certainly conducted using pencil and paper tests, and those administered on computers were administered to test-takers who likely have a fundamentally different relationship with computers in general and computerized testing in particular. It is plausible that the relationship between time perception and motivation while computing is different than otherwise.

The fifth contribution of this paper is that it revealed an indirect effect of game-framing on motivation via perceived length ($\beta = -0.028$, 95% CI [-0.078, -0.002]), supporting AR theory more broadly and setting the foundation for future studies of stronger gamification

interventions. . The point estimate of an indirect effect is difficult to interpret (Hayes, 2009), but this finding does indicate that game-framing, perceived length, and test motivation are linked, and that their relationships are consistent with the causal AR model suggested by Hausknecht et al. (2004). The direction of these relationships, however, requires further research.

The final contribution of this paper is that game-framing did not have a direct effect on either test anxiety or test motivation. Game-framing did not significantly affect test anxiety or test motivation. These results were surprising, especially in light of game-framing's apparent effect on perceived length. Specifically, since that effect was observed, it indicates that participants in the game-frame condition did indeed perceive the cognitive tests as puzzles and games. However, even though the framing effect seems to have worked, it did not affect key test-taking variables.

Limitations

Neither game-framing nor perceived length demonstrated a statistically significant relationship with test anxiety. This result does not support the greater applicant reactions model; however, these data should be interpreted with caution. The experiment may not have simulated the psychological environment of an actual selection context. In order to maximize the game-framing effect, participants were not told until after the cognitive tests were completed to think of their performance in the context of employee selection. Thus, the levels of anxiety, competitiveness, and need present during cognitive testing for selection purposes may not have been present in either condition. Furthermore, Mturk workers, who often fill out online surveys, may be more comfortable doing so than a typical job applicant. Their levels of test anxiety may be generally lower than the greater population of applicants and other test-takers. Lastly, the sample consisted mostly of full-time workers with an average job-tenure of 7.47 years. The

typical respondent in this sample may be so far removed from the application process that they did not experience anxieties more common among those having more recently entered the workforce, a common sample for prior applicant reactions studies.

Additionally, perceived length and motivation may have a non-recursive relationship such that time-perception affects motivation and motivation also affects time perceptions. This relationship is not dictated by the AR model, but is logically sound and follows from research in other areas of psychology. Levels of approach motivation may affect time perceptions, as suggested by Gable and Poole (2012), while tests that seem overly long can demotivate applicants, as suggested by Ryan and Ployhart (2000). To address this, some researchers have employed a reward system (e.g., Guillory & Hancock, 2012). A replication of the present study, in which the instructions read “top performers will be awarded a \$100 gift card” may help to recreate the psychological realism of high-stakes testing. To increase generalizability, an organization could attempt a game-framing manipulation among actual applicants. However, this may be risky given the current lack of research into other outcomes in the AR model such as offer-acceptance intentions, withdrawal intentions, and willingness to recommend to others.

The three-dimensional rotation task was exceptionally difficult for participants. Forty-six percent did not get a single item right. As such, it is possible that this task differentially affected motivation and anxiety in the present sample. Specifically, if the unusual difficulty of the task caused participants to become bored or otherwise reduce their effort across conditions, the inclusions of this test may have attenuated motivation and anxiety, reducing variance available to test hypothesized effects. This therefore represents a threat to the internal validity of the study. However, because this was only one test of several in the battery, it is suspected that any test-specific effects on the outcome variables would have been mitigated by the experience of the full

test battery. Even if the effect of the high difficulty of this test on motivation or anxiety was substantial in effect magnitude, the outcome variables as measured appeared normally distributed and with means in the expected range. Thus the overall impact of the inclusion of this variable appears minimal but should be considered carefully in future research incorporating this test battery with similar populations.

The present sample was nonhomogeneous in terms of employment. Specifically, a portion of the participants were unemployed ($N = 76$), which potentially limits generalizability to the employed population. However, a post hoc SEM analysis excluding the unemployed sample ($N = 260$), revealed no major differences in model fit or in path estimates. Thus, it seems that employment status does not affect the relationships observed in the present study.

Future Directions

The findings of this experiment indicate that other (i.e., non-selection) tasks and activities could be similarly game-framed to modify time-perceptions. Cognitive ability tests were chosen due to their importance in employee selection, but also because many of them seem like puzzles or riddles. It is unknown if the effect on PL is present for other measures that are not as naturally game-like, and this should be explored in future research. Similarly, other forms of framing should be studied in this context. For example, changes in the type font and graphical presentation are forms of gamification that could reframe perceptions of test-takers (Downes-Le Guin et al., 2012). Although it would be beneficial to study individual time-length perceptions on a variety of tests, researchers should note that once participants are asked to think about time, their perception of time is thereafter altered for the duration of the study (Wearden, 2005); given this, examining multiple methods of assessing time perception and looking for convergence might reveal additional insights.

Conclusion

The results of this study progress scholarly understanding of gamification in general and in the context selection testing. The data indicate that simply calling something a game makes it seem to take less time, highlighting the effects of game-framing by isolating the effects of a single game element. Furthermore, the experiment represents, to my knowledge, the first study to empirically evaluate time-perception in the context of AR theory. The results indicate that applicant reaction researchers may need to re-evaluate the specific relationships between perceived length of process and core applicant perceptions in the overall AR model.

A core idea in the present study is that game-elements should be isolated to parse their effects. This idea can and should be replicated in the field of gamification research. The shotgun approach, in which multiple game-elements are applied and then afterward interpreted in context of their combined effects, does not further our understanding of how to use gamification in a practical sense. Game elements such as narrative, competition, graphics, or points, should be first isolated in empirical studies (Seaborn and Fels, 2014). Once game elements' singular effects are known, they can then be combined with other single elements to consider interactive effects. It is only through studies like these that gamification will one day be understood well enough for widespread application in a variety of contexts.

REFERENCES

- Anderson, N., Salgado, J. F., & Hülshager, U. R. (2010). Applicant reactions in selection: Comprehensive meta-analysis into reaction generalization versus situational specificity. *International Journal of Selection and Assessment*, *18*, 291-304. doi: 10.1111/j.1468-2389.2010.00512.x
- Armstrong, M. B., Landers, R. N., & Collmus, A. B. (2015). Gamifying recruitment, selection, training, and performance management: Game-thinking in human resource management. In D. Davis & H. Gangadharbatla (Eds.), *Handbook of Research Trends in Gamification*: IGI Global.
- Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology*, *43*.
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*, 431-449. doi: 10.1037/a0038047
- Buford, C. C., & O'Leary, B. J. (2015). Assessment of fluid intelligence utilizing a computer simulated game. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, *7*(4), 1-17.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel Selection in Organizations* (pp. 35-70). San Francisco, CA: Jossey Bass.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159. doi: 10.1037/0033-2909.112.1.155

- Collmus, A. B., Armstrong, M. B., & Landers, R. N. (2016). Game-thinking within social media to recruit and select job candidates. In R. N. Landers & G. B. Schmidt (Eds.), *Social Media in Employee Selection and Recruitment* (pp. 103-124). Switzerland: Springer International Publishing.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.
- Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence, 43*, 52-64. doi: 10.1016/j.intell.2014.01.004
- Delgado, M. T., Uribe, P. A., Alonso, A. A., & Diaz, R. R. (2016). TENI: A comprehensive battery for cognitive assessment based on games and technology. *Child Neuropsychology, 22*, 276-291. doi: 10.1080/09297049.2014.977241
- DeSimone, J. A., Harms, P. D., & DeSimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior, 36*, 171-181. doi: 10.1002/job.1962
- Downes-Le Guin, T., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research, 54*, 613-633. doi: 10.2501/IJMR-54-5-613-633
- Gable, P. A., & Poole, B. D. (2012). Time flies when you're having approach-motivated fun: Effects of motivational intensity on time perception. *Psychological Science, 23*, 879-886. doi: 10.1177/0956797611435817
- Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*, 441-467. doi: 10.1177/1046878102238607

- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *The Academy of Management Review*, *18*, 694-734. doi: 10.2307/258595
- Gottfredson, L. S. (1997). Why g matters: the complexity of everyday life. *Intelligence*, *24*, 79-132.
- Guillory, J., & Hancock, J. T. (2012). The effect of LinkedIn on deception in resumes. *Cyberpsychology, Behavior, and Social Networking*, *15*(3), 135-140.
- 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. doi: 10.1111/j.1744-6570.2004.00003.x
- Hayes, A. F. (2012) PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from <http://www.afhayes.com/public/process2012.pdf>
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of educational research*, *58*(1), 47-77.
- Hicks, R. E. (1992). *Prospective and retrospective judgments of time: A neurobehavioral analysis* (pp. 97-108). Springer Netherlands.
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*, 340-362. doi: [http://dx.doi.org/10.1016/0001-8791\(86\)90013-8](http://dx.doi.org/10.1016/0001-8791(86)90013-8)
- IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.
- ICAR: The International Cognitive Ability Resource Team (2014) <http://icar-project.com>

- 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. doi: 10.1037/0022-3514.86.1.148
- Landers, R. N., Bauer, K. N., & Callan, R. C. (2015). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*. doi: 10.1016/j.chb.2015.08.008
- Levin, I., & Gaeth, G. J. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research*, *15*, 374-378.
- Levin, I., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*, 149-188.
- Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture*, *10*, 229-248. doi: 10.1177/1555412014559978
- 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*(1), 67-77.
- Mavletova, A. (2014). Web surveys among children and adolescents: Is there a gamification effect? *Social Science Computer Review*. doi: 10.1177/0894439314545316
- McLachlan, G. J. (1999). Mahalanobis distance. *Resonance*, *4*(6), 20-26.
- McPherson, J., & Burns, N. R. (2008). Assessing the validity of computer-game-like tests of processing speed and working memory. *Behavioral Research Methods*, *40*, 969-981. doi: 10.3758/BRM.40.4.969

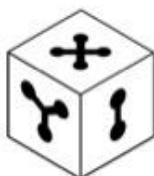
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*, 437-455. doi: 10.1037/a0028085
- Muthén, L.K. and Muthén, B.O. (1998-2015). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén
- Ryan, A. M., & Huth, M. (2008). Not much more than platitudes? A critical look at the utility of applicant reactions research. *Human Resource Management Review, 18*, 119-132. doi: 10.1016/j.hrmr.2008.07.004
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26*, 565-606.
- Ryan, A. M., Sacco, J. M., McFarland, L. A., & Kriska, S. D. (2000). Applicant self-selection: Correlates of withdrawal from a multiple hurdle process. *Journal of Applied Psychology, 85*, 163-179. doi: 10.1037/0021-9010.85.2.163
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78. doi: 10.1037/0003-066x.55.1.68
- Sackett, A. M., Meyvis, T., Nelson, L. D., Converse, B. A., & Sackett, A. L. (2010). You're having fun when time flies: The hedonic consequences of subjective time progression. *Psychological Science, 21*, 111-117. doi: 10.1177/0956797609354832
- 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. (2004). General mental ability in the world of work: occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162-173. doi: 10.1037/0022-3514.86.1.162
- Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14-31.
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, *46*, 49-76. doi: 10.1111/j.1744-6570.1993.tb00867.x
- Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. New Brunswick, NJ: Transaction Books.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Sucala, M., Scheckner, B., & David, D. (2011). Psychological time: Interval length judgments and subjective passage of time judgments. *Current psychology letters. Behaviour, brain & cognition*, *26*(2, 2010).
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.
- Wearden, J. H. (2005). The wrong tree: Time perception and time experience in the elderly. *Measuring the mind: Speed, age, and control*, 137-158.
- Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., . . . Conkey, C. (2008). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, *40*, 217-266. doi: 10.1177/1046878108321866

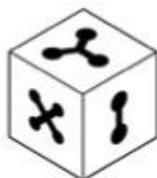
APPENDICES

APPENDIX A

3D ROTATION SAMPLE ITEM



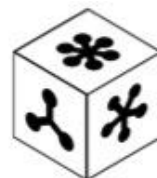
X



A



B



C

None of
the cubes
could be a
rotation.

D



E



F



G

I do not
know
the
solution.

H

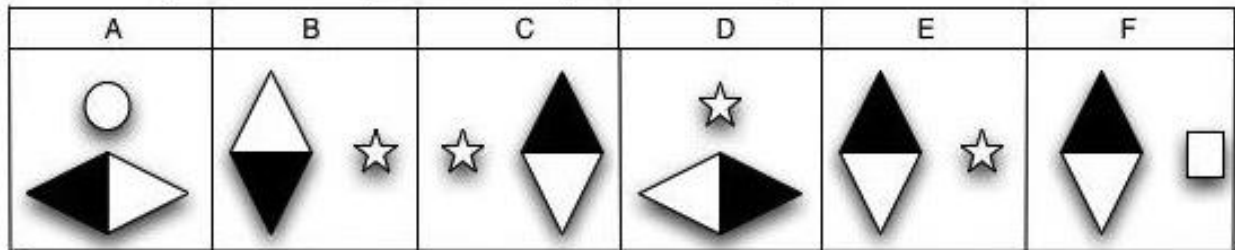
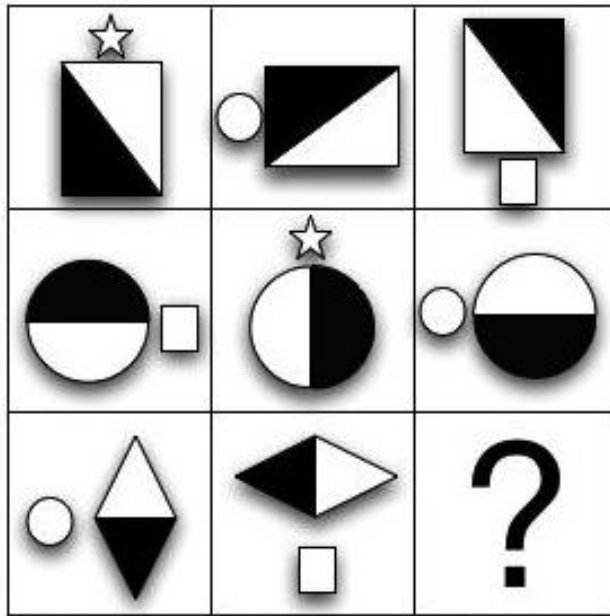
APPENDIX B**LETTER AND NUMBER SERIES SAMPLE ITEM****LN.01 (q_12001)**

In the following number series, what number comes next? 64, 81, 100, 121, 144, ...

(1) 154 (2) 156 (3) 162 (4) 169 (5) 178 (6) 196 (7) None of these (8) I don't know

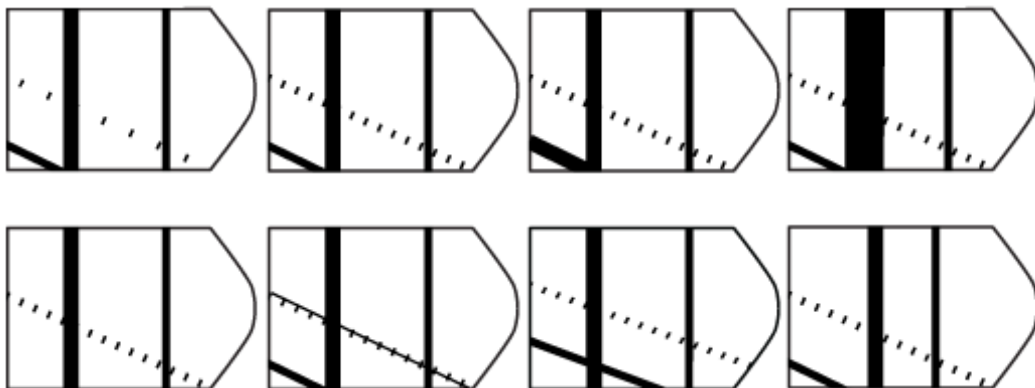
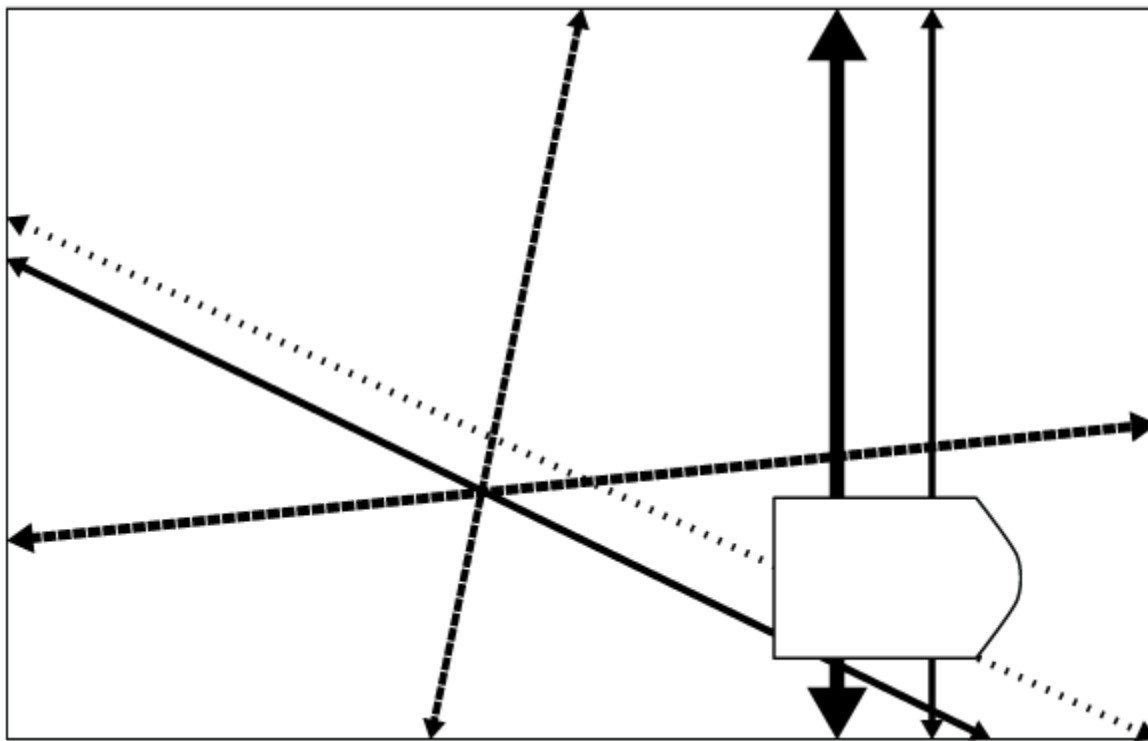
APPENDIX C

MATRIX REASONING SAMPLE ITEM



APPENDIX D

PROGRESSIVE MATRICES SAMPLE ITEM



APPENDIX E**VERBAL REASONING SAMPLE ITEM****VR.31 (q_12031)**

Isaac is shorter than George and Phillip is taller than George. Which of the following statements is most accurate?

- (1) Phillip is taller than Isaac (2) Phillip is shorter than Isaac (3) Phillip is as tall as Isaac (4) It is impossible to tell (5) Isaac is taller than George (6) George is taller than Phillip (7) None of these (8) I don't know

APPENDIX F**TEST ATTITUDES SURVEY****1. Motivation**

- a. Doing well on this test (or these tests) is important to me.
- b. I wanted to do well on this test or tests.
- c. I tried my best on this test or tests.
- d. I tried to do the very best I could to on this test or tests.
- e. While taking this test or tests, I concentrated and tried to do well.
- f. I want to be among the top scorers on this test (or these tests).
- g. I pushed myself to work hard on this test or these tests.
- h. I was extremely motivated to do well on this test or tests.
- i. * I just didn't care how I did on this test or tests.
- j. * I didn't put much effort into this test or tests.

2. Lack of Concentration

- a. It was hard to keep my mind on this test or tests.
- b. I found myself losing interest and not paying attention to the test.
- c. During the test session, I was bored.
- d. I get distracted when taking tests of this type.

3. Belief in Tests

- a. * This test or tests was a good reflection of what a person could do in the job.
- b. * Tests are a good way of selecting people into jobs.
- c. This kind of test or tests should be eliminated.
- d. I don't believe that tests are valid.

4. Comparative Anxiety

- a. I probably didn't do as well as most of the other people who took these tests.
- b. I am not good at taking tests.
- c. During the testing, I often thought about how poorly I was doing.
- d. I usually get very anxious about taking tests.
- e. * I usually do pretty well on tests.
- f. * I expect to be among the people who score really well on this test.
- g. My test scores don't usually reflect my true abilities.
- h. I very much dislike taking tests of this type.
- i. During the test or tests, I found myself thinking of the consequences of failing.
- j. During the testing, I got so nervous I couldn't do as well as I should have.

5. Test Ease

- a. This test was (or these tests were) too easy for me.
- b. I found this test or tests too simple.
- c. * I found this test or tests interesting and challenging.
- d. * I felt frustrated because many of the test questions were too difficult.

6. External Attribution

- a. I became fatigued and tired during the testing.
- b. The questions on this test or tests were ambiguous and unclear.
- c. I have not been feeling well lately and this affected my performance on the test or tests.

- d. While taking the test or tests, I was preoccupied with how much time I had left.
 - e. I felt a lot of time pressure when taking this test or tests.
- 7. General Need Achievement**
- a. Once I undertake a task, I usually push myself to my limits.
 - b. I try to do well in everything I undertake.
 - c. * In general, I like to work just hard enough to get by.
- 8. Future Effects**
- a. * My performance on this test will not affect my chances for obtaining a job or gaining a promotion.
 - b. Scores from this test or tests will probably affect my future.
 - c. These test scores will be used in future decisions made about me.
- 9. Preparation**
- a. I spent a good deal of time preparing for this test or tests.
 - b. I prepared a lot for this test or tests.

VITA

Andrew B. Collmus
 Old Dominion University
 Department of Psychology
 Mills Godwin Building, Room 226
 Norfolk, VA 23529

Education

Bachelor of Science, *magna cum laude*, Psychology, Colorado State University, May 2014

Thesis: “Swift Trust in Virtual Worlds: Building Team Trust with a Videogame”

Publications

- Armstrong, M. B., Ferrell, J. Z., **Collmus, A. B.**, & Landers, R. N. (2016). Correcting misconceptions about gamification of assessment: More than SJTs and Badges. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 9(3), 671-677.
- Landers, R. N., Armstrong, M. B., & **Collmus, A. B.** (2016). How to use game elements to enhance learning: Applications of the theory of gamified learning. In M. E. Ma & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications (Vol. 2)*. Springer.
- Landers, R. N., Brusso, R. C., Cavanaugh, K. J. & **Collmus, A. B.** (2016). A primer on theory-driven web scraping: Automatic extraction of big data from the internet for use in psychological research. *Psychological Methods*.
- Landers, R. N., Fink, A. & **Collmus, A. B.** (*forthcoming*). Big data and other technology-enhanced quantitative methods for selection. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of Employee Selection*. New York, NY: Routledge.
- Collmus, A. B.**, Armstrong, M. B., & Landers, R. N. (2016). Game-thinking within social media to recruit and select job candidates. In R. N. Landers & G. B. Schmidt (Eds.), *Social Media in Employee Selection and Recruitment*. Springer.
- Armstrong, M. B., Landers, R. N., & **Collmus, A. B.** (2015). Gamifying recruitment, selection, training, and performance management: Game-thinking in human resource management. In D. Davis & H. Gangadharbatla (Eds.), *Handbook of Research on Trends in Gamification*. Hershey, PA: Information Science Reference

Presentations

- Cavanaugh, K. J., Brusso, R. C., **Collmus, A. B.** & Landers, R. N. (2016, April). *Web scraping: Automatic extraction of big data for I/O psychology*. Poster presented at the 31st Annual Conference of the Society for Industrial and Organizational Psychology, Anaheim, CA.
Note: This paper placed as a Featured Top Rated Poster at the 2016 SIOP Annual Conference and was a finalist for the John Flanagan Award.
- Collmus, A. B.**, Landers, R. N. (2015). Game narrative in personality assessment: Development of a scale. Presentation at Virginia Psychological Association’s Spring Convention & Educational Conference, Virginia Beach, VA.
- Armstrong, M. B., Landers, R. N., **Collmus, A. B.** (2015). Game thinking in human resource management. Poster presentation at the 30th Annual Conference of the Society for Industrial and Organizational Psychology, Philadelphia, PA.
- Collmus, A. B.**, Kraiger, K. (2015). Swift trust in virtual worlds: Building trust with a videogame. Poster presentation at the 27th Annual Convention of the Association for Psychological Science, New York, NY.
- Turner, S. J., **Collmus, A.**, Lam, R., & Rhodes, M. G. (2013). Using value to examine the own-race bias: Evidence for the social-categorization model. Poster presented at the 83rd annual meeting of the Rocky Mountain Psychological Association, Denver, CO.