5-1-2019

Assessing Factors of Physical Risk-Taking in a Novel Behavioral Measure

Edward A. Smith

ASSESSING FACTORS OF PHYSICAL RISK-TAKING IN A NOVEL BEHAVIORAL MEASURE

By

Edward A. Smith

Bachelor of Arts - Psychology
University of Notre Dame
2017

A thesis submitted in partial fulfillment
of the requirements for the

Master of Arts - Psychology

Department of Psychology
College of Liberal Arts
The Graduate College

University of Nevada, Las Vegas
May 2019

Thesis Approval

The Graduate College
The University of Nevada, Las Vegas

April 12, 2019

This thesis prepared by

Edward A. Smith

entitled

Assessing Factors of Physical Risk-Taking in a Novel Behavioral Measure

is approved in partial fulfillment of the requirements for the degree of

Master of Arts - Psychology
Department of Psychology

Stephen Benning, Ph.D.
Examination Committee Chair

Andrew Freeman, Ph.D.
Examination Committee Member

Kimberly Barchard, Ph.D.
Examination Committee Member

Peter Gray, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Dean

Abstract

Risk taking is a complex heterogeneous construct that has proven difficult to assess, especially when using behavioral tasks. A new measure, the Assessment of Physical Risk Taking (APRT) is presented as a comprehensive assessment of several factors of risk. Specifically, the measure seeks to examine the effects of several decision-making elements (e.g., probability of success and failure, magnitude of reward and punishment) of different types of physically risky behaviors and produce a variety of different outcome scores. Participants (*N* = 256) completed APRT in a laboratory setting, with half being assigned to an enforced Delay condition. Main effects, two-way interactions among five within-subject factors, and interactions between the within-subject factors and Delay were estimated for four APRT outcome scores using Generalized Estimating Equations. Results indicated that Injury Magnitude and Injury Probability exerted much stronger effects than any of the other independent variables. The implications of these results are discussed in the context of the future of behavioral risk-taking tasks and studies

Table of Contents

Introduction

The concept of risk taking is a broad and often quite heterogenous variable used in many areas of research. Because of its relationship to dangerous or otherwise problematic behaviors, many researchers seek to accurately assess risk-taking as a means of both predicting these future behaviors, as well as understanding how risk-taking interacts with other psychosocial variables and outcomes, such as psychopathology or personality traits. Non-academics also could benefit greatly from these kinds of measures as well. For example, healthcare providers may be interested in an individual's risk-taking propensity when recommending treatments or preventative care measures. Drug rehabilitation centers could more easily identify individuals at the highest risk for relapse with an adequate measure of risk taking, and law enforcement or government officials might more easily target certain individuals or population groups that could benefit most from crime prevention programs. To create such measures, however, risk-taking needs to be clearly defined and subsequently operationalized.

Definitions of risk-taking have varied, but a common theme in the literature seems to be the idea of engaging in behaviors or decisions that involve multiple possible outcomes, some of which are negative (Slovic, 1964). This broad definition is unfortunately vague and can lead to a variety of interpretations. For some, risk is only measured as a function of the magnitude and/or probability of these negative outcomes, with the type of risk being included occasionally as a moderating factor as well (Gardoni & Murphy, 2014). Indeed, even most dictionary definitions of risk focus exclusively on potential negative consequences (Carmichael, 2016). Consider the following two examples: both playing poker and drunk driving would fit the definition given by Slovic, as both activities involve making decisions that have potentially negative outcomes (e.g., losing money, crashing and/or being arrested), yet most people would likely judge drunk driving as being a more risky decision to make, and thus both scenarios also fit the conceptualization presented by Gardoni and Murphy (as the type of behavior alters the magnitude of the potential consequences).

This consequence-centered conceptualization seems incomplete, however, as the likelihood and size of rewards seem to be important factors in risky decision-making as well. For example, experienced gamblers understand when they have a good chance of winning a round of their game of choice and can alter their betting based on their perceived chance of success (Spurrier & Blaszczynski, 2014). Similarly, the likelihood of being arrested or in an accident due to drunk driving increases with the number of alcoholic drinks imbibed, a fact of which convicted drunk drivers tend to admit to have had some awareness, resulting in fewer accidents/arrests for individuals who had consumed a very high number of drinks, as they are aware of their increased impairment and choose to drive less often than individuals who have had a medium number of drinks (Phelps, 1987). Individuals who are high in sensation seeking, impulsivity, and other risk-related personality constructs also tend to overestimate the extent of potential rewards of risky behaviors (Horvath & Zuckerman, 1993; Ryb, Dischinger, Kufera, & Read, 2006). Thus, the likelihood of *both* success and failure, as well as the magnitude of each, all seem to play a role in the decision to engage in risky behaviors.

These four factors (the separate probabilities of injury and reward, and the separate magnitude of each) are generally considered to be the key components that affect risky decision making (though the magnitude and probability of reward are ignored more often than not), and each are associated with different (though related) neurological and cognitive processes (Ernst & Paulus, 2005; Smith et al., 2009). Despite the prevalence of this knowledge in the neuroscientific literature, little is done to operationalize these factors within behavioral assessments of risk-taking in psychological studies. Indeed, while current behavioral tasks correlate with real-world risky behaviors broadly, they have not been decomposed into these more fundamental factors that influence the decision to engage in such behaviors, making study of the underlying cognitive or neurological processes of risky decision making difficult, if not impossible (Schonberg, Fox, & Poldrack, 2011). There is thus a strong need for better, more complex behavioral tasks that can both predict real risky behaviors and precisely examine the effects of these underlying factors on risky decision making.

Finally, while these four factors may underlie risky decision making, one must also consider how these factors interact with different risky behaviors, as the strength and extent of the effects of these factors on risk-taking may vary between specific risky behaviors. For example, the magnitude of reward is likely to play a larger role in behaviors such as gambling, where individuals are tempted to risk losing more money for the chance to win a large sum. In contrast, an intoxicated individual may be more concerned with the likelihood of being caught or crashing their vehicle (i.e., injury probability) if they were to attempt to drive. Thus, new behavioral tasks should also cover conceptually distinct types of risky behavior in order to better illustrate how the effect sizes of these factors tend to vary. Therefore, a good behavioral task should examine a total of five factors that may be impacting scores: the effects of the four underlying influences on decision-making, plus the effects of the specific risky behavior being examined.

The subsequent review of the various existing risk-taking measures is focused on briefly examining how currently available measures fall short in their ability to examine these more precise factors, as well as other weaknesses related to the assessment of risk-taking. The details of the current study are then discussed. The study primarily sought to establish a new behavioral measure that is centered around these four factors and covers multiple examples of risky behaviors.

Literature Review

Self-Report Measures of Risk Taking

Before examining behavioral measures of risk-taking, a brief review of self-report measures is necessary, as they are more frequently used in studies and are often the metric (in addition to actual risk-taking behavior) used to evaluate the construct validity of behavioral tasks. The two selected for review here are among the most commonly cited and therefore provide a fair approximation of how most researchers conceptualize and assess risk-taking.

One of the most popular self-report risk-taking measures is the Domain Specific Risk-Taking Scale (DOSPERT), which assesses reported levels of risk taking in 5 domain areas: ethical, financial, health and safety, social, and recreational (Weber, Blais, & Betz, 2002). In addition to the usual assessment of actual engagement in risk behaviors, DOSPERT also examines participants' perceived-risk attitudes in each of these domains, which the authors defined as a participant's willingness to engage in a given activity as a function of the perceived level of risk in that activity.

DOSPERT has good internal consistency and moderate test-retest reliability, and its factor structure has been replicated in a variety of settings and populations (Blais & Weber, 2006; Weber et al., 2002). The construct validity of DOSPERT risk domains is also high (Blais & Weber, 2006; Hanoch, Johnson, & Wilke, 2006). Not only do scale scores correlate to relevant past and present risk behaviors (i.e., someone with high scores on the financial scale has indeed made frequent risky financial decisions in the past or present), but the scale scores do not automatically generalize across domains (e.g., a high score on the recreational scale does not necessarily predict high scores on the financial scale). Additionally, because of its ability to assess multiple types of everyday risk taking and its dual focus on both perceived risk attitudes and actual risk behaviors, DOSPERT has been commended in a systematic review as one of the most relevant and useful risk taking scales for use in clinical environments (Harrison, Young, Butow, Salkeld, & Solomon, 2005).

DOSPERT also seems to assess a few of the five factors of risky decision-making as well. Unlike most risk-taking measures (self-report and behavioral), DOSPERT does assess multiple types of risky behavior as well as risk perception. However, while the risk-perception scales seem to assess the injury magnitude element of risky decision making, DOSPERT does not assess any of the other factors (i.e., injury probability, reward magnitude, reward probability) in its scales. Thus, while this is not necessarily an issue with DOSPERT itself (as self-report measures generally measure a specific construct or constructs, which is in this case are risk-taking and risk perception, and not the factors that influence them), the use of DOSPERT as a gold standard to compare against new behavioral tasks may be inappropriate if such tasks do indeed decompose risk-taking into a function of these factors, as there is only overlap between a few of these factors and the DOSPERT scales.

Despite the rise of more contemporary measures, the Sensation Seeking Scale (SSS; Zuckerman, Eysenck, & Eysenck, 1978) has also remained a popular self-report measure for researchers studying risk taking. While not technically a measure of risk-taking itself, SSS examines the personality construct of sensation seeking, which is related to risk taking by its own definition: “A trait defined by the seeking of varied, novel, complex, and intense sensations and experiences, and the willingness to take physical, social, legal, and financial risks for the sake of such experience” (Deditius-Island & Caruso, 2002; Zuckerman, 1994). The fifth version of SSS (SSS-V) breaks down this construct into four factor scales (Zuckerman et al., 1978): thrill and adventure seeking (whose items reflect a desire to engage in high risk sports like skydiving or mountain climbing), experience seeking (whose items relate to a desire for novel sensory or mental experiences), disinhibition (whose items express the respondent's desire to engage in commonly studied risk behaviors such as drug use and reckless sex), and boredom susceptibility (which assesses aversion to repetitious or routine behaviors/lifestyles).

Despite its popularity over the years, the psychometric data on SSS-V is mixed. A systematic review of each factor scale's reliability found that the boredom susceptibility scale consistently failed to meet reliability standards (as assessed by Cronbach's α) across hundreds of studies, while the other three scales had borderline (though still acceptable) reliabilities on average (Deditius-Island & Caruso, 2002). Despite this weakness, however, numerous studies support the overall construct validity, convergent validity, and factor structure of SSS-V (Roberti, Storch, & Bravata, 2003; Zuckerman, 1994). The largest criticism of the measure is instead usually its use of antiquated jargon in many of its items (Arnett, 1994). For example, few young adults today will understand what the term "jet set" means. While the author has acknowledged this issue and even published revised versions of some items (Zuckerman, 1996, 2007), many contest that further edits are still needed to maintain the scale's validity in contemporary use. Specifically, some have suggested that the items that involve activities that many no longer consider particularly risky or exotic, such as marijuana use or non-traditional sexual practices, require updating (Gray & Wilson, 2007).

Even if the SSS-V did not have the reliability or content/wording problems it seems to have, the measure would still be an inappropriate metric against which to compare a behavioral task when examining such a task's construct validity. Unlike DOSPERT, which measures actual risky behaviors and risk-perception, SSS-V primarily measures sensation seeking, a personality construct that, while perhaps conceptually related to risk-taking, is certainly not equivalent to it.

While personality researchers may find this measure to be of interest, it would be inappropriate to use this measure as a self-report measure of risk-taking, yet researchers have continually used SSS-V to test the convergent validity of behavioral risk-taking tasks (Lejuez et al., 2002; Wardle, Gonzalez, Bechara, & Martin-Thormeyer, 2010). This issue likely has led to erroneous conclusions about the construct validity of existing behavioral measures, which may help explain why behavioral tasks tend to have lower correlations than expected with better risk-taking measures like DOSPERT (Buelow & Blaine, 2015; Hopko et al., 2006; Upton, Bishara, Ahn, & Stout, 2011). While addressing this issue was beyond the scope of the current study, this problem underscores the need for higher quality measures of risk-taking in general, both in the self-report and behavioral domains.

Behavioral Measures of Risk

Behavioral tasks to measure risk-taking are meant to be an alternate form of assessment, in that they enable researchers to observe a standardized example of an individual's actual risk-taking behavior instead of having to rely on retrospective self-report. One of the most popular behavioral measures is the Balloon Analogue Risk Task (BART), a computerized task in which participants earn imaginary money by continually inflating a virtual balloon (Lejuez et al., 2002). Participants may stop inflating a balloon at any time to "bank" their current earnings and start a new balloon trial. Successive pumps increase the risk of the balloon exploding, however, in which case the participant loses any earnings from that trial.

The original authors of the task found several significant correlations between BART risk-taking scores and other measures of risk-related constructs. These included self-report measures of behavioral constraint (which had a negative association with BART risk taking), impulsivity (positive), and sensation seeking (positive). The original researchers also found that BART risk-taking scores had positive associations with substance use and delinquent behavior in the past year (Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Lejuez et al., 2002). BART scores were able to differentiate between smokers and non-smokers, as well as cocaine users/non-users (Bornovalova, Daughters, Hernandez, Richards, & Lejuez, 2005; Lejuez, Aklin, Jones, et al., 2003). These findings replicated in adolescent and clinical samples (Hopko et al., 2006; Lejuez, Aklin, Zvolensky, & Pedulla, 2003). The test-retest reliability of BART scores was also shown to be acceptable, $r(37) = .77$, $p < .001$ (White, Lejuez, & de Wit, 2008). Due to this initial strong body of evidence, BART became very popular, as it apparently lacked many of the weaknesses found in previous behavioral tasks designed to assess risk taking, as evidenced by its significant correlations with self-report measures of risk taking (Lejuez et al., 2002). It thus even began to be used in subfields outside of personality or risk assessment, such as in cognitive neuroscience in conjunction with fMRI (Rao, Korczykowski, Pluta, Hoang, & Detre, 2008).

More recent evidence from other researchers, however, have called many of these findings into question. A recent meta-analysis indicated that the overall effect size for the relationship of BART scores with self-report measures of risk-taking across nearly two dozen studies was rather small ($r = .14$; (Lauriola, Panno, Levin, & Lejuez, 2014). Even in subfields such as personality, this correlation size is considered unimpressive and hardly noteworthy. Thus, the current evidence suggests that the relationship of BART to self-report measures of risk-taking may not be as impressive as was originally thought.

Newer findings have also criticized the BART's supposed relationship to real-world risky behaviors. Studies conducted without the involvement of the BART authors found that there was no significant relationship between BART scores and smoking or drinking, and that there were no differences in BART scores between users and non-users of tobacco or alcohol. Furthermore, the BART has been shown to have no significant association with other behavioral measures of risk taking (Buelow & Blaine, 2015). This last point highlights a persistent issue with different behavioral risk-taking measures, in that they appear to be measuring differing risk-related constructs, instead of acting as comprehensive measures of the elements of risk-taking (Aklin et al., 2005).

Regarding how well BART manipulates the elements of risky decision making, the task certainly manipulates the probability of injury as part of the task (as the likelihood of the balloon popping increases with each pump, and the starting probability varies between trials in most versions of the test), but it does nothing to manipulate any of the other factors that influence risky decision making. The amount of "money" earned with each pump does not vary within or across trials, and the probability of earning money on a pump is simply the inverse of the probability of the balloon popping (i.e., there is no neutral outcome where money is neither earned or lost, thus making the probabilities of reward and injury two ways of looking at the same variable). One might argue that the magnitude of reward does vary, as it increases with each successive pump in a trial, but this makes it manipulated by the participant rather than the experimenter, and thus is not truly an independent variable. Furthermore, the injury probability condition is almost never even utilized in analyses of BART, with scores being typically reported as average earnings across trials or average pumps per trial, rather than being broken down into scores organized by the different levels of the initial popping probability (Lauriola et al., 2014; Lejuez et al., 2002). While this limited form of analysis may be sufficient for personality or clinical studies, BART in its current form is difficult to utilize effectively in studies examining cognitive processes or brain activity, studies which need a means of parsing scores along the lines of these other elements of risky decision making (Schonberg et al., 2011).

Another popular behavioral risk-taking task is the Iowa Gambling Task (IGT; Antoine Bechara, Damasio, Damasio, & Anderson, 1994). The task was originally intended to examine decision-making deficits in brain-damaged populations, but has since seen use in a variety of research subfields, including personality and risk-taking (Xu, Korczykowski, Zhu, & Rao, 2013). In the IGT, participants select a total of 100 cards by choosing from 4 decks, with no more than 60 cards being allowed to come from the same deck. Each card gives or takes away a certain amount of "money" to the participant. The decks are arranged such that two give an overall net positive gain (advantageous) but lower individual "win" values, while the other two give an overall net loss (disadvantageous) but with larger individual "win" values (Bechara, 2007). Scores for decision making are calculated by subtracting the number of choices from disadvantageous decks from the number of choices made from advantageous decks. Healthy individuals generally are able to figure out that the advantageous decks are better in the long run, while neurologically impaired individuals typically are unable to recognize this over the course of the task (Antoine Bechara, Tranel, & Damasio, 2000).

Despite there being evidence for the construct validity of IGT's original purpose (assessing decision-making deficits in clinical populations), there are a number of issues regarding its utility for other purposes (Buelow & Suhr, 2009). For example, while IGT performance has been shown to be impacted negatively by cocaine and marijuana use (compared to abstinent controls), there were significant learning effects that occur in all individuals, regardless of substance use (Verdejo-Garcia et al., 2007). While these effects are expected, if not part of the task design itself, participants have demonstrated the ability to become consciously aware of these effects after only one completion of the task and can explicitly describe that some decks are better than others in the long run, thereby making retesting far less valid than the initial run (Bechara, Damasio, Tranel, & Damasio, 2005). Additionally, healthy individuals also typically show this reversal learning effect midway through the task, at which point they recognize that the disadvantageous decks are negatively impacting their overall performance (despite their larger "win" amounts) and consciously switch to the net positive, but smaller rewarding, advantageous decks (Clark, Cools, & Robbins, 2004). This learning effect (or lack thereof, in the case of neurologically impaired individuals) has thus been hypothesized to be the cause of group differences in IGT performance, as opposed to differences in risk-taking propensity (Maia & McClelland, 2004).

Some have used this idea to argue that the IGT does not actually assess risk-taking at all, instead assessing only an individual's learning abilities, and by extension how different variables (e.g., substance use, neurological functioning, personality) affect this ability (Buelow & Suhr, 2009). Indeed, newer evidence has suggested that IGT performance might be only partially indicative of propensity for risk taking, and even then only in individuals who are both low in trait impulsivity and are able to show the reversal learning effect midway through the task (Upton et al., 2011). This directly contradicts previous theories about how the IGT measures risk-taking, as it was thought that individuals more prone to risky behavior and impulsiveness would fail to show the learning effect and continue drawing from the disadvantageous decks (A. Bechara et al., 2005).

Like the BART, IGT does not seem to correlate strongly (if at all) with other measures of risk taking (Buelow & Suhr, 2009; Lejuez, Aklin, Jones, et al., 2003; Upton et al., 2011). This supports the notion that these tasks are likely examining different aspects of decision making

(Buelow & Blaine, 2015). With respect to its manipulation of the various elements of risk, the four decks in the task actually do have varying magnitudes of both rewards and penalties, as well as differing reward and penalty probabilities (Bechara, 2007). However, as with BART, the effects of each of these factors are rarely (if ever) part of analyses of task performance. Examining each of their effects individually would be difficult anyway, as the task is essentially one very long trial, with the variations in each of these factors being controlled by the participant (via the deck they choose for each of their 100 selections) and not the experimenter directly. Thus, IGT is difficult to use in cognitive neuroscience studies that wish to examine the cognitive processes underlying risky behavior. To be fair to the authors, this was not the intended purpose of the task, and thus this limitation does not make IGT a bad measure of risk-taking per se. As with BART, the task is simply unsuitable for studies trying to examine the precise cognitive constructs influencing risky behavior (Schonberg et al., 2011).

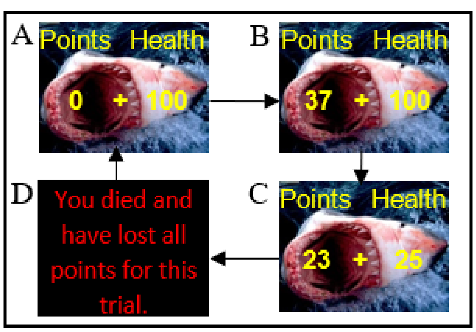
Few other behavioral tasks seem to explicitly assess the underlying elements of risk-taking. Some newer tasks, such as the Columbia Card Task (CCT), show promise, but their validity and relationship to self-report measures of risk taking and other relevant variables is not yet established, and their relationship to BART and IGT has been shown to be non-significant (Buelow & Blaine, 2015; Figner, Mackinlay, Wilkening, & Weber, 2009). There is consequently a need in the field of cognitive neuroscience for specialized behavioral tasks that not only assess the different cognitive constructs that comprise risky decision making, but that also show appropriate correlations with self-report measures of risk-related constructs, as well as real-world risky behaviors. Such tasks should also assess a variety of risk-taking behaviors, as using only a single type of task inhibits the generalizability of any effects of these factors on risky behavior.

The present study attempted to provide a bridge to this gap in the literature by introducing a new, more specialized type of behavioral task that incorporates imagery of actual physically risky activities, as opposed to a novel game that does not have a real-world equivalent (such as selecting cards or pumping up balloons for money). While no measure of risk-taking can truly simulate most risk-taking behaviors, especially those that risk physical injury, the hope was that at least reflecting multiple real-world behaviors in a task would more accurately and reliably predict real-world risk-taking, as well as correlate better with other measures of risk-taking and risk-related constructs.

Assessment of Physical Risk Taking (APRT)

APRT is a computerized task in which participants complete 64 unique trials. Each trial involves the participant viewing a picture that depicts a first-person view of a physically dangerous activity from one of four categories (encountering a dangerous animal, attempting to photograph natural disasters, attempting to help someone in physical danger, and standing on the edge of a high ledge or cliff). While these are not exactly everyday behaviors, they are still intended to represent real-world activities in which an individual high in physical risk-taking propensity may engage when the opportunity presents itself. At the start of the trial, the picture is accompanied by text along the top of the screen describing what the participant is "doing" (e.g., "You are attempting to photograph natural disasters" would accompany a picture of a tornado). The screen also shows the participant's current health and points scores for that trial; each trial begins at 100 health and 0 points.

During each trial, participants choose whether or not to proceed with the displayed activity to try and earn points. Choosing to proceed (by pressing the "p" key on a standard keyboard) results in either gaining points and a reward message that replaces the initial activity description (“Lucky You – Points Awarded!”), gaining no points but also not losing health (“Sorry, no points awarded”), or losing both points and health (“Oh no! You have been injured – choose again”). The condition with no change in points was implemented to further distinguish the differing probabilities of rewards and injuries, as opposed to tasks like BART where one is simply the inverse of the other (Lejuez et al., 2002). If the participant’s health drops to zero as a result of a press, they are presented with a “death” screen, in which the entire screen changes to only display the following message: “You have died. You have lost all of your points for this trial. Please wait for a few seconds while the computer resets for the next trial.” Choosing to quit the current trial (by pressing the “q” key) allows the participants to save all their current points and safely end the current “activity”, automatically proceeding to the next trial (with a new picture, text description, and reset health and points). Figure 1 shows an example of how a trial may go for the participant, showing a few of the possible outcomes of each “p” press.



A) Participant begins each new trial with 0 points, 100 health, and is shown an image of a physical risk-taking activity.
B) Participant pressed the “p”roceed key, gaining some points.
C) Another “p” press caused an injury, losing health and points.
D) Another “p” press caused another injury, losing enough health to drop below 0. They are shown a “death” screen and lose any points gained during that trial. If the participant does *not* “die,” they may press the “q”uit key to save all points earned on that trial and begin a new one with reset health and a new image.

Figure 1. Example of an APRT trial

The probabilities of reward or injury, as well as the relative ranges of possible rewards and injuries, are randomly assigned to each trial but are consistent throughout a trial.

Additionally, respondents are (accurately) informed that the border pattern surrounding each trial picture indicates the general reward probability of that trial (solid = high reward probability, striped = low reward probability), while the color of the border indicates the injury probability (blue = low risk of injury, yellow = high risk of injury). This is done to protect the task from learning effects present in other behavioral measures of risk taking, such as the IGT (Maia & McClelland, 2004). By informing the participants of some of the manipulations of each trial, there is less of an ability for some participants to "figure out" the task and use that knowledge to their advantage where an unaware participant could not, therefore making their behavior during the task more indicative of their actual risk-taking propensity, rather than their ability to learn or adapt to the task. Furthermore, neurobiological research has shown that there is stronger activation of brain areas associated with risk perception when the probability of injury and/or reward are somehow indicated compared to when the probabilities are left ambiguous (Burke & Tobler, 2011). Thus, participants would hopefully take these probabilities more seriously when making decisions in the task than if they had to infer them from their own performance. Participants are not, however, given any information about the severity of rewards or injuries in each trial (i.e., how many points/health they could win or lose from a single press). Though they could technically determine this from closely watching how their scores change with each press, it was never explicitly explained to them. Table 1 summarizes the exact probabilities for the different reward and injury conditions, as well as the different ranges of health/points lost per injury and points gained per reward.

As part of the instructions explaining the task, participants are shown previous high scores of other individuals who have completed APRT (which are real scores taken from early pilot studies using a prototype of APRT). They are also told that their total score will be shown to them at the very end of all trials, and that they should try to earn as many points as possible. The participants are also given four practice trials (one from each picture category) to become accustomed to the task before the actual 64 trials begin. Additionally, for this study specifically, half of the participants were assigned to a 1500ms delay condition that slows the speed at which the points and health are updated (the delay occurs after the key press and before the updated health and points scores are displayed), during which they are unable to continue to their next key press decision (as opposed to the non-delay condition, which updates and allows the next press instantaneously). The introduction of delays during risk-taking tasks, such as card games, has been previously shown to reduce maladaptive preservations in task behavior in participants with personality traits related to high levels of risk taking (Newman, Patterson, & Kosson, 1987). I therefore examined whether a delay would also impact APRT performance, regardless of personality. Additionally, studies using psychophysiological measurements, such as an electroencephalogram (EEG), require short delays between recorded responses in order to properly capture event-related potentials (ERP). Thus, I had also wished to know whether a delay in future psychophysiological studies that used APRT waslikely to result in different effects for the various manipulations on participant scores than those observed when no delay was present.

There are four outcomes scores derived from APRT performance: average total Points across all trials, average number of Go (“p”) Presses across all trials, points acquired per injury sustained, average Injuries per trial, and average Remaining Health after each trial. Lastly, the within-subject factors of APRT are Picture Type (animal, heroism, natural disaster, and cliff), Reward Probability (low/high), Reward Amount (low/high), Injury Probability (low/high), and Injury Amount (low/high), which are designed to represent the underlying elements of risk. The only between-subject factor for this study was the Delay condition.

Aims of the Study

The primary aim of the study was to introduce and explore the independent variable-dependent variable relationships of a new behavioral measure of risk-taking, the Assessment of Physical Risk Taking (APRT). Unlike existing behavioral tasks, APRT was designed to assess specific aspects of risk taking that might affect decision-making (e.g., probability and magnitude of rewards and punishments), as opposed to using a much more general, unspecified operationalization of risk-taking as other behavioral measures do.

Because of the multitude of both independent and dependent variables within APRT, there are an extremely high number of predictions and research questions that might have been answered using the measure by itself, to say nothing of the plethora of potential relationships between APRT variables and other risk-taking measures. Rather than produce an excessive number of hypotheses that cover every possible relationship between each of the APRT variables, I instead chose to focus on a select few questions that I felt would provide the strongest evidence for APRT's utility as a behavioral measure of physical risk-taking. They were as follows:

Q1: Do the severity of penalties, frequency of such penalties, extent of rewards, frequency of rewards, and type of displayed activity all independently impact an individual's performance in a simulation task designed to assess that person's propensity to engage in similar, real-world physically risky activities?

Q2: Does the inclusion of a momentary delay in such a task affect the relationship between any of the aforementioned variables and an individual's performance in the game?

Q3: Do any of the variables in Q1 interact with each other to alter relationships these variables have to performance in this task?

Earlier pilot studies that utilized older versions of APRT also sought to answer these questions, and while I updated and altered several features of APRT for the present study, I still wished to replicate some of the more interesting and notable findings from those studies. With both the research questions and prior data in mind, my hypotheses for this study were as follows:

H1: Main effects for all five within-subjects independent variables would be observed for all 4 dependent variables (Points, Injuries, Remaining Health, and Go Presses). The Reward Amount condition was of particular interest, as the difference between the low and high condition point ranges was increased from previous pilot studies to be an entire order of magnitude greater, in order to better contrast the potential outcome differences between the two conditions of this variable, as the original difference in ranges was apparently too small to result in any observable effect (see Table 1 for the exact values used in this study). Additionally, three of the Picture Type categories had also been changed from previous iterations of APRT (the cliff pictures being the only preserved category), to better reflect conceptually different, but still real-world risk behaviors. Thus, I expected there to be strong main effects for this condition and the Reward Amount condition.

H2a: The Delay variable (the only between-subject manipulation) was expected to show significant interactions with several of the other independent variables across many of the dependent variables. Specifically, prior research suggests that an instituted delay can cause decreased perseveration in maladaptive response patterns (Newman et al., 1987), so I expected that the participants in the delay condition would exhibit decreased participation (as measured by Go Presses), as well as lower averages for Points and Injuries, in trials programmed with either high Injury Amount or high Injury Probability compared to trials with the "low" settings for these conditions. The disparity between the high and low conditions for these outcomes (Points, Injuries, and Go presses) was expected to be greater for delayed participants than non-delayed participants, as the former were forced to stop and consider their decision for an additional moment and may potentially have recognized the greater risk of injury (or severity of injury) in some trials, while the latter were free to act impulsively and immediately if they so wished.

H2b: I expected that Delay would also interact significantly with Picture Type. Prior analyses on earlier versions of APRT found that the extent to which the different picture categories affected the average number of Points, Go Presses, and Remaining Health was dependent on whether or not the delay was present, with the presence of the delay usually decreasing the average value of Points, Go Presses, and Health, while also increasing the difference in averages between the picture categories. Forcing some participants to pause and consider their actions seems to have resulted in decreased participation in the task during trials displaying certain (but not all) types of pictures compared to the same trials in non-delayed participants. Because the replacement of three of the picture categories in this update of the APRT was designed to increase the conceptual differences between the categories, I expected a similar, if not stronger, interaction to arise. Specifically, the Picture Types involving actions related to another entity (the hero and animal conditions) should have been more impacted by the Delay (with respect to average Points, Go Presses, and Remaining Health) than more neutral picture types, where an additional moment of consideration was less likely to affect decision-making (i.e., the cliff and disaster photography conditions). Decision-making when another entity is present is inherently more complex than when the individual is alone, and thus additional time to consider the former type of situation was presumably more likely to result in changes to an individual's behavior in those situations compared to the latter, where additional consideration time will likely not change the individual's ultimate decision.

H3a: I anticipated that for the Points outcome, the effect of Reward Amount would differ depending on the level of Injury Amount, with the difference in points earned between high and low Reward Amount trials being greater during trials that also have the low Injury Amount condition, as participants would have been able to capitalize on the reduced injury during high Reward Amount trials and quickly attain much higher scores than during high Injury Amount trials. The risk of losing more health (and potentially all points if they die) would likely result in lower Go Presses (and therefore fewer chances to earn points) in those trials, which would consequently make the difference between high and low Reward Amount trial Points much smaller. The new changes to the Reward Amount conditions would likely result in this interaction occurring.

H3b: I also predicted interactions between Picture Type and Injury Amount, as well as Picture Type and Injury Probability. As stated previously, the new picture categories were intended to represent a wider range of physically risky behaviors, and therefore the increased severity or chance of being harmed during these activities (i.e., in the high condition of Injury Amount/Probability) may have been more apparent to participants in some of the picture types versus others. Specifically, I expected the animal pictures with high Injury Amount or high Injury Probability to result in lower scores for the Points, Injuries, and Go Presses outcomes compared to animal pictures with low Injury Amount/Probability. The participants were likely to assume the more severe or more frequent injuries received during the former were the result of the particular animal they are encountering being more dangerous than others (i.e., those animals displayed during low Injury Amount/Probability trials) and would have therefore quickly moved on from these trials, resulting in fewer Points, Injuries, and Go Presses. I predicted Remaining

Health would be higher in these high-risk animal trials, as engaging in the task less would (on average) result in fewer lost points and health.

I did not expect to see a similar effect for the cliff and natural disaster photography categories, however, as these pictures do not involve an agent creating danger for the player nor are the pictures as varied in content, so participants were probably less likely to attribute noticed differences in injury frequency/amount to the situation depicted by the particular trial picture in these categories. They were thus less likely to change their playstyle as radically as I believed they would in the high-risk animal trials. Finally, for the heroic picture I category I anticipated that high Injury Probability would not differ significantly from its low counterpart on any of the outcome scores, but that high Injury Amount trials would result in fewer Points, Go Presses, and Injuries than low Injury Amount trials in this picture category. This was due to the altruistic nature of the category (trying to save someone from danger) likely causing the majority of participants to be willing to risk acquiring a few additional smaller injuries for the sake of the "person" in danger, but more severe injuries resulting in greater hesitancy to continue to engage in the task (i.e., few would risk death to save a stranger from danger).

H3c. There would be a significant interaction between Injury Probability and Reward Probability (the only dichotomous independent variables fully visible to the participant in each trial) for all four outcome variables. Because the conditions (low/high) of these variables are both indicated by the border of each trial's picture, participants would likely use this information to adjust their playstyle. Low Injury Probability and high Reward Probability would result in greater participation and success in those trials (resulting in higher Points, Go Presses, and Remaining Health, in addition to fewer Injuries), while high Injury Probability and low Reward Probability trials would result in little time spent on those trials (resulting in fewer Points, Go

Presses, and Injuries, along with greater Remaining Health). Trials in which both variables are high or low would likely result in the variables cancelling the other's effects out, resulting in outcome scores that are roughly in between those obtained in the other two types of trial combinations for these variables.

Method

Participants

Participants were 256 students recruited through the psychology subject pool at UNLV. They received class credit as compensation. The sample was 71% female and had a mean age of 19.71 (*SD* = 3.03, Range = 18-42). The racial and ethnic distribution for the sample was: 40.1% Caucasian, 38.5% Asian, 13.9% African-American, 6.4% Pacific Islander, 1.2% Native American, and 9.9% choosing not to provide information regarding their race, with 35.7% of participants also identifying as Hispanic. A pre-hoc power analysis revealed that, in order to observe the smallest significant main effect size found in an earlier pilot study (d = 0.167) at the recommended power level of .80, a sample size of 281 would have been required. However, the next smallest main effect size (d = 0.232) only required a sample of 202. Thus, I had a sufficiently powered sample to detect effects of similar size to the pilot study, except for the very smallest effect from that sample.

Procedure

Participants completed APRT on computers located in the UNLV Psychophysiology of Emotion and Personality (PEP) Lab as part of a larger study on personality and behavior. The task was programmed in PsychoPy and the trials were counterbalanced across all 5 independent variables. Each participant was assigned a subject number, with even-numbered subjects being placed into the Delay condition. Their completion of the study was overseen by a trained research assistant, who sat them at one of five lab computers and give each participant the opportunity to ask questions whilst undergoing the consent process. All participants received a verbal debriefing once the study procedure was complete.

Analyses

Main Effects and Interactions. To model the effects of the independent variables on the four different dependent variables, a series of generalized estimating equations (GEEs) were utilized. GEEs are an alternative to Generalized Linear Models that allow for correlations among the observations (such as those present in repeated measure tasks such as APRT) and do not need to meet the same assumptions about the distribution of the variables or residuals that traditional linear regression models do, while still providing robust and unbiased estimates of the standard errors for the estimated coefficients (Burton, Gurrin, & Sly, 1998; Hanley, Negassa, Edwardes, & Forrester, 2003). A separate GEE was created for each dependent variable to model the main effects and interactions of the 6 independent variables. All main effects, two-way interactions among the 5 within-subject factors, and interactions between within-subject factors and Delay (the sole between-subject factor) were estimated assuming an unstructured estimated covariance matrix (i.e., a specific pattern was not assumed for the relationships between observations, a method which results in more robust standard errors but slightly less statistical power). Due to heavy skew in the dependent variables, a log transformation was performed prior to running each GEE. While this was not technically required to run the GEEs properly, it helped retain some of the statistical power that would have been otherwise lost from assuming the unstructured covariance matrix.

Statistically significant main effects for Picture Type and interactions were decomposed using *post hoc* pairwise comparisons utilizing a sequential-Sidak (also known as Holm-Sidak) correction to control for multiple comparisons. This method resembles a sequential Bonferroni correction in method but employs a more complex formula to determine the critical α for each comparison, in addition to assuming the comparisons are independent of each other (Holm, 1979). A starting critical α of .05 was used for all analyses, and all analyses were conducted in SPSS.

Results

Generalized Estimating Equations

The results of the GEEs are summarized in Table 2. Significant main effects and interactions are explored in detail below. Table 3 summarizes the relevant means, standard errors, and Wald χ^2 values for all statistically significant interactions. Figures 2 through 4 illustrate a few of the most important (and largest) interaction effects that emerged from the data. These types of interactions form the basis of the subsequent discussion of how APRT seems to be assessing risk-taking. Figure 2 shows how the slopes in both conditions are discernible and that the slope in one condition is discernibly greater than that in another condition. Figure 3 displays an interaction in which there is no discernible slope between two levels of one factor in one condition but there is a discernible slope in the other. Figure 4 demonstrates how slopes of one factor are in opposite directions based on levels of another factor.

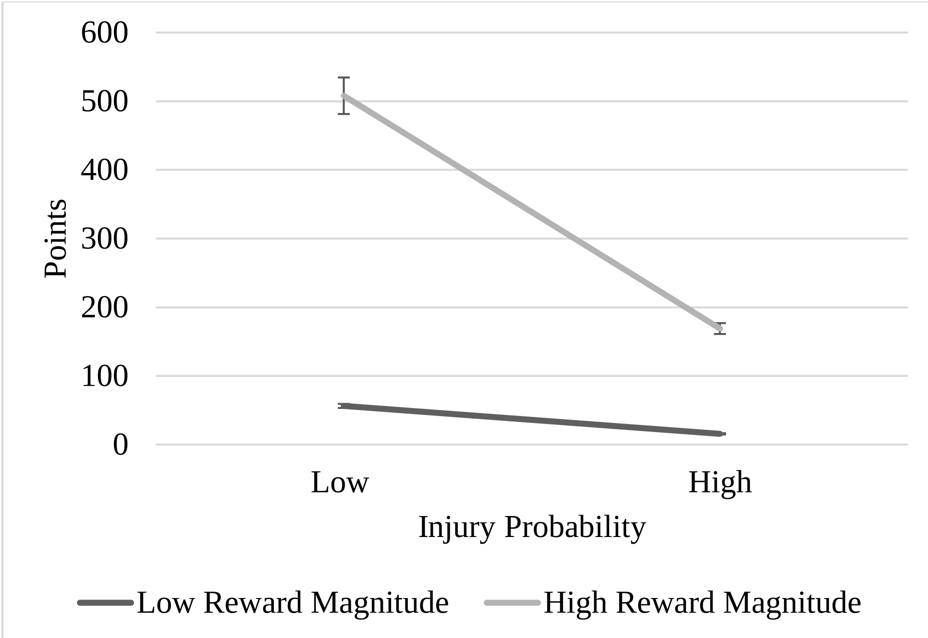


Figure 2. Two-way interaction between Reward Magnitude and Injury Probability on APRT Points score.

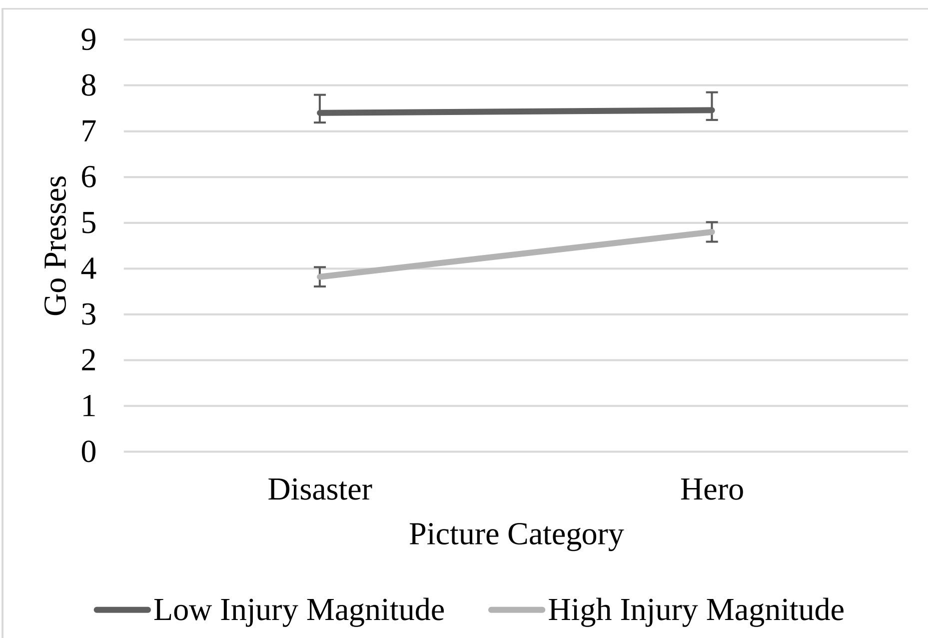


Figure 3. Two-way interaction between Injury Magnitude and Picture Category (Disaster-Hero) on APRT Go Presses score.

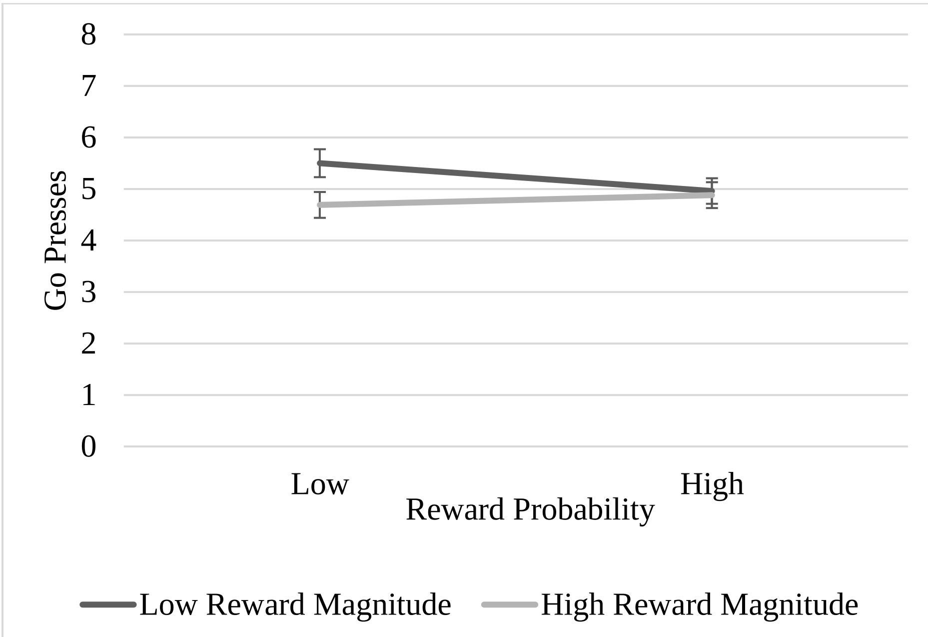


Figure 4. Two-way interaction between Reward Magnitude and Reward Probability on APRT Go Presses score.

Points.

***Main Effects*.** There was a main effect for Picture Type. The Hero condition resulted in the largest number of Points earned on average (*M* = 114.23, *SE* = 5.42), which was significantly greater than the other three conditions. The Animal condition resulted in the lowest number of Points on average (*M* = 69.09, *SE* = 5.42), which was significantly lower than all the other conditions. The average Points earned during Cliff (*M* = 95.28, *SE* = 5.11) and Disaster (*M* = 100.94, *SE* = 5.70) trials did not significantly differ, however.

Participants earned more Points in low Injury Magnitude trials (*M* = 143.53, *SE* = 7.62) than in high Injury Magnitude trials (*M* = 60.70, *SE* = 2.84) and in low Injury Probability trials (*M* = 169.19, *SE* = 8.24) than in high Injury Probability trials. They also earned more Points in high Reward Magnitude trials (*M* = 292.98, *SE* = 13.34) than in low Reward Magnitude trials (*M* = 29.74, *SE* = 1.44) and in high Reward Probability trials (*M* = 124.80, *SE* = 5.66) than in low Reward Probability trials (*M* = 69.08, *SE* = 3.47). After correcting for multiple comparisons, there was no main effect of Delay on Points.

***Within-subject interactions*.** There was a Picture Type x Injury Probability interaction,. There was no difference in average Points per trial between Disaster and Hero trials when Injury Probability was low, but when Injury Probability was high, Hero trials resulted in higher Points scores on average compared to Disaster trials. Additionally, the increase in Points going from the Animal to the Cliff trials was slightly sharper when Injury Probability was high compared to when it was low.

There was an Injury Magnitude x Reward Magnitude interaction. When Reward Magnitude was high, the decrease in average Points per trial going from the low Injury Magnitude to the high Injury Magnitude condition was substantially sharper compared to when Reward Magnitude was low.

There was an Injury Magnitude x Injury Probability interaction. When Injury Probability was low, the decrease in average Points going from the low Injury Magnitude condition to the high Injury Magnitude condition was slightly sharper compared to when Injury Probability was high.

There was a Reward Magnitude x Injury Probability interaction. When Reward Magnitude was high, the decrease in average Points going from the low Injury Probability condition to the high Injury Probability condition was substantially sharper compared to when it was low.

Finally, there was an Injury Probability x Reward Probability interaction. When Injury Probability was low, the increase in average Points going from the low Reward Probability condition to the high Reward Probability condition was slightly sharper compared to when it was high. No other significant within-subject interactions for Points were observed.

***Interactions involving Delay*.** After correcting for multiple comparisons, no significant interactions involving Delay were observed.

Go Presses.

Main Effects. There was a main effect for Picture Type. All picture categories had significantly different Go Presses from each other. The Hero condition had the most Go Presses on average (M = 5.98, SE = 0.27), followed by Disaster (M = 5.31, SE = 0.27), Cliff (M = 4.85, SE = 0.24), and Animal (M = 4.05, SE = 0.26).

Participants made more Go Presses in low Injury Magnitude trials (M = 6.55, SE = 0.35) than in high Injury Magnitude trials (M = 3.81, SE = 0.17) and in low Injury Probability trials (M = 7.36, SE = 0.38) than in high Injury Probability trials (M = 3.40, SE = 0.16). Participants also made more Go Presses in low Reward Magnitude trials (M = 5.23, SE = 0.25) than in high Reward Magnitude trials (M = 4.78, SE = 0.24). Finally, participants who were given the forced delay between responses had fewer Go Presses (M = 4.35, SE = 0.29) on average than participants that were not delayed (M = 5.75, SE = 0.39). After correcting for multiple comparisons, there was no main effect for Reward Probability.

***Within-subject interactions*.** There was a Picture Type x Injury Magnitude interaction. There was no significant difference in Go Presses between Disaster and Hero trials when Injury

Magnitude was low, but when Injury Magnitude was high, Hero trials resulted in significantly more Go Presses compared to Disaster trials. Additionally, when Injury Magnitude was high, the increase in Go Presses when going from Animal trials to Cliff trials was slightly sharper compared to when it was low.

There was a Picture Type x Injury Probability interaction. When Injury Probability was low, there was no difference in Go Presses between Disaster and Hero trials, but Hero trials resulted in significantly more Go Presses compared to Disaster trials when Injury Probability was high. Additionally, when Injury Probability was low, the increase in Go Presses going from Animal to Cliff trials was slightly sharper compared to when it was high.

There was an Injury Magnitude x Injury Probability interaction. When Injury Probability was low, the decrease in Go Presses going from the low Injury Magnitude to the high Injury Magnitude condition was slightly sharper compared to when it was high.

Finally, there was a Reward Magnitude x Reward Probability interaction. When Reward Probability was high, there was no difference in Go Presses between the low and high Reward Magnitude conditions, but high Reward Magnitude trials resulted in significantly fewer Go Presses compared to low Reward Magnitude trials when Reward Probability was low.

***Interactions involving Delay*.** After correcting for multiple comparisons, no significant interactions involving Delay were observed for Go Presses.

Remaining Health.

***Main Effects*.** There was a main effect for Picture Type. All picture categories had significantly different amounts of Remaining Heath from each other on average. Animal trials had the highest Remaining Health on average (M = 76.41, SE = 1.26), followed by Cliff trials (M

= 73.26, *SE* = 1.19), Disaster trials (*M* = 70.27, *SE* = 1.31), and lastly Hero trials (*M* = 67.02, *SE* = 1.25).

Participants had higher Remaining Health on average in low Injury Magnitude trials (*M* = 80, *SE* = 1.06) than in high Injury Magnitude trials (*M* = 63.48, *SE* = 1.33) and in low Injury Probability trials (*M* = 75.80, *SE* = 1.13) than in high Injury Probability trials (*M* = 67.69, *SE* = 1.24). Participants also had greater Remaining Health in high Reward Magnitude trials (*M* = 72.56, *SE* = 1.16) than in low Reward Magnitude trials (*M* = 70.92, *SE* = 1.20) and slightly more Remaining Health in high Reward Probability trials on average (*M* = 72.34, *SE* = 1.17) than in low Reward Probability trials (*M* = 71.15, *SE* = 1.18). Finally, participants who were given the delay had higher Remaining Health on average (*M* = 77.87, *SE* = 1.43) than participants not given the delay (*M* = 65.61, *SE* = 1.82).

***Within-subject interactions*.** There was a Picture Type x Injury Magnitude interaction. When Injury Magnitude was low, there was no significant difference in average Remaining Health between Disaster and Hero trials, but Hero trials resulted in significantly lower Remaining Health scores compared to Disaster trials when Injury Magnitude was high. A similar effect was observed for the difference between Animal and Cliff trials, wherein when Injury Magnitude was low, there was no significant difference in Health scores between the two, but when Injury Magnitude was high, Cliff trials resulted in lower Remaining Health on average compared to Animal trials.

There was a Picture Type x Reward Magnitude interaction. When Reward Magnitude was low, there was no significant difference in Remaining Health scores between Disaster and Hero trials, but when Reward Magnitude was high, Hero trials resulted in significantly lower Remaining Health scores on average compared to Disaster trials. A similar effect was observed for Animal and Cliff trials, wherein there was no significant difference between them when Reward Magnitude was low, but when Reward Magnitude was high, Cliff trials resulted in significantly lower Health scores on average compared to Animal trials.

There was a Picture Type x Injury Probability interaction. When Injury Probability was low, there was no significant difference in Remaining Health scores between Disaster and Hero trials, but when Injury Probability was high, Hero trials resulted in significantly lower Health scores on average compared to Disaster trials. Additionally, when Injury Probability was low, Cliff trials resulted in significantly lower Health scores compared to Animal trials, but there was no significant difference in scores between Cliff and Animal trials when Injury Probability was high.

Finally, there was an Injury Magnitude x Injury Probability interaction. When Injury Magnitude was high, the decrease in Remaining Health score when going from the low Injury Probability condition to the high Injury Probability condition was slightly sharper compared to when Injury Magnitude was low.

***Interactions involving Delay*.** There was a Picture Type x Delay interaction. For non-delayed participants, Hero trial Remaining Health scores were not significantly different from Disaster trial scores, but Hero trials resulted in significantly higher Health scores compared to Disaster trials in the delayed participants.

Injuries.

Main Effects. There was a main effect for Picture Type. Each picture category differed significantly from the other three in average Injuries for each trial. Participants had significantly greater Injuries in low Injury Magnitude trials (M = 0.91, SE = 0.05) than in high Injury Magnitude trials (M = 0.50, SE = 0.02) and in high Injury Probability trials (M = 0.82, SE = 0.03) than in low Injury Probability trials (*M* = 0.56, *SE* = 0.03). Participants suffered significantly more Injuries in low Reward Magnitude trials (*M* = 0.70, *SE* = 0.03) than in high Reward Magnitude trials (*M* = 0.65, *SE* = 0.03). There was a main effect for Injury Probability. Finally, participants that were given the delay had significantly fewer Injuries on average (*M* = 0.55, *SE* = 0.03) than participants that were not given the delay (*M* = 0.83, *SE* = 0.05). After controlling for multiple comparisons, there was no main effect for Reward Probability.

***Within-subject interactions*.** There was a Picture Type x Injury Magnitude interaction. When Injury Magnitude was low, there was no significant difference in average Injuries per trial between Disaster and Hero trials, but there were significantly more Injuries during Hero trials compared to Disaster trials when Injury Magnitude was high. Additionally, the increase in average Injuries score when going from Animal to Cliff trials was slightly sharper when Injury Magnitude was high compared to when it was low.

There was a Picture Type x Injury Probability interaction. When Injury Probability was low, there was no significant difference in Injury score between Disaster and Hero trials, but Hero trials resulted in significantly more average Injuries compared to Disaster trials when Injury Probability was high. Additionally, when Injury Probability was low, there was no significant difference in Injuries score between Animal and Cliff trials, but Cliff trials resulted in significantly more Injuries on average compared to Animal trials when Injury Probability was high.

There was a Picture Type x Reward Probability interaction. When Reward Probability was high, the increase in average Injuries per trial when going from Disaster to Hero trials was much sharper compared to when Reward Probability was low. Additionally, the increase in Injuries score when going from Animal to Cliff trials was also much sharper when Reward Probability was high compared to when it was low.

Finally, there was a Reward Magnitude x Reward Probability interaction. When Reward Probability was high, there was no significant difference in Injuries score between the low and high Reward Magnitude conditions, but high Reward Magnitude trials resulted in significantly fewer injuries compared to low Reward Magnitude trials when reward Probability was high.

***Interactions involving Delay*.** There was an Injury Probability x Delay interaction. The increase in average Injuries score when going from low Injury Probability trials to high Injury Probability trials was sharper in the delayed participants compared to the non-delayed participants.

Discussion

The goal of the present study was to better understand how specific aspects of risk influenced scores on a new behavioral measure of risk-taking. Results indicate mixed support for my hypotheses. Table 4 summarizes how the current results compared to my hypotheses. While there were significant and strong main effects for nearly all the within-subject variables across all four APRT scores, many of the interactions I anticipated were not statistically significant. Specifically, the Reward Magnitude x Injury Magnitude and Reward Probability x Injury Probability interactions were both only significant for Points scores, while all but two of the anticipated interactions involving Delay were not statistically significant (though there were significant main effects across all scores except Points). On the other hand, the anticipated Picture Type x Injury Magnitude and Picture Type x Injury Probability interactions were significant across all scores, except for Picture Type x Injury Magnitude on Points. Of these findings, the lack of Delay interactions was most surprising, as while there was a basic difference in most scores between the delayed and non-delayed participants (i.e., main effects), Delay did essentially nothing (with two exceptions, see Table 3) to moderate the influence of the other manipulations on scores. This bodes well for APRT, as it indicates a small delay does not drastically alter APRT scores or the effects of other variables that influence these scores. This is essential for study designs that may require short delays between responses, such as studies using psychophysiological measurements.

As expected, there were also a number of other significant interactions that were not included in my hypotheses. Most of these were fairly intuitive, such as the significant interaction between Injury Magnitude and Injury Probability on Points, Go Press, and Remaining Health scores, which essentially caused each variable to strengthen the main effect of the other (i.e., the difference in scores between the low and high conditions of each variable increased when the other was at the same low or high condition in a given trial). Interestingly, the corresponding interaction between Reward Magnitude and Reward Probability was only significant for Go Press and Injury scores, and even then, the effect size was much smaller (as evidenced by the Wald χ^2 values).

The clearest pattern from the GEE results was the superiority of the Injury-related variables compared to the Reward-related variables. There were more significant interactions involving Injury Magnitude and Probability than Reward Magnitude and Probability (as shown in Table 3), and the majority of the Injury-related interactions were larger in effect size than the majority of the Reward-related interactions. Simply put, Injury Magnitude and Probability seem to matter much more to participants than Reward Magnitude and Probability when making risky decisions. This is especially striking when comparing Reward Probability (which had few significant effects and weak effect sizes for those that were significant) to Injury Probability (which had the largest number of main effects/interactions and some of the strongest effect sizes in all four GEE models).

Picture Type fell somewhere between these two variables, having strong main effects but generally having interactions with only the Injury Magnitude and Probability variables. As predicted, the difference in Animal and Cliff trial scores tended to increase when either Injury Magnitude or Injury Probability was high, suggesting increased defensive emotional processing when faced with a dangerous animal, a response that is consistent with similar findings in studies on startle blink potentiation that also used pictures of threatening animals as stimuli (Bradley, Codispoti, Cuthbert, & Lang, 2001; Quevedo, Benning, Gunnar, & Dahl, 2009). Additionally, the difference in Hero and Disaster trial scores was only significantly different when Injury

Probability or Magnitude was high, with all APRT scores being higher (except Health, which is expected from receiving more injuries) in the Hero condition. These Hero-Disaster interactions were particularly interesting, as they suggest that participants were more willing to risk harm to self when the stakes were higher, but only if the risk was related to attempting to help another person in danger. This is consistent with some of the literature on risk-taking and altruism, which posits that individuals are more willing to accept higher risks if it is for the sake of aiding another person (Fagin-Jones & Midlarsky, 2007; Wu et al., 2009).

The overall pattern of results across each of the APRT scores seems to suggest that the task assesses risk-taking in a manner different to that of other tasks, like the BART. Specifically, higher risk-taking in the context of APRT performance seems to be the failure of a respondent to alter their behavior (e.g., Go Presses) in response to information indicating an increased chance of Injury (e.g., there is little to no difference in Go Presses when Injury Magnitude or Probability switches from low to high), whereas individuals with low or normal levels of risk-taking would tend to heed this information more readily. In other words, risk-taking in APRT is indicated as the magnitude of the difference in scores between the Injury variable conditions, rather than as an overall score.

This tendency to ignore increases in risk can be further understood as the opposite of a satisficing drive that seems to be present in the normal population. This is evidenced within APRT through the pattern of the Points scores, which indicated that most participants would become satisfied with the amount of Points they have earned in a trial after reaching a certain limit, at which point they would quit the trial, even if Injury Probability and/or Magnitude was low and Reward Magnitude was high (i.e., when they had low-risk opportunities to earn large amounts of Points indefinitely). Thus, APRT presents a novel conceptualization of risk-taking: a contextual insensitivity to increases in risk. This definition is a direct complement to the existing concept of satisficing, which describes the tendency of most individuals to pursue a suboptimal target (i.e., not maximizing reward) in exchange for decreased risk (Mohamed, 2006). Satisficing has presently been more heavily examined (with respect to risk-taking) in behavioral economics than in other branches of psychology (Parker, de Bruin, & Fischhoff, 2007). APRT thus potentially merges risk-taking research across multiple domains under a uniform conceptualization of risky behavior by providing separate operationalizations of risk-taking and satisficing (Go Presses and Points, respectively).

There are a number of limitations to consider when interpreting the findings of the current study. First and foremost, the sample was likely underpowered. When I conducted my power analysis to determine sample size, I was only considering main effects and did not account for the fourfold increase required for interactions of the same effect size (Leon & Heo, 2009). Thus, the actual sample size required to detect the anticipated effect sizes of any interactions would have been over 800, much more than the 256 participants I had. In this light, there may have been additional interaction effects that were not detected by my analyses due to insufficient power.

Another limitation was the limited engagement of many of the participants in the study. While I initially had considered excluding participants who had zero Go Presses for half or more of the APRT trials, I recognized that this would have eliminated around a third of the sample, which would have resulted in the sample being insufficiently powered to detect even the main effect sizes I anticipated. Thus, while I kept these participants in the sample for the sake of preserving power (as well as to not potentially eliminate a subgroup of participants that may have indeed simply been very cautious, as opposed to disinterested, during the task), their high number of zero-scored trials likely skewed the means for all four outcome scores.

The significant main effects for Delay may also have been a function of these zero Go Presses trials, as participants that were given the response delay tended to have more of these trials on average. Delay may have thus acted as an unintentional motivator for an unmeasured "reward" to participants: the ability to finish the experiment sooner. A study has previously shown that participants were willing to sacrifice further potential reward on BART trials when they were given a similar type of delay, if that sacrifice meant being able to complete the experiment in a significantly faster time (Young, Webb, Rung, & McCoy, 2014). Since APRT involves a similar repetitive, repeated measure type task, a similar effect may have been present in the current study, especially if the delayed participants were feeling particularly bored or fatigued by the task. This potential effect served as an additional reason to not exclude participants with low engagement in APRT.

Beyond this unintentional effect of the delay condition, the high number of participants refusing to make even a single Go Press on a very large number of trials, regardless of whether they received the delay or not, suggests possible issues with the sample population itself being generally uninterested or disengaged while completing the study, a problem frequently found in psychological research using college students (Peterson & Merunka, 2014). Furthermore, the sample in the study was a convenience sample drawn from university psychology majors, a group not particularly known to be high in risk-taking behavior, compared to other groups such as criminals or substance abusers (Knust & Stewart, 2002; Wills, Vaccaro, & McNamara, 1994). There is thus far less variation in APRT scores than might be expected from a more diverse sample or a sample drawn from populations known to be high in risk-taking. This limitation likely has impacted the strength and significance of the observed effects by estimating these parameters from only a small portion of the true distribution of APRT scores.

There are also several limitations stemming from the APRT trials themselves. Firstly, each of the 64 trials represents a unique combination of the within-subject manipulations. This means that there may be more complex interaction effects present (3-way or even up to 5-way), which were not examined due to the difficulty (if not impossibility) of conceptually explaining them in ecologically useful terms (i.e., what they mean in terms of real-world risky decision making). A second issue with the APRT trials is the design of the Picture Type condition. The picture categories were ultimately determined on the basis of convenience, as I wished to avoid copyright issues and was limited by what pictures were available for free use. Thus, despite my best efforts, the categories may not be the best representation of different types of physically risky behaviors. Images of more disparate types of physically risky behaviors may have resulted in stronger effect sizes, or at least may have allowed for a more structured or meaningful approach to the design of these variable conditions.

A final limitation to this study is its failure to include any assessment of life history factors that might have impacted the results of some participants. Life history theory is a commonly used framework for understanding and analyzing human behavior in evolutionary biology, anthropology, and other fields, in addition to psychology (Mittal & Griskevicius, 2014). Previous research on life history variables and risk-taking has suggested that in addition to a basic gender difference in risk-taking propensity, factors such as age, parental status, birth order, number of siblings, and subjective life expectancy all substantially impact an individual's risk-taking, particularly in the physical risk domain that APRT assess (Wang, Kruger, & Wilke, 2009). Many of these have an evolutionary explanation (e.g., a mother taking less risks to ensure she can continue caring for her children or taking more risks in order to protect them from harm). Future studies should include measures of these factors and attempt to account for them when using the APRT to assess risk-taking.

Despite the limitations of the study and its task, this study marks the comprehensive assessment of the elements of risk taking across a variety of risky behaviors using a single behavioral task. While some of these factors are present in other behavioral risk-taking tasks, they are rarely examined or analyzed in any way, and these other tasks each only examine a single risky behavior (Bechara, 2007; Lejuez et al., 2002). APRT directly addresses these shortcomings while retaining the positive features of these measures (e.g., the simplicity of the task).

Many of the present issues with APRT could also likely be improved by revising certain aspects of the measure (e.g., creating more conceptually distinct and meaningful categories for Picture Type). If such improvements are made, APRT may mark the next major leap in research concerning behavioral assessment of risk-taking, as it would be the first to provide a comprehensive assessment of the elements of risk across a variety of different risky activities, in addition to providing a novel conceptualization of risk-taking itself.

Appendix

Table 1

Summary of Reward/Injury Probabilities and Severities

Variable	Low	High
Injury Magnitude (per injury)	Randomized Range: 5-35	Randomized Range: 60-90
Reward Magnitude (per reward)	Randomized Range: 5-25	Randomized Range: 50-250
Injury Probability	1/15	1/5
Reward Probability	2/3	9/10

Table 2

Summary of Generalized Estimating Equation Model Effects (Wald χ^2) for Points, Go Presses, Injuries, and Health

Model Effect	*df*	Points	Go Presses	Remaining Health	Injuries
Picture Type	3	45.60**	87.07**	109.74**	77.94**
Injury Magnitude	1	323.89**	533.05**	579.21**	339.64**
Reward Magnitude	1	4879.81**	17.95**	11.21**	11.50**
Injury Probability	1	863.75**	1172.02**	231.73**	254.92**
Reward Probability	1	258.11**	2.75	8.51**	2.69
Delay	1	3.59	9.00**	27.75**	22.13**
Picture Type x Injury Magnitude	3	3.94	22.38**	76.59**	23.39**
Picture Type x Reward Magnitude	3	2.26	3.98	21.84**	7.68
Picture Type x Injury Probability	3	14.65**	31.91**	17.13**	33.02**
Picture Type x Reward Probability	3	1.05	5.80	11.96*	27.72**

Injury Magnitude x Reward Magnitude	1	24.27**	0.30	0.04	6.14*
Injury Magnitude x Injury Probability	1	34.07**	71.19**	9.44**	1.46
Injury Magnitude x Reward Probability	1	3.51	1.91	0.38	0.001
Reward Magnitude x Injury Probability	1	10.98**	2.59	1.14	0.02
Reward Magnitude x Reward Probability	1	0.90	14.87**	1.56	16.00**
Injury Probability x Reward Probability	1	15.74**	0.06	0.01	0.01
Picture Type x Delay	3	4.62	8.62*	17.08**	3.54
Injury Magnitude x Delay	1	5.76*	4.66*	0.01	0.45
Reward Magnitude x Delay	1	0.03	0.17	0.05	0.03
Injury Probability x Delay	1	5.00*	3.15	0.97	11.02**
Reward Probability x Delay	1	0.48	0.03	5.74*	0.07

df= degrees of freedom. * $p < .05$ (uncorrected). ** $p_{adj} < .05$ (sequential Sidak corrected).

Table 3

Decompositions of Statistically Significant APRT Interactions

	Low v1/Low v2		Low v1/High v2			High v1/Low v2		High v1/High v2		
Interaction (v1/v2)	*M*	*SE*	*M*	*SE*	Wald $\chi^2(1)$	*M*	*SE*	*M*	*SE*	Wald $\chi^2(1)$
					POINTS					
Injury Probability/Disaster-Hero	196.94	10.80	196.35	10.26	0.00	51.74	3.87	66.45	3.83	8.70^{*}
Injury Probability/Animal-Cliff	131.28	10.20	161.43	9.24	9.00^{*}	36.36	3.55	56.24	3.88	18.16^{***}
Injury Magnitude/Reward Magnitude	49.06	2.76	419.91	22.81	326.62^{***}	18.02	1.11	204.41	9.16	467.55^{***}
Injury Magnitude/Injury Probability	236.00	13.6	87.29	4.68	215.96^{***}	121.29	5.586	30.38	1.98	273.27^{***}
Reward Magnitude/Injury Probability	56.34	2.95	15.69	0.92	242.57^{***}	508.05	26.58	168.95	7.94	238.68^{***}
Injury Probability/Reward Probability	133.38	6.76	214.62	11.12	138.29^{***}	36.54	2.29	72.57	3.42	150.27^{***}
					GO PRESSES					

Injury Magnitude/Disaster -Hero	7.38	0.40	7.46	0.39	0.15	3.82	0.21	4.80	0.21	13.80***
Injury Magnitude/Animal-Cliff	5.47	0.36	6.12	0.36	10.46**	3.00	0.21	3.84	0.19	21.78***
Injury Probability/Disaster -Hero	8.48	0.44	8.54	0.43	0.05	3.33	0.19	4.19	0.19	29.57***
Injury Probability/Animal-Cliff	5.90	0.42	6.85	0.38	11.51**	2.78	0.18	3.43	0.18	21.56***
Injury Magnitude/Injury Probability	9.02	0.51	4.79	0.25	202.59***	6.00	0.30	2.43	0.11	264.80***
Reward Magnitude/Reward Probability	5.50	0.27	4.96	0.25	14.88***	4.69	0.25	4.88	0.25	2.14
					REMAINING HEALTH					
Injury Magnitude/Disaster -Hero	77.99	1.20	78.03	1.19	0.003	62.56	1.59	56.03	1.47	35.24***
Injury Magnitude/Animal-Cliff	82.39	1.11	81.61	1.11	1.40	70.43	1.54	64.91	1.44	23.45***
Reward Magnitude/Disaster -Hero	67.22	1.52	68.02	1.38	0.35	73.32	1.36	66.04	1.35	51.08***
Reward Magnitude/Animal-Cliff	75.19	1.45	73.28	1.31	2.15	77.64	1.34	73.25	1.34	12.74**

Injury Probability/Disaster-Hero	72.76	1.36	72.17	1.31	0.36	67.79	1.44	61.88	1.37	29.30***
Injury Probability/Animal-Cliff	81.04	1.24	77.23	1.23	14.52***	71.79	1.42	69.29	1.35	5.84
Injury Magnitude/Injury Probability	83.49	1.04	76.52	1.45	186.78***	68.11	1.32	58.85	1.45	148.13***
Delay/Disaster-Hero	64.11	2.00	61.25	2.00	5.87	64.11	1.65	72.80	1.53	9.05***
Delay/Animal-Cliff	71.94	2.00	65.15	2.12	28.14**	80.88	1.56	81.37	1.8	0.18
					INJURIES					
Injury Magnitude/Disaster-Hero	1.00	0.05	1.03	0.05	1.00	0.51	0.02	0.63	0.02	30.25***
Injury Magnitude/Animal-Cliff	0.77	0.05	0.86	0.05	7.11**	0.40	0.03	0.50	0.02	20.25***
Injury Probability/Disaster-Hero	0.65	0.03	0.65	0.03	0.004	0.79	0.04	1.00	0.04	29.30***
Injury Probability/Animal-Cliff	0.46	0.03	0.52	0.03	9.00	0.68	0.04	0.82	0.04	21.78***
Reward Probability/Disaster-Hero	0.74	0.04	0.77	0.03	16.00***	0.70	0.04	0.84	0.04	21.78***

Reward Probability/Animal-Cliff	0.60	0.04	0.65	0.03	4.00	0.52	0.03	0.66	0.03	49.00***
Reward Magnitude/Reward Probability	0.74	0.03	0.67	0.03	16.89***	0.64	0.03	0.66	0.03	1.67***
Delay/Injury Probability	0.75	0.05	0.96	0.06	144.00***	0.44	0.03	0.70	0.04	169.00***

Note. For interactions involving Delay, “Low” is considered the non-delayed participants, while “High” denotes delayed participants.

For interactions involving Animal/Cliff/Hero/Disaster, Animal and Disaster are considered “Low” and Cliff and Hero are considered “High.” *p_{adj} < .05. ** p_{adj} < .01. *** p_{adj} < .001.

Table 4

Summary of Support for Hypotheses

Hypothesis	Overall Support	Comment
1. Main effects for all within-subject IVs across all four DVs.	Yes	Except for Reward Probability on Go Presses and Injuries
2. Delay x Injury Magnitude and Delay x Injury Probability interactions on Go Presses, Points, and Injuries.	Mixed	Only for some DVs
3. Delay x Picture Type interaction on Points, Go Presses, and Remaining Health	No	Only for Go Presses
4. Reward Magnitude x Injury Magnitude interaction on Points	Yes	
5. Picture Type x Injury Magnitude and Picture Type x Injury Probability interaction on Points, Injuries, and Go Presses	Yes	Except for Injury Magnitude on Points
6. Injury Probability x Reward Probability interaction across all four DVs	No	Only for Points

References

Aklin, W. M., Lejuez, C. W., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour Research and Therapy*, *43*(2), 215–228. https://doi.org/10.1016/j.brat.2003.12.007

Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, *16*(2), 289–296. https://doi.org/10.1016/0191-8869(94)90165-1

Bechara. (2007). *Iowa gambling task professional manual*. Psychological Assessment Resources.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). *Research Focus The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers*.

Bechara, Antoine, Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bechara, Antoine, Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*(11), 2189–2202. https://doi.org/10.1093/brain/123.11.2189

Blais, A.-R., & Weber, E. U. (2006). A Domain-Speci c Risk-Taking (DOSPERT) Scale for Adult Populations. *Judgment and Decision Making*, *1*(1), 15.

Bornovalova, M. A., Daughters, S. B., Hernandez, G. D., Richards, J. B., & Lejuez, C. W. (2005). Differences in impulsivity and risk-taking propensity between primary users of crack cocaine and primary users of heroin in a residential substance-use program. *Experimental and Clinical Psychopharmacology*, *13*(4), 311–318. https://doi.org/10.1037/1064-1297.13.4.311

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298. https://doi.org/10.1037//1528-3542.1.3.276

Buelow, M. T., & Blaine, A. L. (2015). The assessment of risky decision making: A factor analysis of performance on the Iowa Gambling Task, Balloon Analogue Risk Task, and Columbia Card Task. *Psychological Assessment*, *27*(3), 777–785. https://doi.org/10.1037/a0038622

Buelow, M. T., & Suhr, J. A. (2009). Construct Validity of the Iowa Gambling Task. *Neuropsychology Review*, *19*(1), 102–114. https://doi.org/10.1007/s11065-009-9083-4

Burke, C. J., & Tobler, P. N. (2011). Coding of Reward Probability and Risk by Single Neurons in Animals. *Frontiers in Neuroscience*, *5*. https://doi.org/10.3389/fnins.2011.00121

Burton, P., Gurrin, L., & Sly, P. (1998). Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multi-level mixed modelling. *Statistics in Medicine*, *17*(11), 1261–1291. https://doi.org/10.1002/(SICI)1097-0258(19980615)17:11<1261::AID-SIM846>3.0.CO;2-Z

Carmichael, D. G. (2016). Risk – a commentary. *Civil Engineering and Environmental Systems*, *33*(3), 177–198. https://doi.org/10.1080/10286608.2016.1202932

Clark, L., Cools, R., & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition*, *55*(1), 41–53. https://doi.org/10.1016/S0278-2626(03)00284-7

Deditius-Island, H. K., & Caruso, J. C. (2002). An Examination of the Reliability of Scores from Zuckerman's Sensation Seeking Scales, Form V. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 7.

Ernst, M., & Paulus, M. P. (2005). Neurobiology of Decision Making: A Selective Review from a Neurocognitive and Clinical Perspective. *Biological Psychiatry*, *58*(8), 597–604. https://doi.org/10.1016/j.biopsych.2005.06.004

Fagin-Jones, S., & Midlarsky, E. (2007). Courageous altruism: Personal and situational correlates of rescue during the Holocaust. *The Journal of Positive Psychology*, *2*(2), 136–147. https://doi.org/10.1080/17439760701228979

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: Age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 709–730. https://doi.org/10.1037/a0014983

Gardoni, P., & Murphy, C. (2014). A Scale of Risk: Scale of Risk. *Risk Analysis*, *34*(7), 1208–1227. https://doi.org/10.1111/risa.12150

Gray, J. M., & Wilson, M. A. (2007). A detailed analysis of the reliability and validity of the sensation seeking scale in a UK sample. *Personality and Individual Differences*, *42*(4), 641–651. https://doi.org/10.1016/j.paid.2006.08.019

Hanley, J. A., Negassa, A., Edwardes, M. D. deB, & Forrester, J. E. (2003). Statistical Analysis of Correlated Data Using Generalized Estimating Equations: An Orientation. *American Journal of Epidemiology*, *157*(4), 364–375. https://doi.org/10.1093/aje/kwf215

Hanoch, Y., Johnson, J. G., & Wilke, A. (2006). Domain Specificity in Experimental Measures and Participant Recruitment: An Application to Risk-Taking Behavior. *Psychological Science, 17*(4), 300–304.

Harrison, J. D., Young, J. M., Butow, P., Salkeld, G., & Solomon, M. J. (2005). Is it worth the risk? A systematic review of instruments that measure risk propensity for use in the health setting. *Social Science & Medicine, 60*(6), 1385–1396. https://doi.org/10.1016/j.socscimed.2004.07.006

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. Retrieved from JSTOR.

Hopko, D. R., Lejuez, C. W., Daughters, S. B., Aklin, W. M., Osborne, A., Simmons, B. L., & Strong, D. R. (2006). Construct Validity of the Balloon Analogue Risk Task (BART): Relationship with MDMA Use by Inner-City Drug Users in Residential Treatment. *Journal of Psychopathology and Behavioral Assessment, 28*(2), 95–101. https://doi.org/10.1007/s10862-006-7487-5

Horvath, P., & Zuckerman, M. (1993). Sensation seeking, risk appraisal, and risky behavior. *Personality and Individual Differences*, *14*(1), 41–52. https://doi.org/10.1016/0191-8869(93)90173-Z

Knust, S., & Stewart, A. L. (2002). Risk-Taking Behaviour and Criminal Offending: An Investigation of Sensation Seeking and the Eysenck Personality Questionnaire. *International Journal of Offender Therapy and Comparative Criminology*, *46*(5), 586–602. https://doi.org/10.1177/030662402236742

Lauriola, M., Panno, A., Levin, I. P., & Lejuez, C. W. (2014). Individual Differences in Risky Decision Making: A Meta-analysis of Sensation Seeking and Impulsivity with the

Balloon Analogue Risk Task: Personality and Risky Decision Making. *Journal of Behavioral Decision Making*, *27*(1), 20–36. https://doi.org/10.1002/bdm.1784

Lejuez, C. W., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). The balloon analogue risk task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology*, *11*(1), 26–33. https://doi.org/10.1037/1064-1297.11.1.26

Lejuez, C. W., Aklin, W. M., Zvolensky, M. J., & Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of Adolescence*, *26*(4), 475–479. https://doi.org/10.1016/S0140-1971(03)00036-8

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., … Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84. https://doi.org/10.1037//1076-898X.8.2.75

Leon, A. C., & Heo, M. (2009). Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. *Computational Statistics & Data Analysis*, *53*(3), 603–608. https://doi.org/10.1016/j.csda.2008.06.010

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences*, *101*(45), 16075–16080. https://doi.org/10.1073/pnas.0406666101

Mittal, C., & Griskevicius, V. (2014). Sense of control under uncertainty depends on people's childhood environment: A life history theory approach. *Journal of Personality and Social Psychology*, *107*(4), 621–637. https://doi.org/10.1037/a0037398

Mohamed, R. (2006). The Psychology of Residential Developers: Lessons from Behavioral Economics and Additional Explanations for Satisficing. *Journal of Planning Education and Research*, *26*(1), 28–37. https://doi.org/10.1177/0739456X05282352

Newman, J. P., Patterson, C. M., & Kosson, D. S. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology*, *96*(2), 145–148. https://doi.org/10.1037/0021-843X.96.2.145

Parker, A. M., de Bruin, W. B., & Fischhoff, B. (2007). Maximizers versus satisficers: Decision-making styles, competence, and outcomes. *Judgment and Decision Making*, *2*(6), 9.

Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, *67*(5), 1035–1041. https://doi.org/10.1016/j.jbusres.2013.08.010

Phelps, C. E. (1987). Risk and Perceived Risk of Drunk Driving among Young Drivers. *Journal of Policy Analysis and Management*, *6*(4), 708–714. https://doi.org/10.2307/3323529

Quevedo, K., Benning, S. D., Gunnar, M. R., & Dahl, R. E. (2009). The Onset of Puberty: Effects on the Psychophysiology of Defensive and Appetitive Motivation. *Development and Psychopathology*, *21*(1), 27–45. https://doi.org/10.1017/S0954579409000030

Rao, H., Korczykowski, M., Pluta, J., Hoang, A., & Detre, J. A. (2008). Neural correlates of voluntary and involuntary risk taking in the human brain: An fMRI Study of the Balloon Analog Risk Task (BART). *NeuroImage*, *42*(2), 902–910. https://doi.org/10.1016/j.neuroimage.2008.05.046

Roberti, J. W., Storch, E. A., & Bravata, E. (2003). Further Psychometric Support for the Sensation Seeking Scale-Form V. *Journal of Personality Assessment*, *81*(3), 291–292. https://doi.org/10.1207/S15327752JPA8103_12

Ryb, G. E., Dischinger, P. C., Kufera, J. A., & Read, K. M. (2006). Risk perception and impulsivity: Association with risky behaviors and substance abuse disorders. *Accident Analysis & Prevention*, *38*(3), 567–573. https://doi.org/10.1016/j.aap.2005.12.001

Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends in Cognitive Sciences*, *15*(1), 11–19. https://doi.org/10.1016/j.tics.2010.10.002

Slovic, P. (1964). Assessment of risk taking behavior. *Psychological Bulletin*, *61*(3), 220–233. https://doi.org/10.1037/h0043608

Smith, B. W., Mitchell, D. G., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J., & Ernst, M. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *NeuroImage*, *44*(2), 600–609. https://doi.org/10.1016/j.neuroimage.2008.08.016

Spurrier, M., & Blaszczynski, A. (2014). Risk Perception in Gambling: A Systematic Review. *Journal of Gambling Studies*, *30*(2), 253–276. https://doi.org/10.1007/s10899-013-9371-z

Upton, D. J., Bishara, A. J., Ahn, W.-Y., & Stout, J. C. (2011). Propensity for risk taking and trait impulsivity in the Iowa Gambling Task. *Personality and Individual Differences*, *50*(4), 492–495. https://doi.org/10.1016/j.paid.2010.11.013

Verdejo-Garcia, A., Benbrook, A., Funderburk, F., David, P., Cadet, J.-L., & Bolla, K. I. (2007). THE DIFFERENTIAL RELATIONSHIP BETWEEN COCAINE USE AND MARIJUANA USE ON DECISION-MAKING PERFORMANCE OVER REPEAT

TESTING WITH THE IOWA GAMBLING TASK. *Drug and Alcohol Dependence*, *90*(1), 2–11. https://doi.org/10.1016/j.drugalcdep.2007.02.004

Wang, X. T., Kruger, D. J., & Wilke, A. (2009). Life history variables and risk-taking propensity. *Evolution and Human Behavior, 30*(2), 77–84. https://doi.org/10.1016/j.evolhumbehav.2008.09.006

Wardle, M. C., Gonzalez, R., Bechara, A., & Martin-Thormeyer, E. M. (2010). Iowa Gambling Task performance and emotional distress interact to predict risky sexual behavior in individuals with dual substance and HIV diagnoses. *Journal of Clinical and Experimental Neuropsychology*, *32*(10), 1110–1121. https://doi.org/10.1080/13803391003757833

Weber, E. U., Blais, A.-R., & Betz, N. E. (2002). A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, *15*(4), 263–290. https://doi.org/10.1002/bdm.414

White, T. L., Lejuez, C. W., & de Wit, H. (2008). Test-Retest Characteristics of the Balloon Analogue Risk Task (BART). *Experimental and Clinical Psychopharmacology*, *16*(6), 565–570. https://doi.org/10.1037/a0014083

Wills, T. A., Vaccaro, D., & McNamara, G. (1994). Novelty seeking, risk taking, and related constructs as predictors of adolescent substance use: An application of Cloninger's theory. *Journal of Substance Abuse*, *6*(1), 1–20. https://doi.org/10.1016/S0899-3289(94)90039-6

Wu, P., Fang, Y., Guan, Z., Fan, B., Kong, J., Yao, Z., … Hoven, C. W. (2009). The Psychological Impact of the SARS Epidemic on Hospital Employees in China: Exposure, Risk Perception, and Altruistic Acceptance of Risk. *The Canadian Journal of Psychiatry*, *54*(5), 302–311. https://doi.org/10.1177/070674370905400504

Xu, S., Korczykowski, M., Zhu, S., & Rao, H. (2013). Assessment of risk-taking and impulsive behaviors: A comparison between three tasks. *Social Behavior and Personality*, *41*(3), 477–486. https://doi.org/10.2224/sbp.2013.41.3.477

Young, M. E., Webb, T. L., Rung, J. M., & McCoy, A. W. (2014). Outcome Probability versus Magnitude: When Waiting Benefits One at the Cost of the Other. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0098996

Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge University Press.

Zuckerman, M. (1996). Item revisions in the Sensation Seeking Scale Form V (SSS-V). *Personality and Individual Differences*, *20*(4), 515. https://doi.org/10.1016/0191-8869(95)00195-6

Zuckerman, M. (2007). The sensation seeking scale V (SSS-V): Still reliable and valid. *Personality and Individual Differences*, *43*(5), 1303–1305. https://doi.org/10.1016/j.paid.2007.03.021

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology*, *46*(1), 139–149. https://doi.org/10.1037/0022-006X.46.1.139

Curriculum Vitae

Edward A Smith Jr.

E-mail: esmithjr995@gmail.com

Education

2017- Clinical Psychology PhD student
University of Nevada Las Vegas
Current GPA: 3.97

2017 **B.A. Psychology,** *magna cum laude*
University of Notre Dame

Senior Thesis: *Early Childhood Experience, Religious Upbringing, and the Development of Psychopathic Traits*
Supervisor: Darcia Narvaez, PhD

GPA: 3.86
Major GPA: 3.94

Grants and Honors

2019 UNLV Patricia Sastaunik Scholarship
Award: $2500

UNLV Summer Graduate Research Fellowship
Award: $7000

2018 UNLV Patricia Sastaunik Scholarship
Award: $2500

UNLV Summer Graduate Stipend
Award: $2500

UNLV Summer Session Scholarship
Award: $2000

UNLV Clinical Student Committee, Diversity Liaison

2017 UNLV Clinical Student Committee, Treasurer

University of Notre Dame Psychology Department Senior Recognition Award

Psi Chi (Psychology academic honor society)

2016 Undergraduate Research Opportunity Program (UROP) Summer Comprehensive Grant
Award: $4000

Institute for Scholarship in the Liberal Arts (ISLA) Conference Presentation Grant
Awards: $4500 (3@$1500 each)

2015 UROP Summer Comprehensive Grant
Award: $4500

2015-16 ISLA Award for Substantial Undergraduate Research

2014-2017 Dean's List, College of Arts and Letters, University of Notre Dame

Research Experience

Graduate Assistant/Lab Manager
Department of Psychology
University of Nevada Las Vegas
2017- Present

Responsibilities:

Since the Fall 2017 semester, I have worked with Dr. Stephen Benning in his Psychophysiology of Emotion and Personality (PEP) lab. I was instructed in the use of EEG and other psychophysiological equipment and software, and I have coordinated and supervised data collection for four different studies. Since Fall 2018, I have also been responsible for training new undergraduate Research Assistants and managing the lab in general. Additionally, I have been compiling several years' worth of study data for two other projects Dr. Benning is planning on publishing, for each of which I will serve as one of the principal authors. One of these papers is currently undergoing revision for publication to the *Journal of Abnormal Psychology*. The paper details the need for increased transparency and preregistration practices in clinical research and provides a guide on how to do so for academic psychologists.

Research Assistant
Department of Psychology
University of Notre Dame
2014-2017

Responsibilities:

While working in Dr. Darcia Narvaez's Moral Psychology lab, I designed and conducted several new studies for the lab. The first such project compared the usefulness of simple, 1-item assessments of religiosity/spirituality to more complex, multidimensional measures of these constructs, including of a new measure that was primarily of my own design. Under Dr. Narvaez's supervision, I performed the literature review, constructed and edited the new measure, wrote grant and conference proposals, collected pilot and validation study data, performed statistical analyses of data (including factor analysis), and created a poster and manuscript for the completed project.

The second project I completed with Dr. Narvaez focused on the relationship between moral personality/character and religiosity. Duties included programming a lexical decision task, constructing an online survey, running participants, instructing/supervising other lab members in data collection, and statistical analysis of data.

Finally, my Summer 2016 and 2016-2017 academic year responsibilities were split between my own senior thesis on the relationship between early experience and psychopathic traits, for which Dr. Narvaez served as supervisor, and statistically validating a new Civic Narcissism measure for a potential journal publication. I performed background research, wrote manuscripts, conducted statistical analyses (including factor analysis), constructed surveys, and ran participants.

Publications and Presentations

Benning, S., Bachrach, R., Freeman, A., Smith, E., & Wright, A. (Under Revision). The Registration Continuum in Clinical Science: A Guide toward Transparent Practices. *Journal of Abnormal Psychology.*

Smith, E., Baggio, M., Pedregon, C., & Benning, S. (May 2019). *A Cold, Fearful Touch: Social Support and Psychopathic Meanness.* Biennial Conference for the Society for the Scientific Study of Psychopathy, Las Vegas, NV.

Smith, E., & Benning, S. (May 2019). *The Assessment of Physical Risk Taking and Fearless Dominance in Psychopathy.* Biennial Conference for the Society for the Scientific Study of Psychopathy, Las Vegas, NV.

Smith, E., & Narvaez, D. (August, 2017). *The Evolved Developmental Niche, Religious Upbringing and the Development of Psychopathic Traits.* The Annual Convention of the American Psychological Association, Washington, DC.

Smith, E., & Narvaez, D. (April, 2017). *Early Experience, Civic Narcissism and Voting Preference in the 2016 Election.* Midwest Political Science Association Annual Meeting, Chicago, IL.

Smith, E., Kurth, A., & Narvaez, D. (December, 2016). *Moral Chronicity and Religiosity.* The Association for Moral Education Annual Conference, Cambridge, MA.

Smith, E., Kurth, A., & Narvaez, D. (August, 2016). *Examining the Dimensions of Religiosity: Comparing single-item and multifactor models.* The Annual Convention of the American Psychological Association, Denver, CO