Perceptions of effort and risk assessment

by

Lisa Lynn Vangsness

B.A., University of Iowa, 2013
B.S., University of Iowa, 2013

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2017

Approved by:

Major Professor
Michael E. Young

Copyright

Abstract

Although risky decision-making tasks present some a priori risk (i.e., base-rate), decision makers often have an opportunity to modify this level of risk through their behaviors. Broadly speaking, risk can be modified by assigning additional resources to an ongoing task or by engaging in specific risk-mitigation strategies before or after the risky decision is made. The modification of risk requires ongoing awareness of task demands, resource constraints, and risk-mitigation strategies that can be used to adapt behavior over time. This thesis explores risk modification that occurs during difficult tasks. Difficult tasks hold greater risks because they fall at the edge of the decision maker's abilities and are likely to require a greater number of resources to overcome. As resources are engaged they become unavailable for other tasks or strategies to cope with changing task demands. I studied how individuals monitor risks and develop risk mitigation strategies using a videogame task designed to mirror contingencies that would be encountered in the real world. Results from two experiments that involve this task suggest that decision-makers adequately monitor and develop active strategies for dealing with risks. These strategies change over time and vary as a function of task difficulty and experience.

key words: RDOs, risk mitigation, dynamic decision making, metacognitive judgments, difficulty

Table of Contents

List of Figures

List of Tables

Acknowledgements

"*1. If an undertaking was easy, someone else already would have done it.*

2. If you follow in another's footsteps, you miss the problems really worth solving."

- John Chatterton

Early in the process of writing my thesis, I encountered John Chatterton's Rules and Observations in a book about deep sea diving. At the time, I was still early in the research process and had not yet encountered the difficulties associated with studying human behavior in rapidly changing environments. As my research grew, I became indebted to the friends and family who encouraged me through each phase of this project.

Although I find that appreciation is best conveyed in person, thank you for helping me to identify and pursue the problems that are really worth solving.

Perceptions of Effort and Risk Assessment

Physical and cognitive resources are necessary for survival (Krebs, 1977; Shaw & Shaw, 1977) and are integral to growth and exploration (Kurzban, Duckworth, Kable, & Myers, 2013). These resources are limited and finite across a variety of domains, including food (Krebs, 1977), consumer goods (Cobb & Douglas, 1928), attention (Wickens, 2002), and working memory (for a review see Cowan, 2010). The limited nature of physical and cognitive resources requires organisms to strategically allocate these resources during the pursuit of specific goals, such as energetic homeostasis (Krebs, 1977) or the successful completion of multiple tasks (Wickens, 2002). Although much attention has been dedicated to evaluating organisms' ability to monitor and manage physical resources within a dynamic context, similar research on cognitive resources is noticeably lacking. In this paper, I will outline a model of cognitive resource management and test predictions related to two components of this model, specifically humans' ability to monitor task difficulty and to evaluate and mitigate risks with respect to changing internal and external conditions.

A Model of Cognitive Resource Management

The strategic allocation of cognitive resources in pursuit of a goal requires an awareness of environmental changes and the ability to self-assess cognitive resource availability over time, a process referred to as metacognition. Metacognitive knowledge is acquired after an initial allocation of cognitive resources towards elements of a task. As the task progresses, individuals acquire metacognitive knowledge about their abilities, the demands and constraints of the task, and the efficacy of the strategies they employed in pursuit of task-related goals. This metacognitive knowledge can be used to iteratively "select, evaluate, revise, and abandon cognitive tasks, goals, and strategies in light of their relationships with one another and with [one's] own abilities and interests with respect to that enterprise" (Flavell, 1979, p. 908). In other words, cognitive resources are reallocated in response to new information several times throughout the duration of a task.

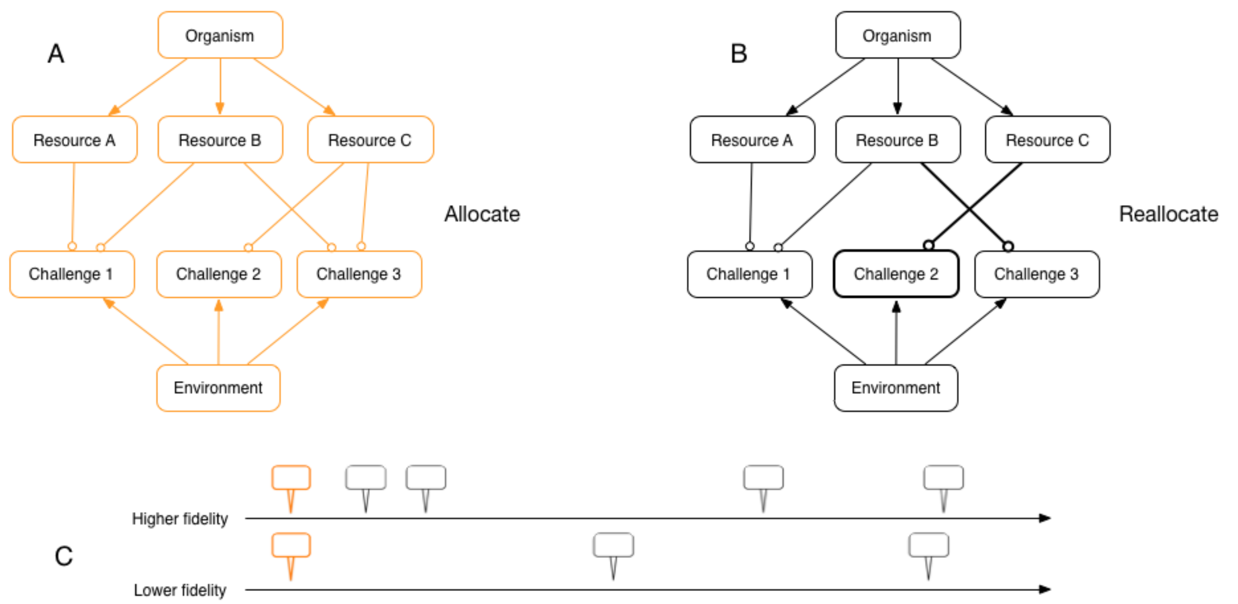

Figure 1. An iterative model of cognitive resource allocation.

The process of resource allocation is illustrated in Figure 1, which depicts resource allocation across three cognitively demanding tasks (Challenges 1, 2, and 3). Initially, a decision-maker uses existing internal and external information to estimate and distribute resources among tasks (see Figure 1A). As tasks progress, the decision-maker gathers additional information that may influence cognitive resource allocation. For example, Figure 1B depicts a decision to shift cognitive Resource C away from Challenge 3 to Challenge 2 and to increase the amount of Resource B allocated to address Challenge 3. This shift could be motivated by task characteristics or by an increase in the difficulty level of Challenge 2. Each reallocation of cognitive resources has the potential to alter performance on other tasks. Thus, reallocation is an iterative process that occurs throughout the course of task performance.

Because this type of cognitive resource management occurs within a dynamic environment, decision-makers must update their representation of the environment frequently enough to support their predictions about task difficulty, cognitive resource allocation, and the resulting probability of success and/or failure (i.e., risk) associated with potential actions. The fidelity of this representation depends on the frequency of an individual's difficulty and risk assessments and the rate at which the environment changes. In general, higher fidelity representations result in more accurate metacognitive judgments, the appropriate allocation of cognitive resources, and successful risk mitigation (see Figure 1C); however, not all decisions require high fidelity information (e.g., making long-term business plans).

The process of resource reallocation becomes particularly important when there is great disparity between an individual's abilities and a task's demands. In these circumstances, strong metacognitive skills can be used to assess risks and to reallocate resources appropriately (per Brunswik, 1956). For instance, decision-makers that recognize when a task nears or exceeds the level of their skills and abilities can address this challenge through the use of strategies such as: seeking additional help from a decision aid (Slovic, Fischhoff, & Lichtenstein, 1977) or colleague (for a review of advice-taking see Bonaccio & Dalal, 2006), conserving cognitive resources during less effortful tasks (similar to the conservation of physical resources observed by Smith, Marcoroa, & Coutts, 2015), improving the accuracy of one's perceptions of effort (e.g., Bjorkman, 1994) and risk assessment (e.g., Erev et al., 2009; Camilleri & Newell, 2013), and engaging in additional behaviors to reduce the risks associated with negative outcomes (i.e., Risk Diffusing Operators per Huber, 2012).

For example, an air traffic controller must maintain cognitive vigilance throughout a shift as well as respond quickly to acute challenges that threaten the security of passengers and aircraft. These challenges involve many dimensions of difficulty: an unanticipated weather event may tax controllers' working memory as they struggle to maintain alternative flight paths for many aircraft, but an errant pilot may require controllers to emphasize speed and allocate attentional resources to quickly clear the airspace. Regardless the nature of the challenge, a controller's goal is to minimize the negative impact of an event outcome. The most successful air traffic controllers will identify the dimensions of difficulty presented by a task and balance these characteristics with their own abilities, skills, and resources. This may be accomplished by utilizing available tools (e.g., radar) and colleagues, reallocating cognitive resources toward addressing effortful tasks, engaging in intensive simulation and on-the-job training, and actively seeking alternative strategies to reduce risks during difficult and/or risky situations.

The skills and awareness required to accurately predict task difficulty (e.g., Bjorkman, 1994), risk (e.g., Erev et al., 2010; Camilleri & Newell, 2013), and cognitive resource allocation (Kanfer & Ackerman, 1989) can be learned, especially when individuals gain repeated experience and are highly motivated. Still, even task "experts" retain some degree of miscalibration (for a review see Koehler, Brenner, & Griffin, 2002; but see also Keren, 1987; Murphy & Winkler, 1977). This poses a conflict: unsuccessful risk mitigation strategies may occupy cognitive resources and thereby reduce an individuals' ability to mitigate future risks – including those which were unsuccessfully diminished.

For this reason, metacognitive awareness and appropriate resource allocation is particularly relevant to workers who must sustain cognitive vigilance over long shifts, such as nurses and police officers. People working long shifts may not notice changes in their level of attentional focus, memory, or physical energy. The effects of fatigue on decision-making performance are widely noted. For example, fatigue increases on-the-job incidents among air traffic controllers (FAA, 2009), biases police officers toward the use of force (Kop & Euwema, 2001), and increases the likelihood of nursing errors (for a review see Rogers, 2008). These examples underscore the need to better understand the relationship between metacognition, resource allocation, and performance so as to provide workers with strategies that reduce losses and successfully mitigate risk.

This paper will focus on the relationship between perceptions of effort and risk-mitigation strategies on performance in a dynamic task. I will begin by outlining three theoretical perspectives (Rational Choice, Prospect Theory, and Ecological Rationality) on resource allocation in risky environments, and consider the generalizability of these findings to decision making in dynamic environments. I will conclude with a review of existing literature on effortful choice and will advance the perspective that metacognitive processes, such as perceptions of difficulty, play a key role in an individual's ability to mitigate risks during difficult or challenging tasks.

Theoretical Perspectives on Decision Making

In an ideal environment, the probability and value of outcomes are known and the relationship between one's resources and desired outcomes is well-understood. Decision-makers that engage in accurate metacognitive monitoring can develop sustainable resource allocation and risk mitigation strategies that maximize gains and minimize losses. Yet many researchers acknowledge that the ideal environments that facilitate such strategies do not exist and that decision-makers appeal to alternative strategies to make decisions. Here, I will review three perspectives on decision making that predict different patterns of resource allocation and risk mitigation.

Rational Choice Theory

Rational Choice Theory suggests that a decision-maker considers the consequences of all possible actions and outcomes and compares them to available resources before pursuing an outcome that maximizes total gains and minimizes losses (Arrow, 2004). Specifically, actions are evaluated with respect to their expected value or subjective utility, a function of both the value of and the probability with which an outcome will occur (von Neumann & Morgenstern, 1947). These evaluations are then used to create a preference ordering of potential actions that remains consistent over time (Arrow, 2004).

Because Rational Choice Theory assumes that decision-makers consistently maximize gains and minimize losses, resource allocation and risk mitigation strategies should be driven by the expected utility of the strategies that are available, irrespective of individual differences in metacognitive abilities. Thus, Rational Choice Theory predicts that consistent and optimal resource allocation and risk mitigation strategies will always occur. Because Rational Choice Theory reflects optimal decision making, many researchers believe it lacks utility (for a review of Rational Choice Theory and its limitations see Burns & Roszkowska, 2016) and does not accurately predict human behavior (e.g., Harsanyi, 1994). These researchers argue that constraints, such as time and memory capacity, undermine a decision-maker's ability to form an accurate representation of the decision space (Simon, 1956). This revelation has led researchers to propose alternative models of decision making that consider systematic biases and misestimations that may underlie human behavior.

Prospect Theory

One systematic bias that seems to affect decision making under conditions of risk is the tendency to overweight losses and underweight gains. Prospect Theory predicts that this asymmetry leads individuals to make riskier choices when they anticipate losses and to prefer more certain outcomes when anticipating gains. This behavior can be modeled using the equation:

$$V = \sum_{i=1}^{n} \pi(p_i) v(x_i) \tag{1}$$

where v_i is the value of a particular outcome associated with a decision, x; p_i is the probability of this outcome occurring; and π is a decision weight that reflects humans' propensity to overweight unlikely events and underweight those that are likely to occur. The perception of an outcome as a loss or gain is determined by a reference point, which is commonly described as an individual's current assets or as an "expectation or aspiration level" they hope to attain. Gains follow a concave function with diminishing return, while losses follow a convex function that strongly overvalues small losses. This value is further modified by a decision weighting function which reflects humans' tendency to overweight the probability of unlikely outcomes and to underweight the probability of more certain outcomes (Kahneman & Tversky, 1979).

In decisions involving resource allocation and risk mitigation, losses take two forms: resources are temporarily unavailable when they are engaged by a task or risk mitigation strategy, and collateral losses are incurred when decision outcomes are unsuccessful. If decisions are primarily driven by an avoidance of any type of loss, decision-makers may pursue "easier" strategies that occupy fewer resources even if those strategies increase the risk of collateral losses. In a dynamic environment, this behavioral pattern may manifest as a cycle of poor decisions: a decision-maker allocates fewer resources to challenging tasks, which increases the probability of failure and encourages the decision maker to engage in suboptimal resource allocation and risk mitigation strategies that further promote loss. For example, a police officer nearing the end of a shift may insufficiently attend to a suspect's non-verbal cues and misinterpret anxiety as a "loss." This perception may lead the officer to look more favorably upon riskier choices, such as an escalation of force which could further upset the suspect. Thus, Prospect Theory predicts that decision-makers will pursue "easier" tasks and will pursue riskier mitigation strategies when losses are unavoidable.

Support for Prospect Theory originates in research conducted within a static environment. In these experiments, participants read a vignette before deciding to pursue one of two actions that were identical in expected value, but that differed in how this value was expressed. That is, one action represented a "risky" choice while the other offered certainty. Take, for instance, the options offered to Tversky and Kahneman's (1985) participants that were instructed to prevent the outbreak of a deadly Asian disease. These participants learned that "[i]f Program C is adopted, 400 people will die. If Program D is adopted there is a 1/3 probability that nobody will die, and 2/3 probability that 600 people will die." The expected value of each option can be calculated by multiplying the probability with which an outcome will occur by its overall value. Because the expected value of Program D, in terms of "people [that] will die," is $\frac{2}{3} \times 600 = 400$, the two options are equal in objective value. Still, participants prefer riskier Program D in the face of loss. When the wording of the question is expressed in terms of the number of "people [that] will be saved," participants reverse their preference and become more likely to choose the certain gain (Kahneman & Tversky, 1985).

The findings from empirical studies involving hypothetical scenarios, such as Kahneman and Tversky's (1985) Asian Disease Problem, are largely assumed to generalize to everyday decisions involving everything from sunscreen application (Detweiler, Bedell, Salovey, Pronin, & Rothman, 1999) to the under-utilization of annuities (Brown, Kling, Mullainathan, & Wrobel, 2008). Yet emerging research suggests that external framing may be context-dependent (e.g., Neale, Huber, & Northcraft, 1987) or limited to descriptive contexts (Rakow & Newell, 2010; Hau, Pleskac, & Hertwig, 2010) and that decision-makers may create their own hedonic frames (e.g., Wang, 2004). Thus, it is unclear whether Prospect Theory is a viable theoretical model for decisions that occur within dynamic environments.

Ecological Rationality and the Description-Experience Gap

A third alternative is that humans evolved to interpret probabilities in specific contexts and cannot accurately do so when risks are presented as single-event probabilities (Cosmides & Tooby, 1996). This theory suggests that description-based decision making scenarios are difficult to interpret because they associate a single probability with each outcome. The reliability of described outcome estimates is relatively unknown and cannot be updated with new information gained through experience. Thus, the risky decision making of Prospect Theory arises from the artificial context in which the problem is encountered (e.g., Cosmides & Tooby, 1990; Gigerenzer, 1991; Barron & Erev, 2003).

In dynamic environments, information is encountered iteratively and representations can be updated as individuals integrate feedback with information about their internal and external environment to adapt behavior to more closely align with optimal strategies. When decision-makers actively experience losses and gains as a series of interactive gambles, the preferences exhibited in static environments reverse such that participants prefer certain losses and risky gains (Rakow & Newell, 2010) and behave conservatively when faced with losses (Barron & Erev, 2003). Although this “description-experience gap” is well-documented (for a review see Camilleri & Newell, 2013), the underlying cause is unclear. Researchers whose findings strongly support the description-experience gap have appealed to distributional properties, cognitive limitations, presentation formats (e.g., Hau, Pleskac, & Hertwig, 2010), and the salience of individual outcomes (Spetch & Ludvig, 2011) to explain the phenomenon. Other researchers propose that working memory influences decision-making, but not enough to effectively or consistently reverse choice preference (e.g., Rakow, Demes, & Newell, 2008).

Although there is considerable diversity in how this preference reversal is explained, researchers largely agree the phenomenon is related to the uncertainty present in decisions made from experience. A participant who encounters a descriptive decision is provided with a known probability with which each outcome will occur. This degree of certainty is rarely available during dynamic decision making tasks because working memory limits the number of samples that can be used to estimate the probabilities associated with each decision's outcome (Kareev, 2000). This distinction between *a priori* (i.e., known) and *statistical* (i.e., experienced) probabilities (Knight, 1921) has been modeled across studies involving decks of cards (Hau, Pleskac, Kiefer, & Hertwig, 2008), virtual monetary gambles (Camilleri & Newell, 2013; Ludvig & Spetch, 2011), and point-based gambles (Rakow, Demes, & Newell, 2008). The findings of these studies all provided at least some support for the description-experience gap.

Thus, the ecological rationality account would seem to suggest that while decision-makers will not minimize losses and maximize gains flawlessly (per Rational Choice Theory), they should be sensitive to changes in the environment and capable of monitoring outcome success. Additionally, decision-makers should exhibit risk aversion in the face of losses (per Barron & Erev, 2003) and develop conservative resource allocation and risk mitigation strategies as situations become more challenging. However, the experiments that support this prediction do not allow individuals to influence the probability or risks associated with future outcomes. Extending this possibility allows participants to change the contingencies of the task and may alter subsequent behavior.

Perceptions of Difficulty and Risk

Regardless of the underlying mechanism, all three competing perspectives on resource allocation and risk mitigation assume decision-makers estimate and track performance losses and gains. “Difficult” tasks that near or exceed the level of an individual’s skills, abilities, or available cognitive resources result in poorer performance than do less challenging tasks. Therefore, estimations of task difficulty and performance should underlie resource allocation and risk mitigation decisions. In the section that follows, I will briefly discuss how perceptions of difficulty inform risk mitigation strategy and suggest potential patterns of behavior that emerge from decision-makers engaged in difficult tasks.

Perceptions of Difficulty

Perceptions of difficulty are important because they serve as an indicator of the number of physical and cognitive resources that remain available for completing a task (Kahneman, 1973; Kanfer & Ackerman, 1989; for a review see Kurzban, 2016). Perceptions of difficulty can also help individuals select and prioritize tasks and strategies that will allow them to achieve their goals (Kurzban, Duckworth, Kable, & Myers, 2013). Traditionally, this is reflected in an aversion towards effortful tasks (Hull, 1932): engaging fewer resources in a task reserves them for future deployment. Although this behavior is often viewed as detrimental to one’s safety (Sigurdsson, Taylor, & Wirth, 2013) and achievement (Ostaszewski, P., Bąbel, P., & Swebodzińsk, 2013), it can be adaptive under certain conditions.

An aversion towards effortful tasks can be reframed as an aversion towards the under- or over-allocation of resources (Kurzban, Duckworth, Kable, & Myers, 2013). Effortful tasks that are unnecessary or insufficiently rewarded should be avoided because they consume resources that could be otherwise allocated toward more rewarding endeavors. Choosing to engage in these tasks represents an over-allocation of cognitive resources. Similarly, if a task nears or exceeds the level of an individual's skills, abilities, or available cognitive resources, no amount of resource allocation will sufficiently improve performance: it will be consistently under-allocated. In accordance with these principles, the value of effort is discounted rapidly (e.g., Freeman, Morgan, Brandner, Almahdi, & Curran, 2013; Nishiyama, 2014; Hartmann, Hager, Tobler, & Kaiser, 2013; Sugiwaka & Okouchi, 2004; for a review of the animal literature see Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006) and people quickly become demotivated by very difficult tasks (for a review, see Locke & Latham, 2002).

Research on effortful decision making often occurs in a laboratory context that does not provide individuals with recourse to strategies that reduce risk by changing the difficulty level of the situation or improving the decision-maker's abilities. As such, these experiments may provide limited information about how and when people choose to pursue effortful tasks. There is a great need to investigate human behavior in situations that involve realistic scenarios, particularly those that provide decision-makers with an opportunity to mitigate risks (Huber, 2012), and to understand how risk mitigation decisions relate to perceptions of task difficulty.

Risk Mitigation Strategies: Effective use of Risk Diffusing Operators (RDOs)

In realistic situations, decision-makers can mitigate risk by employing a risk diffusing operator (RDO) before or after a loss occurs (Huber & Huber, 2008). For example, a dentist might reduce a patient's risk of tooth decay by applying a dental sealant, which forms a shield that protects teeth from bacteria and food waste (NIH, 2016). Alternatively, the dentist may wait and halt decay with a dental amalgam filling after the decay has occurred (USFDA, 2015). Pre-event RDOs, such as the dental sealant, have a high initial cost but a certain outcome (i.e., they prevent a negative outcome from occurring). Post-event RDOs, such as the dental amalgam, do not have an initial cost but are inherently riskier because they do not change the probability of a negative outcome but seek to reduce its impact if it occurs. In the event that a negative outcome occurs, post-event RDOs are associated with greater overall costs because decision makers experience a negative outcome and must reallocate resources to tasks and strategies that will help them recover from that loss.

RDOs can be used to mitigate the risks associated with challenging tasks. That is, a decision-maker can seek to prevent the negative consequences of poor performance in a challenging task by employing a pre-event RDO. Similarly, decision-makers can rely on post-event RDOs when engaged in low-risk, simple tasks, and free resources for alternative pursuits. In both of these situations, the best decisions are made by those with strong metacognitive skills that allow them to gauge the difficulty level of the task and, in turn, which RDO is most appropriate.

The competing perspectives on resource allocation and risk mitigation predict different patterns of RDO selection by decision-makers engaged in difficult tasks. Ecological rationality predicts that decision-makers' risk mitigation strategy is driven by previous successes and failures with RDOs: individuals will consistently use risk mitigation strategies that are most effective given their skills and abilities. As tasks become increasingly difficult and losses become more likely, decision-makers should gradually shift towards taking preventative measures (pre-event RDOs) to reduce their losses. In less-challenging situations, decision-makers should rely on post-event RDOs that only expend resources in the unlikely event of a loss. In contrast, Prospect Theory predicts that decision-makers will be strongly motivated by loss avoidance and will resist the certain loss associated with a pre-event RDO in favor of the probabilistic loss of a post-event RDO. Preventative actions should only be taken if they are not perceived as costly or when overshadowed by the expected value of the loss associated with post-event RDOs. A third alternative, supported by Rational Choice Theory, is that decision-makers' risk mitigation strategy will be directly related to anticipated and experienced losses. That is, decision makers will utilize pre-event RDOs when tasks are challenging and the probability of loss is high but prefer post-event RDOs in alternative circumstances.

The Effect of Time-on-task on the Decision-making Process

Research exploring the relationship between time-on-task and strategic planning underscores the importance of selecting appropriate risk mitigation strategies during sustained, challenging tasks. Evidence from EEG research suggests that as time-on-task increases participants become less likely to take corrective actions to reduce future errors (Lorist, Boksem, & Ridderinkhof, 2005) and become more likely to perseverate in the use of inappropriate strategies (van der Linden, Frese, & Meijman, 2003). Together, this evidence seems to suggest that time-on-task will negatively impact task performance and reduce decision-makers' ability to adapt risk mitigation strategies in response to changes in performance. Given that fatigue is responsible for the loss of about 4 hours of productivity per work week per person (Ricci, Chee, Lorandeau, & Berger, 2007) and nearly doubles an individual's risk of being involved in a workplace accident (Swaen, Amelsvoort, Bultmann, & Kant, 2003), it is important to explore the effects of time-on-task on risk mitigation strategy. Addressing the relationship between effort, loss, and risky choice in a realistic environment requires the use of a dynamic experimental task.

The Experimental Environment

The experiments were conducted in a virtual environment developed with the Unity Game Engine (Unity, 2016), a freely available software package that allows users to design videogames using existing or custom-made graphic elements. Designing a videogame was advantageous because it allowed for the selection and control of all properties of the virtual environment, with the exclusion of the player’s behavior. In this section, I will describe the experimental environment in brief detail; readers interested in a more thorough description may view a video clip at https://www.k-state.edu/psych/people/graduatestudents/vangsness.html.

The Videogame Task

In the videogame, participants controlled the avatar of a young boy and explored a large-scale version of a bedroom while avoiding the stuffed-animal zombies that aimed to attack the avatar. Participants guided the avatar through the game space using the arrow/ASWD keys and could eliminate enemies using a laser cap gun that was controlled with the computer mouse. The goal of the videogame was to successfully progress through as many levels as possible within a 40 min session. Experimenters encouraged participants to pursue this goal by stating that “most participants clear eight levels before the session ends.”

To achieve the goal, participants needed to prioritize their performance in the game because death was a time-costly event that reduced their likelihood of successfully completing 8 levels. Death occurred once the avatar’s health was completely depleted by enemy attacks, which occurred each time an enemy character touched the avatar (see Figure 2). When the avatar died, the game was paused and a loading screen appeared. After approximately 30 s, the game resumed and the avatar was resurrected (with full health) in a random location within the game space.

Figure 2. Enemy attacks negatively impact the avatar's health, as reflected by a full (top) and partially depleted (bottom) health bar.

Much like in a traditional videogame, participants advanced to a new level (or block of trials) by eliminating enemies. Once participants eliminated 30 enemies from the game space, the game was paused and a loading screen appeared. The level reloaded and the game resumed after approximately 7 s. This new level could be easier or harder than the last: unlike a traditional videogame, the degree of difficulty was randomly assigned at the beginning of each level.

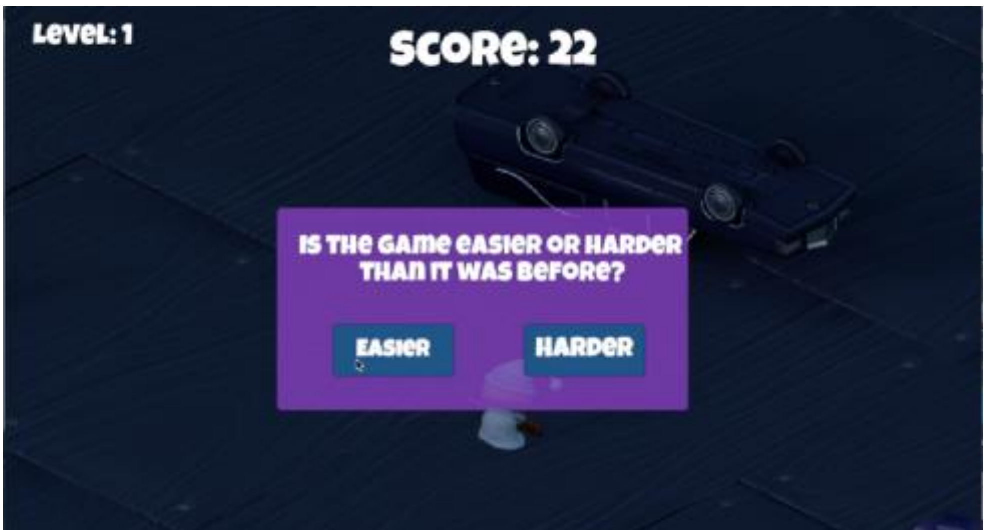

Figure 3. Participants indicated their perceptions of difficulty using a dialogue box.

At the beginning of each level, and every subsequent 2 min, a pop-up window invited participants to indicate whether the videogame was “easier” or “harder” than it was before (see Figure 3). This format allowed participants to make comparative assessments without interpreting scale anchors or values and without making assumptions about the scaling of subjective perceptions of difficulty (for additional information, see Böckenholt, 2004). Once participants selected one of the two options with the computer mouse, the pop-up window disappeared from the screen. Gameplay remained paused for 3 s before and after the pop-up window’s appearance to the reduce the performance costs associated with task interruption (Altmann & Trafton, 2007).

After 40 min of gameplay, the videogame ended and participants completed a demographic questionnaire that included questions about sex and videogame experience. Participants also completed a modified version of the Game Engagement Questionnaire (Brockmyer et al., 2009; see Appendix A). Researchers debriefed participants and allowed them to leave the session.

Task difficulty. Task difficulty was manipulated by adjusting one characteristic of the enemy characters' behavior at the start of each level. All other characteristics of the enemy characters' behavior remained constant during the session. For example, participants in the "speed" condition saw the enemy characters' rate of movement change from one level to the next but did not experience changes in the enemy characters' population rate, attack damage, hit points, or line-of-sight. Changes in enemy characters' behavior were randomly determined at the start of each level using an algorithm developed during the pilot study. For a full list of the conditions included in these thesis experiments, see Table 1.

Tutorial level. The videogame task included a tutorial level to help participants become familiar with the controls of the game. The tutorial level was identical to the videogame task in all respects but contained only three enemies. Players did not progress to the first level of the videogame task until they successfully eliminated all three enemies from the tutorial level. Experimenters encouraged participants to ask questions during the tutorial.

Table 1

The videogame task had five possible conditions.

Condition	Description	Unit of measurement
Population	The rate at which enemies appeared in the game space.	seconds
Damage	The amount of damage that enemies could inflict upon the avatar in a single attack.	hit points
Strength	The number of hit points enemies had when they first appeared in a level.	hit points
Line-of-sight	How near the avatar needed to be before an enemy began to move towards the avatar.	Unity units
Speed	The speed at which enemy characters could travel.	Unity units

Note. Unity units are an arbitrary measure that can be used to scale game objects with respect to one another.

Pilot Study

Task characteristics influence which cognitive resources must be directed towards navigating a challenge and overcoming risks (Wickens, 2002); however, the relationship between difficulty, performance, and subjective perceptions of difficulty may change depending on the cognitive resources required by the task. Before exploring the nature of this relationship across tasks with differing characteristics, I sought to first ensure these tasks presented similar levels of challenge. Few – if any – other research studies conduct preliminary analyses to ensure that manipulating a particular task dimension does, indeed, make that task more difficult, and many assume experimenters' perceptions of difficulty will align with participants' task performance. Because intuitions about the characteristics of effortful tasks are not always accurate (Ganskopp, Cruz, & Johnson, 2000), and that multiple factors (e.g., exertion, sleepiness, discomfort) contribute to individual perceptions of effort (Åhsberg, 2000), I decided to conduct a pilot study to explore the effects of task manipulations on participant performance to ensure that subsequent studies (Experiments 1 & 2) involved an accurate and systematic variation of task difficulty across condition.

Method

Participants. A total of 31 participants (15 female) from the graduate student body of the Department of Psychological Sciences at Kansas State University completed 15-min or 40-min segments of the experimental task. These students volunteered and did not receive compensation for their participation in the study.

Procedure. To calibrate the difficulty of the videogame task, I conducted a series of pilot experiments that took place across four groups of participants. In these studies, participants completed short, 15-min segments in each condition of the videogame task (group 1) or completed a full-length, 40-min session in a single condition (groups 2-4). Other than the minor change in session duration for group one, the videogame task was identical to that discussed previously. Performance during these experiments informed the range of difficulty used in Experiment 1.

Experimental Manipulation. Task difficulty was manipulated in each condition by changing a single characteristic of the enemy characters' behavior at the start of each level (for a list of these characteristics, see Table 1. All other characteristics of the enemy characters' behavior remained the same throughout the session.

After collecting each group of calibration data, I standardized the difficulty of each condition (parameter value/parameter range) and calculated participants' rate of damage received (damage of a single attack/time since last attack), an objective measure of task performance. I then created a graph relating task difficulty and task performance and used this graph to determine changes in the random algorithm that manipulated task difficulty at the start of each level. A full disclosure of these changes can be found in Table 2.

Table 2

The videogame task was calibrated across four groups. The difficulty of each condition was determined by range and constant values of each characteristic of enemy characters' behavior.

Calibration group	Condition	Range	Constant
1 (*n* = 4)	population	5 - 20	10
	damage	5 - 20	10
	strength	90 - 250	100
	line-of-sight	5 - 26	100
	speed	1 - 7	3
2 (*n* = 7)	population	1 - 15	10
	damage	5 - 75	35
	strength	70 - 300	185
	line-of-sight	3 - 30	100
	speed	1 - 12	6
3 (*n* = 6)	population	5 - 15	10
	damage	5 - 90	20
	strength	70 - 300	115
	line-of-sight	6 - 30	100
	speed	1 - 10	5
4 (*n* = 14)	population	5 - 17	10
	damage	10 - 90	20
	strength	70 - 270	115
	line-of-sight	6 - 30	100
	speed	1 - 10	5

Discussion

This pilot study demonstrates that objective performance measures (e.g., damage rate) can and should be used to calibrate difficulty across conditions or tasks, especially when challenges involve different dimensions of difficulty. Task characteristics that underlie dimensions of difficulty may depend on relative scaling or non-comparable scales. For example, enemy characters' speed is contingent upon the size of the game space and how quickly the participant's avatar moves. Standardizing difficulty across task characteristics (here, conditions) ensures that performance differences are due to similar changes in task difficulty rather than a byproduct of scaling complexity.

Experiment 1

Appropriate resource allocation strategies are sensitive to the impact that changing internal and external factors have on task performance (i.e., gains and losses) and risk. Specifically, the relative success of a decision-making strategy depends on a decision-maker's metacognitive ability: how closely subjective perceptions align with objective measures of performance and by extension, risk. While objective measures of performance improve over time (for a review see Ericsson, Krampe, & Tesch-Romer, 1993), it is not clear whether decision-makers recognize these performance gains. Subjective perceptions of effort are impacted by task characteristics unrelated to performance such as boredom (Gilbertova & Glivicky, 1967), attention (Razon, Hutchinson, & Tenenbaum, 2013), and time-on-task (D'Huyvetter as cited in Ackerman, 2011). As time-on-task increases, participants become less flexible in their decision-making strategies (Lorist, Boksem, & Ridderinkhof, 2005; van der Linder, Frese, & Meijman, 2003). Together, these findings seem to suggest a dichotomy between subjective perceptions and objective measures of difficulty. However, existing research methods involve a post-hoc evaluation of task difficulty and make it difficult to draw conclusions about participants' perceptions of difficulty during task completion.

This experiment sought to establish the relationship between subjective perceptions and objective measures of difficulty, and to determine the effects of time-on-task on this relationship. It was predicted that perceptions and objective measures of performance would be affected by game difficulty: as task difficulty increases, performance (defined as the avatar's damage rate) should suffer (H1) and the task should be perceived as more challenging (H2). However, there should be some degree of miscalibration between subjective perceptions of effort and objective measures of performance (H3), as advanced in the earlier core predictions. The existing literature advanced competing hypotheses regarding the effects of time-on-task on task performance, and it was predicted that time-on-task would affect perceptions of task difficulty (H3a) and perhaps also objective measures of task performance (H3b). Finally, it was anticipated that some dimensions of effort would be more difficult than others (H4).

Method

Participants. A total of 107 participants (65 female) from the General Psychology pool at Kansas State University completed the experimental task and received 1 hr of research credit compensation. One participant experienced a panic attack during the session, and this data has been excluded from analyses. A small subset of the participants (n = 10) did not complete all 40 min of the experiment due to computer issues; however, no participant experienced fewer than 31 min of gameplay. Thus, all remaining data were retained.

Procedure. Participants completed a full-length, 40 min session of the videogame task identical to that discussed previously.

Experimental manipulation. As before, task difficulty was manipulated in each condition by changing a single characteristic of the enemy characters' behavior at the start of each level (for a list of these characteristics and number of participants in each condition, see Table 1). All other characteristics of the enemy characters' behavior remained the same throughout the session. The values assigned to each characteristic at the start of a level can be found in Table 3.

Table 3

The difficulty of each condition was determined by range and constant values of each characteristic of enemy characters' behavior.

Condition	Range	Constant	Unit of measurement
population (n = 16)	1 - 25	10	seconds
damage (n = 23)	10 - 90	20	hit points
strength (n = 22)	20 - 400	115	hit points
line-of-sight (n = 22)	6 - 30	66	Unity units
speed (n = 24)	0.2 - 15.0	5.0	Unity units

Note. Unity units are an arbitrary measure that can be used to scale game objects with respect to one another.

Results

Variables. Participants' rate of damage from enemy characters served as an objective measure of performance during the experimental task. A Box-Cox analysis ($\lambda = 0$) indicated a log transform was necessary to restore the distribution of this variable to normality (Box & Cox, 1964). However, when log-transformed rate of damage was used as a criterion variable in subsequent analyses, it produced a model with non-normal residuals (see Figure 4). Because the residuals appeared to be bimodal, a k-means cluster analysis was used to dichotomize rate of damage received into high and low damage groups without a significant loss of power (Irwin & McClelland, 2003). The reported analyses use cluster group as a dependent measure of task performance (rate of damage) and take appropriate measures to account for the bimodality of the data.

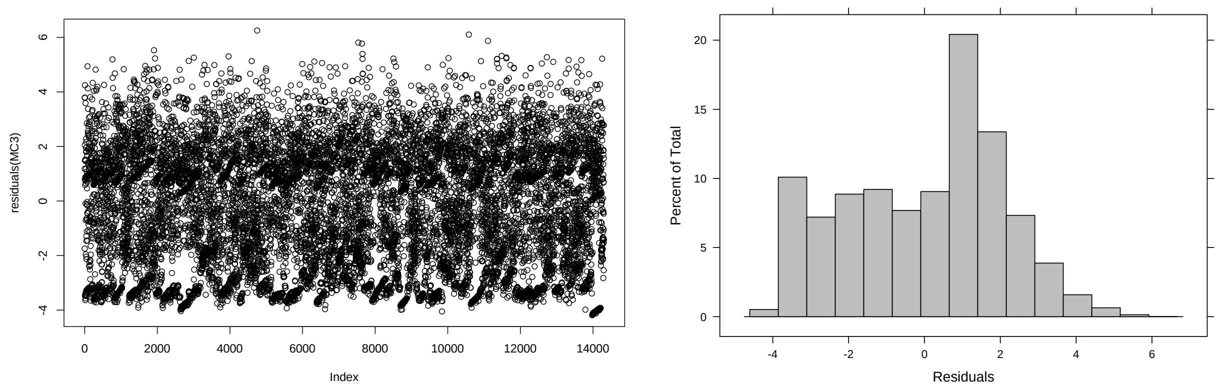

Figure 4. Diagnostic plots of the residuals from a linear mixed effects model that predicts log transformed rate of damage received using main effects of game difficulty, condition, time, videogame experience, and the Difficulty x Condition interaction.

The value assigned to the condition-specific difficulty parameter at the start of each level was used as an objective indicator of task difficulty. Because each difficulty parameter occupied a program-specific scale, task difficulty was standardized across conditions by applying the following equation:

$$\textit{standardized difficulty} = \frac{\textit{programmed difficulty} - \textit{minimum programmed difficulty}}{\textit{minimum programmed difficulty} - \textit{maximum programmed difficulty}} \quad (2)$$

which produces a value in the range of 0 to 1. For example, a participant in the Strength condition that encounters opponents with 100 hit points would experience a standardized difficulty of (100-20)/(400-20), or 0.21. Participants' ratings of game difficulty ("easier/harder than before") were used as a subjective measure of task difficulty, with "harder" responses coded as 1. Participant experience was determined by summing responses to a nine-item videogame questionnaire that asked participants to indicate the frequency of their videogame play on a scale from 0 (*never*) to 4 (*daily*) for several different types of videogames (e.g., "Multiplayer Online:

e.g., World of Warcraft, Everquest"). All predictors were effect coded and centered before analysis, and time-on-task was scaled to occupy a range similar to other predictors.

Because the probability of making a "false discovery" increases with each statistical test, I conducted only those analyses outlined by my hypotheses. Pairwise comparisons are only conducted when necessary to address specific predictions, and appropriate corrections are made to reduce familywise error. Together, these steps effectively reduce researcher degrees of freedom and limit the probability of a Type I error (Gelman & Loken, 2013).

Performance is Affected by Game Difficulty, Time-on-task, and Videogame Experience. A multilevel logistic regression was used to test the predictions of H1, H3a, and H4. The fixed effect structure tested for the main effects of game difficulty, condition (damage, line-of-sight, strength, population, & speed), time-on-task, and videogame experience, as well as the Difficulty × Condition interaction. The random effect structure was selected using the Akaike Information Criteria (Akaike, 1973; see Table 4), a goodness-of-fit measure based on Log Liklihood, and included intercept, game difficulty, and time-on-task to account for participant differences in overall ability, sensitivity to differences in difficulty, and rates of learning.

Table 4

AIC for proposed random effect structures.

Random Effect Structure	AIC
intercept	19037.3
intercept + time	19009.1
intercept + standardized difficulty	18973.0
intercept + standardized difficulty + time	18953.9

Note: A difference in AIC greater than 10 is very strong evidence in favor of the model with the lower AIC (Raferty, 1995).

Game difficulty significantly affected participants' performance in the game. Participants were more likely to experience high rates of damage during difficult levels than during easy levels (B = 0.12, z = 4.62, p < .001), confirming that difficulty can be manipulated across a continuous dimension and that this manipulation is reflected in performance (H1; see Figure 5). Figure 5 also illustrates that performance improved over time as players became less likely to experience high rates of damage in the game (B = -0.16, z = -7.77, p < .001). This suggests that participants gained familiarity with the videogame and developed strategies to become more successful players (H3a). Additionally, videogame experience affected participants' performance, such that players with gaming experience were less likely to experience high rates of damage overall than their less experienced peers (B = -0.04, z = -5.39, p < .001), as shown in Figure 6.

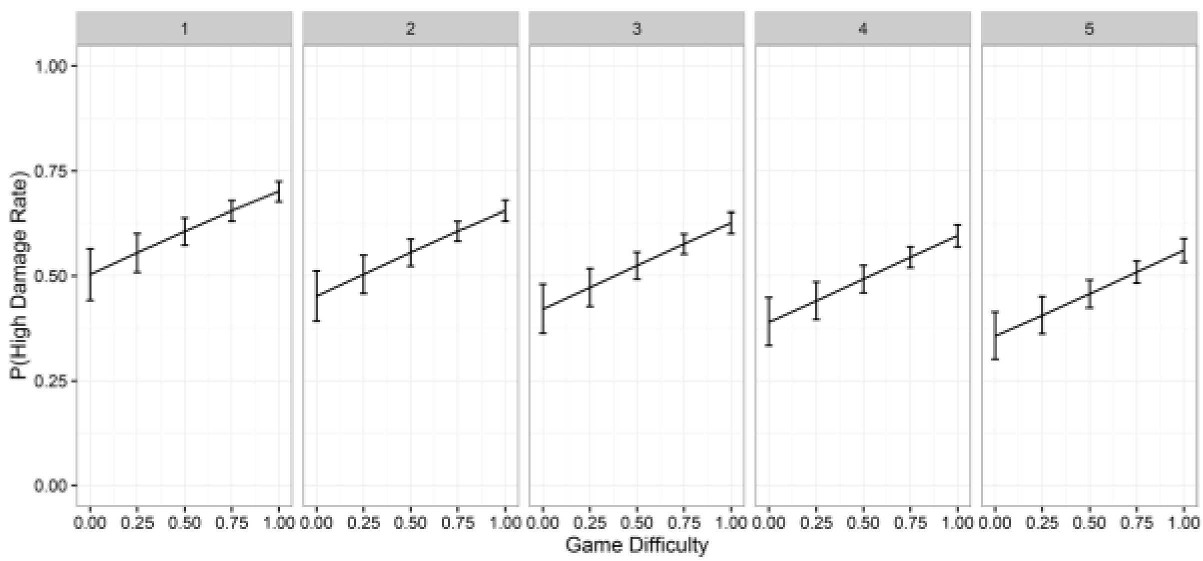

Figure 5. Game (standardized) difficulty and time-on-task significantly affect predicted performance. Time-on-task is depicted across panels, with earlier cross-sections of data appearing on the left and those from alter in the game appearing on the right. Error bars represent 1 standard deviation above and below the predicted values.

Some Conditions were More Difficult than Others. A Wald χ^2 confirmed that the probability of receiving a high (vs. low) rate of damage varied across conditions ($\chi^2(4)$ = 26.57, p < .001). A post-hoc Tukey's HSD revealed significant differences between participants' P(high damage rate) in the Speed and Damage conditions. Specifically, *P*(high damage rate) was greater in the Speed condition than in any other condition, and significantly lower in the Damage condition than in all other conditions except Line-of-sight (see Table 5 and Figure 7).

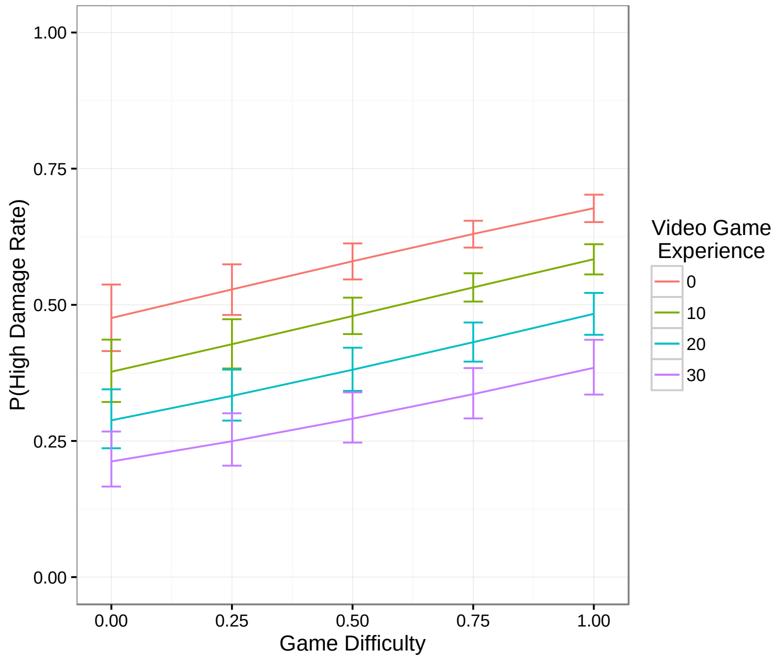

Figure 6. Videogame experience also affected participants' performance in the game. more experienced players performed better (i.e., had lower rate of damage received) than their less experienced peers across all levels of standardized game difficulty. Error bars represent 1 standard deviation above and below the predicted values.

Table 5

Tukey's post-hoc pairwise comparisons illustrate the difference in P(high damage rate) between conditions. Estimates represent the difference in average damage rate across conditions.

comparison	estimate	*z*	*p*
Damage – Line-of-sight	-0.78	-2.66	0.06
Damage – Population	-1.54	-4.53	< .001
Damage – Speed	-2.47	-8.55	< .001
Damage – Strength	-0.99	-3.59	0.003
Line-of-Sight – Population	-0.76	-2.21	0.174
Line-of-Sight – Speed	-1.69	-5.79	< .001
Line-of-Sight – Strength	-0.22	-0.78	0.937
Population – Speed	-0.93	-2.77	0.044
Population – Strength	0.54	1.66	0.459
Speed – Strength	1.48	5.40	< .001

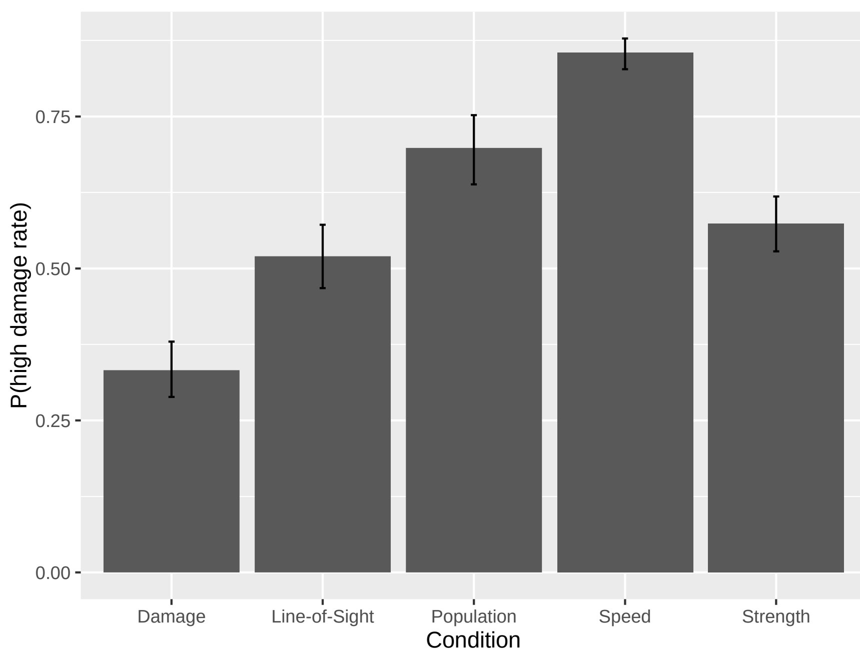

Figure 7. Model-estimated damage rate across conditions. Error bars represent 1 standard deviation above and below predicted values.

Perceptions of Game Difficulty were Affected by Performance Measures and Time-on-task. A second multilevel logistic regression was used to test the predictions of H2 and H3b. The fixed effect structure tested for the main effects of participants' rate of damage received since last question, condition, time-on-task, and videogame experience. The random effect structure included intercept to account for participant differences in overall ability. There was insufficient evidence, as assessed by AIC and convergence errors to include the slope random effects in the model (Bates, Kliegl, Vasishth, & Baayen, 2015).

Performance significantly affected participants' perceptions of game difficulty (see Figure 8). Participants were more likely to say the game was "harder than before" when they recently took lots of damage from enemy characters (B = 0.34, z = 3.74, p < .001). Thus, participants seemed capable of tracking their performance in the game over time (H2). Time-on-

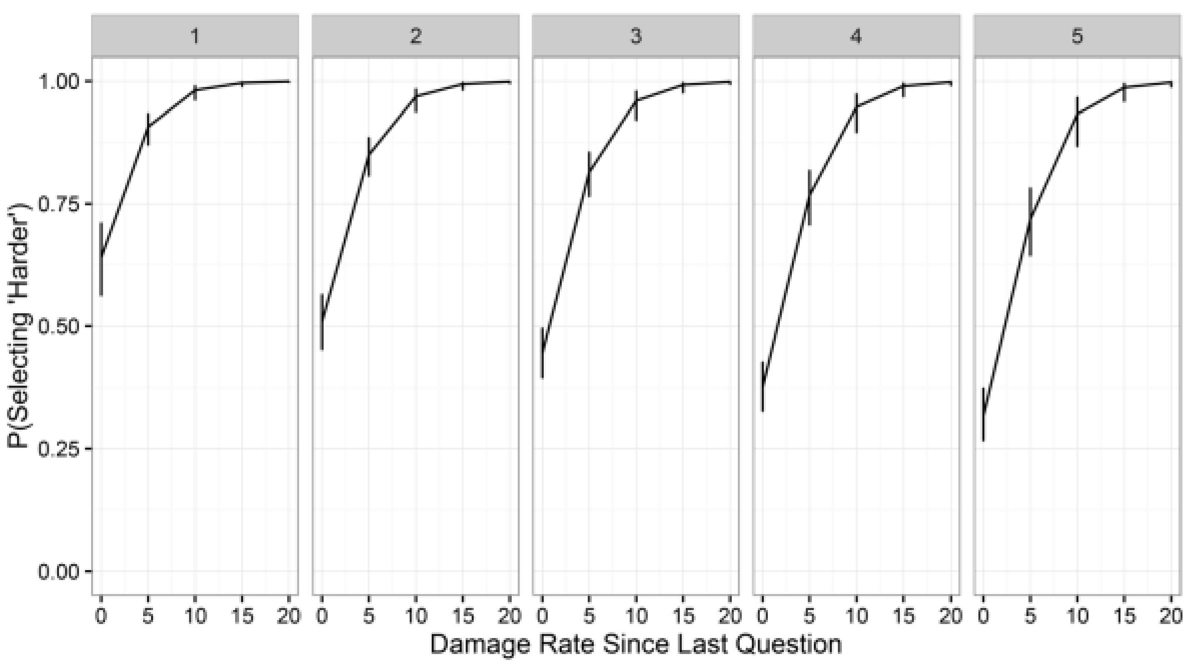

Figure 8. Rate of damage from enemy characters and time-on-task significantly affect participants' perceptions of game difficulty. Time-on-task is depicted across panels, with earlier cross-sections of data appearing on the left and those from later in the game appearing on the right. Each panel represents one-fifth of the total time-on-task. Error bars represent 1 standard deviation above and below the predicted value.

task also affected perceptions of game difficulty such that players became more likely to say the game was "easier than before" as time passed (B = -0.33, z = -3.38, p < .001; see Figure 8). These ratings aligned closely with the practice effect observed in previous analyses, and confirmed hypothesis H3a.

Other Factors Influence Perceptions of Difficulty. An exploratory analysis was conducted to determine whether factors beyond participants' rate of damage received since last question affected perceptions of difficulty. Adding standardized difficulty, an objective measure of level difficulty, to the second multilevel logistic regression significantly improved model predictions ($\Delta AIC = 39.963$). This suggests that several performance-related factors influence participants' perceptions of difficulty during task completion.

Other Factors Influence Perceptions of Difficulty. An exploratory analysis was conducted to determine whether factors beyond participants' rate of damage received since last question affected perceptions of difficulty. Adding standardized difficulty, an objective measure of level difficulty, to the second multilevel logistic regression significantly improved model predictions ($\Delta AIC = 39.963$). This suggests that several performance-related factors influence participants' perceptions of difficulty during task completion.

Discussion

The results of this study confirmed my hypothesis (H1) that specific manipulations of task difficulty affect participants' rate of damage, and further validated the use of the videogame task to study difficulty and risk-related decisions over time. Player performance was also affected by individual differences in videogame experience, a finding that provided additional face validity to the paradigm. More experienced participants had better overall performance, and all participants improved over time. Participants could track these changes in performance, and correctly identified when the game became “easier” or “harder,” confirming hypotheses H2. Contrary to hypotheses H3a and H3b, participants’ perceptions of game difficulty were appropriately calibrated and did not reflect increasing levels of fatigue. Finally, some versions of the videogame were more difficult than others, supporting hypothesis H4.

These findings suggest that decision-makers effectively track characteristics that contribute to task difficulty and may be able to use this information to adjust decision-making strategies over time. Encouragingly, exploratory analyses seemed to suggest that decision-makers’ ratings of game difficulty are based on factors beyond those related to performance. Although my analyses only involved damage rate, it is possible that participants track difficulty using gameplay characteristics (e.g., number of enemies on-screen) that could also be used to anticipate changes in performance and then adjust gameplay strategy before negative consequences occur.

Experiment 2

The purpose of Experiment 2 was to determine whether participants’ performance and perceptions of difficulty affected risk mitigation strategy during dynamic tasks. Previous literature on decision-making in risky situations suggested three competing hypotheses (see Table 6). This experiment seeks to better understand how individuals utilize risk mitigation strategies when making decisions within a videogame environment that is responsive to the previous choices and behavior of the player.

Table 6.

Theoretically motivated competing hypotheses regarding decision-makers' risk mitigation strategies in difficult situations involving risks.

Theoretical interpretation	Hypothesized risk mitigation strategy
Rational Choice Theory (H1a)	Decision-makers engage in optimal risk mitigation strategies. The losses associated with difficult tasks (i.e., levels) are mitigated through the use of pre-event RDOs; post-event RDOs are utilized during less difficult tasks.
Prospect Theory (H1b)	Decision-makers use riskier strategies to mitigate loss. As tasks become difficult and the probability of failure increases, decision-makers turn to post-event RDOs to mitigate their risks; pre-event RDOs are preferred during less difficult tasks
Ecological Rationality (H1c)	Decision-makers' risk mitigation strategies shift over time as they gain experience with the environment and improve their metacognitive abilities.

Method

Participants. A total of 79 participants (43 female) from the General Psychology pool at Kansas State University completed the experimental task and received 1 hr of research credit compensation. One participant experienced a computer malfunction and needed to restart the videogame. The remaining data from this participant is included in the analysis.

Procedure. Experiment 2 involved only three of the five conditions from experiment one. Specifically, participants were assigned to one of three conditions: Speed (n = 23), Strength (n = 30), or Population (n = 26). The videogame task was similar to that used in experiment one, but included an experimental manipulation that affected gameplay and the look-and-feel of the game.

Experimental manipulation. At the beginning of each level, and every 5 min, a pop-up window invited participants to "select a tool" that would assist them in playing the game (see Figure 9). This pop-up window was a slightly different color than the pop-up window that allowed participants to indicate their subjective perceptions of game difficulty, but behaved similarly in all other respects. Once participants selected an option, the avatar was granted a set of five tools that could be used at any point before the tool selection pop-up window appeared again.

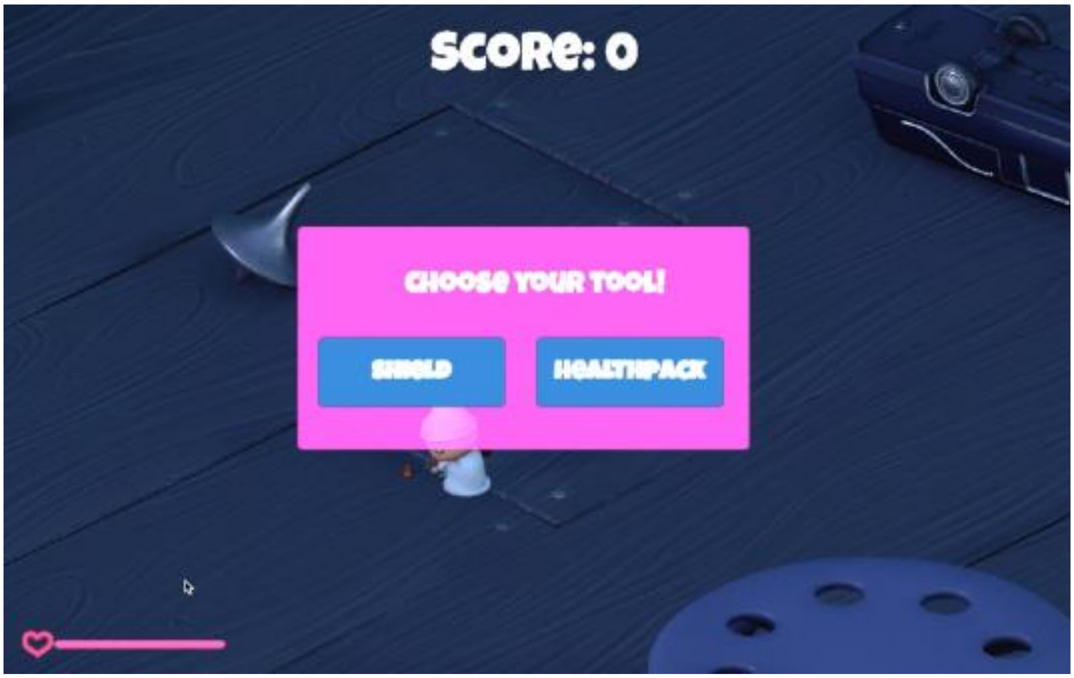

Figure 9. Participants selected a tool to use as part of their gameplay strategy.

Figure 10. Icons near the health bar allowed participants to track the number of tools available for use.

Participants could choose between receiving five shields, each of which could be used to block 20 hit points of damage from enemy characters, or five health packs, each of which could be used to immediately restore 20 hit points of the avatar's health. These tools represented different RDOs. The shield, which needed to be deployed before an enemy attack, represented a pre-event RDO. The health pack, which could only be deployed successfully after an enemy attack, represented a post-event RDO. Thus, each tool had distinct strengths and weaknesses that participants could use to improve their gameplay strategy and task performance.

Once participants selected a tool option, five icons appeared to the right of the avatar's health bar (see Figure 10). Participants could deploy a tool by pressing the F key. Each time a tool was used, participants were notified by visual (see Figure 11) and auditory cues. Specifically, participants heard a 250-ms sound and saw a bubble appear around the avatar for the tool's duration. Both the sound and the bubble were specific to the tool and could be used to differentiate tool choice. The health pack bubble remained engaged for approximately 1 s, while the shield bubble remained engaged for approximately 5 s or until the shield was struck by an enemy character. Health packs were consumed following deployment while shields were only consumed following an enemy attack. When a tool was consumed, one of the five icons disappeared from the screen.

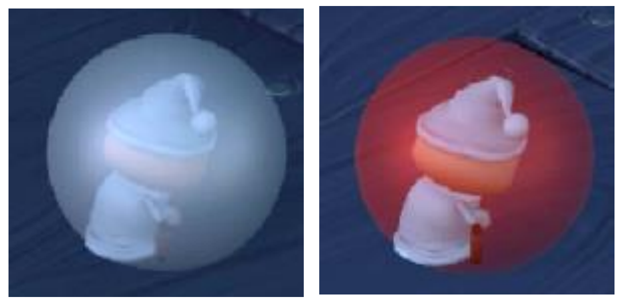

Figure 11. A visual indicator notified participants when they used a shield (left) or a health pack (right).

Because participants were asked to select a tool every 5 min during the task, it was possible for participants to either retain or switch gameplay strategies in the middle of a level. However, participants "lost" all unused tools after making a tool selection. That is, participants could not carry more than five tools at a time, and tools were always of the same type.

Results

Variables. Participants' rate of damage received since the last tool selection served as an objective measure of game performance during the experimental task. Rate of damage received was calculated in three different ways by including only experienced damage ("hits-only rate of damage received"), both experienced damage and damage to the shield ("shielded rate of damage received"), or by adjusting the rate of damage received by adding 20 hit points when a health pack was used ("health pack-adjusted rate of damage received"), excluding instances in which a health pack was deployed after the avatar died. All three predictors were tested to ensure the correct interpretation of the data. Subsequent model comparisons using AIC suggested that the predictors had a similar influence on each measure; for simplicity, only the analysis of experienced damage is included here. For a full disclosure of all model comparisons, see Appendix B.

Risk mitigation strategy was determined by participants' tool choices during the game. Health packs, which represented pre-event RDOs, were coded as 1. Participants' ratings of game difficulty ("easier/harder than before") were used as a subjective measure of task difficulty, with "harder" responses coded as 1. Participants' videogame experience was determined by summing responses to a nine-item videogame questionnaire that asked participants to indicate the frequency of their videogame play on a scale from 0 (*never*) to 4 (*daily*) for several different types of videogames (e.g., "Multiplayer Online: e.g., World of Warcraft, Everquest"). All predictors were effect coded and centered before analysis, and time-on-task was scaled to occupy a range similar to other predictors. As before, an effort was made to conduct only those tests outlined in the hypotheses to reduce the probability of a Type I error (Gelman & Loken, 2013).

Resource Allocation and Risk Mitigation Strategies are Driven by Difficulty but Altered by Time-on-task. A multilevel logistic regression was used to test the predictions of H1 and H3. The fixed effect structure tested for the main effects of rate of damage received, condition (population, speed, & strength), time-on-task, and videogame experience, as well as the Rate of damage received × Time-on-task interaction. The random effect structure was selected using the AIC (see Table 7) and included intercept, rate of damage received, and time-on-task to account for participant differences in overall ability, experiences of difficulty, and rates of learning.

Table 7.

AICs for proposed random effect structures.

Random Effect Structure	AIC
intercept	647.9
intercept + rate of damage received	642.2
intercept + rate of damage received + time	632.5

Note: A difference in AIC greater than 10 is very strong evidence in favor of the model with the lower AIC (Raftery, 1995).

Game performance systematically affected participants' risk mitigation strategy during the game such that participants were more likely to select a shield when they received a higher rate of damage from enemy characters (B = -0.51, z = -2.26, p = 0.024), suggesting that resource allocation and risk mitigation strategies are driven by individual sensitivities to environmental changes (H1c; see Figure 12). Risk mitigation strategy was significantly affected by time-on-task (B = 0.36, z = 2.15, p = 0.03), but not in the direction predicted by hypothesis H3. Participants became less likely to engage in the predominant risk mitigation strategy as they progressed through the game.

Subjective Perceptions of Game Difficulty Do Not Improve Model Predictions Regarding Risk Mitigation Strategy. A second multilevel logistic regression was used to test the predictions of H2. The fixed effect structure tested for the main effects of rate of damage received, condition (Population, Speed, & Strength), time-on-task, videogame experience, and perceptions of game difficulty, as well as the Rate of damage received × Time-on-task interaction. The random effect structure included intercept, game difficulty, time-on-task, and perceptions of game difficulty to account for individual differences in overall ability, experiences of difficulty, rates of learning, and metacognitive abilities across participants.

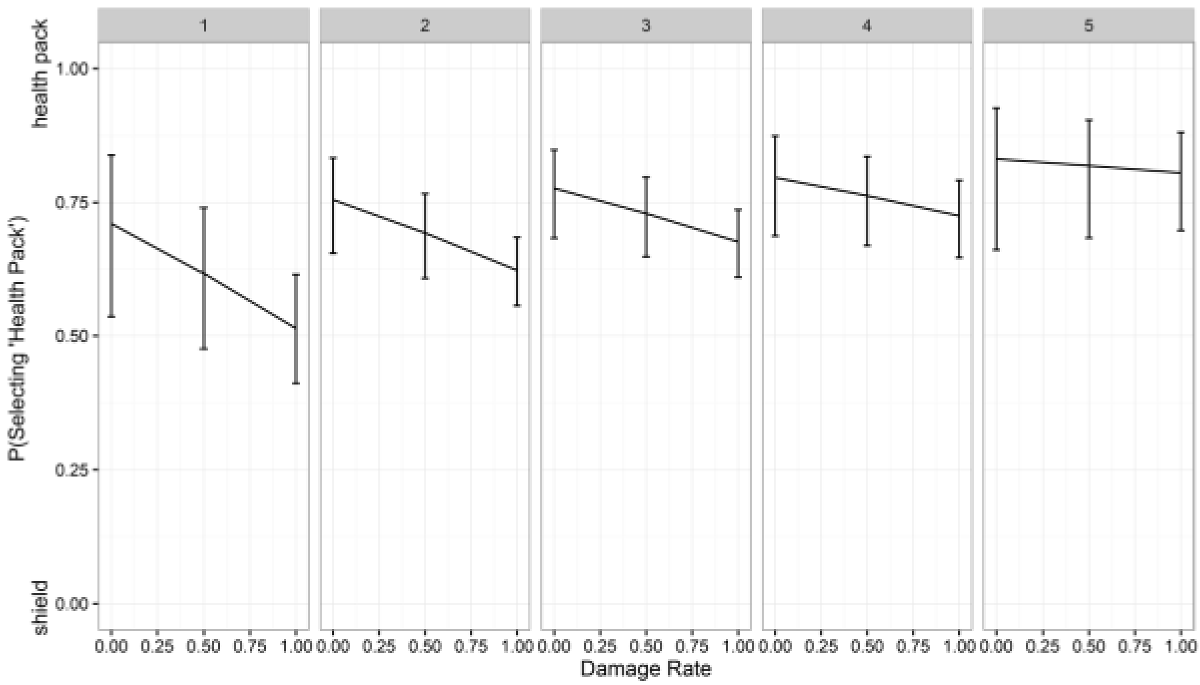

Figure 12. Rate of damage from enemy characters affects the probability that participants will select a health pack during the game. As players move through their game, preference shifts to favor health packs. Each panel represents an equal amount of time-on-task. Error bars represent 1 standard deviation above and below predicted values.

The AIC (Akaike, 1973) can be used to compute the likelihood of one model producing the data as compared to another (Wagenmakers & Farrell, 2004):

$$L(M_i|data) \propto exp\left\{-\frac{1}{2}\Delta_i(AIC)\right\} \tag{2}$$

This measure is advantageous because it provides a concrete measure of comparison between two models without requiring that those models be nested. Adding perceptions of game difficulty to the multilevel model used to test hypothesis H1 did not improve the model's predictions. In fact, a model including perceptions of game difficulty was 2.3 times less likely to have produced the data, providing weak support for a model that excluded this predictor.

Discussion

Several theoretical models predict behavior in risky environments. Rational Choice Theory (Arrow, 2004) predicts that decision makers will behave optimally and develop idiosyncratic strategies that benefit their own abilities and decision making style. Prospect Theory (Kahneman & Tversky, 1979) and Ecological Rationality (Gigerenzer, Todd, & ABC Research Group, 1999) predict that decision makers will engage in preferred risk mitigation strategies by selecting either riskier, post-event RDOs or preventative pre-event RDOs respectively. I investigated whether risk mitigation strategies changed in response to environmental changes, specifically whether changes in task difficulty would affect performance and risk mitigation decisions.

Decision-makers' risk mitigation strategies most closely aligned with those predicted by Ecological Rationality: decision-makers became more likely to select preventative, pre-event RDOs (i.e., shields) as task difficulty increased and performance worsened (H1c). Perceptions of difficulty seemed to explain this behavior as well as objective measures of performance. This finding did not support hypothesis (H2), but instead suggests that decision-makers' perceptions of difficulty may align with objective measures of performance. This congruency emerged in Experiment 1, but remained limited: decision-makers improved in their ability to discriminate "easier" levels but were not completely accurate. Thus, decision-makers track losses and risks in dynamic situations but maintain some degree of error while doing so.

As time-on-task increased, participants became more likely to select riskier post-event RDOs (i.e., health packs). While it is possible that this shift towards riskier mitigation strategies is due to a fatigue effect, results from Experiment 1 supported a practice effect rather than a fatigue effect. Therefore, it seems more likely that participants' shift in risk mitigation strategy reflects a greater efficiency with or preference for that tool. This suggestion aligns with Ecological Rationality, which predicts that behavior will be imperfectly driven by sensitivities to environmental changes including those that are related to the decision maker. It is also possible that participants' shift in risk mitigation strategy was motivated by improvements in metacognitive ability: some research suggests that post-event RDOs become more frequently selected as risk detection ability improves (Kirchler, Hoelzel, & Huber, 2010).

General Discussion

The present findings indicated that videogame tasks can be used to study effort and risk mitigation, and explored the factors that contribute to decision-makers' risk mitigation strategies and metacognition. Decision-makers appear to appropriately gauge task difficulty and performance in real time, updating risk mitigation strategies as situations change. Specifically, decision-makers tended to prefer preventative risk mitigation strategies when task difficulty increased and negatively affected performance. Over time, decision-makers' performance improved and they became less likely to take preventative measures to mitigate risks. These findings align with the theory of Ecological Rationality, which proposes that decision-makers can adequately track losses and manage risks without engaging in optimal practices, affirm the existence of the description-experience gap (Camilleri & Newell, 2013), and support previous work exploring the use of RDOs in description-based decision making (e.g., Huber & Huber, 2008).

It is important to consider how these findings relate broadly to cognitive resource management, specifically with respect to resource reallocation. Resource reallocation is an iterative process that allows organisms to adapt to a changing environment. As decision-makers receive feedback regarding task characteristics and performance, they may shift cognitive resources from one pursuit (e.g., gameplay) to another (e.g., deployment of a health pack). The appropriate allocation of resources is learned through experience with the task, and more "expert" decision-makers can predict which resource allocation strategies will be most successful and when to employ those strategies. Within this experiment, metacognitive abilities seemed to align with game performance and improved over time. How risk mitigation strategies are affected by metacognitive skills is uncertain, an issue that will be visited in discussing the limitations of this experiment.

Although decision-makers' shift in RDO preference seems to diverge from optimality, it is important to recall that optimal performance in a dynamic environment requires the occasional exploration of alternative strategies (Kurzban, Duckworth, Kable, & Myers, 2013). Shifts in preference from pre-event to post-event RDOs may reflect strategic exploration that occurs as decision-makers develop greater familiarity with their environments. Indeed, tool preference and performance both shifted across gameplay: post-event RDOs were most frequently selected in later levels, when task performance was strongest. However, it is also possible that the shift to post-event RDOs represents the exploitation of a successful strategy: perhaps decision makers discovered that their abilities and gameplay strategy led post-event RDOs (i.e., the health pack) to become more efficient than pre-event RDOs (i.e., the shield). Further analyses will be necessary to differentiate these competing hypotheses.

Regardless of decision-makers' motivation for shifting risk mitigation strategies, the results underscore the importance of ecological validity in experimental design. Unlike previous research designs that explored decision-making in a static, descriptive environment (Huber & Huber, 2003), my participants tracked performance and employed risk mitigation strategies during engaging and dynamic tasks. While participants in static tasks exhibit a preference for risky choice (Barberis, 2013) and do not actively seek outcome probability information (Huber & Huber, 2003), my participants were more likely to employ preventative strategies and appeared to be sensitive to probability information relating to overall task performance (i.e., difficulty and probability of failure). This evidence suggests that decision-makers behave differently in dynamic environments and underscores the need to bring methodologies into alignment with the type of decision-making in question. While description-based paradigms may be appropriate to study decisions that are made with written evidence (e.g., policy or future planning decisions), dynamic paradigms should be used to study decisions made in changing environments (e.g., job-related tasks or investment decisions).

This research also appears to contradict existing literature on time-on-task effects in that performance and perceptions of difficulty improved over time (Lorist & Faber, 2011). However, participants became more rigid in their decision-making strategies over time. While it is possible that the increased *P*(health pack selection) is due to exploitation of a successful strategy, it may also represent a perseveration in decision-making strategy (per Lorist, Boksem, & Ridderinkhof, 2005; van der Linden, Frese, & Meijman, 2003). Both may be affected by changes in metacognitive ability, and future research should explore these competing hypotheses in greater depth.

Limitations

Conducting research in a dynamic environment introduces additional variables that may affect performance in systematic ways. In this study, I assume that pre-event RDOs represent a low-risk, high-resource tool for mitigating risk. That is, pre-event RDOs can be implemented before losses occur but require additional cognitive resources to deploy. Conversely, I assume that post-event RDOs represent a high-risk, low-resource tool for mitigating risk because they can only be implemented once losses occur but require fewer cognitive resources to deploy. While this assumption is intuitively appealing, participants' skills may affect the risk and resource trade-offs associated with each option. Pre-event RDOs may be riskier for a player that struggles to coordinate shield deployment because using this tool results in certain loss. Future analyses should consider the relative risks of available RDOs to account for differences in abilities between participants.

It is also possible that risk mitigation strategies are affected by task goals. In this task, participants mitigate risks while eliminating enemy characters from the screen. This behavior is different from that of an air traffic controller or physician who reduces risks while saving lives. Additional research must be conducted to explore the generalizability of these findings to other tasks and environments.

Conclusion

In summary, decision-makers shift risk mitigation strategies when provided an opportunity to acquire information about their environments during dynamic tasks. The ability to shift risk mitigation strategies appears to be rooted in performance and metacognitive abilities, which continue to improve over time as decision-makers gain experience with task and environmental characteristics. This pattern of shifting risk mitigation strategies most closely aligns with the predictions of Ecological Rationality; however, the present analyses were insufficient to determine whether these strategic adjustments were due to a systematic bias that arose as time-on-task increased, an exploratory strategy that arose as decision-makers gained familiarity with their environment, or in response to the probability of success associated with a specific risk mitigation strategy. Future research should explore the factors that motivate these shifts with the intention of identifying optimal behavior and predicting when shifts should occur.

References

Ackerman, P. L. (2011). *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications*. Washington, DC, US: American Psychological Association.

Åhsberg, E. (2000). Dimensions of fatigue in different working populations. *Scandinavian Journal of Psychology, 41,* 231-241.

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika, 60,* 255-265.

Altmann, E. M., & Trafton, J. G. (2006). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review, 14,* 1079-1084.

Arrow, K. J. (2004). Is bounded rationality unboundedly rational? Some ruminations. In M. Auger & J. G. March (Eds.), *Models of a Man: Essays in Memory of Herbert A. Simon* (pp. 47-57). Cambridge, MA: MIT Press.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *The Journal of Economic Perspectives*, *27*, 173-195.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16,* 215-233.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. Unpublished manuscript. arXiv:1506.04967v1.

Björkman, M. (1994). Internal cue theory: Calibration and resolution of confidence in general knowledge. *Organizational Behavior and Human Decision Processes, 58,* 386-405.

Böckenholt, U. (2004). Comparative judgments as an alternative to ratings: Identifying the scale origin. *Psychological Methods, 9,* 453-465. doi:10.1037/1082-989X.9.4.45

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational science. *Organizational Behavior and Human Decision Processes, 101,* 127-151.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B, 26*, 211-252.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*, 624-634. doi:10.1016/j.jesp.2009.02.01

Brown, J. R., Kling, J. R., Mullainathan, S., & Wrobel, M. V. (2008). Why don't people insure late-life consumption? A framing explanation of the under-annuitization puzzle. *American Economic Review, 98*, 304-309. doi:10.1257/aer.98.2.30

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA: University of California Press.

Burns, T., & Roszkowska, E. (2016). Rational Choice Theory: Toward a psychological, social, and material contextualization of human choice behavior. *Theoretical Economics Letters, 6*, 195-207.

Camilleri, A. R., & Newell, B. R. (2013). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In Chandrasekhar Pammi, V. S., & Srinivasan, N. (Eds.), *Progress in Brain Research: Decision Making Neural and Behavioral Approaches* (pp. 55-71). Amsterdam, The Netherlands: Elsevier.

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review, 18*, 139-165.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58,* 1-73.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19,* 51-57.

Detweiler, J. B., Bedell, B. T., Salovey, P., Pronin, E., & Rothman, A. J. (1999). Message framing and sunscreen use: Gain-framed messages motivate beach-goers. *Health Psychology, 18*, 189.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., … Lebiere, C. (2009). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23,* 15-47.

Federal Aviation Administration. (2009). *Flightcrew member duty and rest requirements.* (Docket No. FAA-2009-1093; Notice No. 10-11). Washington, DC: Federal Aviation Administration.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34,* 906-911.

Freeman, T. P., Morgan, C. J., Brandner, B., Almahdi, B., & Curran, H. V. (2013). Dopaminergic involvement in effort-based but not impulsive reward processing in smokers. *Drug and Alcohol dependence, 130*, 109-114.

Ganskopp, D., Cruz, R., & Johnson, D. E. (2000). Least-effort pathways?: A GIS analysis of livestock trails in rugged terrain. *Applied Animal Behaviour Science, 68,* 179-190.

Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". In Stroebe, W., & Hewstone, M. (Eds.), *European Review of Social Psychology,* (Vol. 2, pp. 83-115) Chichester, UK: Wiley.

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). *Simple Heuristics that Make us Smart.* Oxford, England: Oxford University Press.

Gilbertova, S., & Glivicky, V. (1967). Monotony at work. *Studia Psychologia*, *9*, 232-240.

Hartmann, M. N., Hager, O. M., Tobler, P. N., & Kaiser, S. (2013). Parabolic discounting of monetary rewards by physical effort. *Behavioural Processes, 100,* 192-196.

Harsanyi, J.C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge, UK: Cambridge University Press.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making, 23,* 48-68. doi:10.1002/bdm.66

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. Journal of Behavioral Decision Making, 21(5), 493-518. doi:10.1002/bdm.59

Huber, O. (2012). Risky decisions: Active risk management. *Current Directions in Psychological Science, 21*, 26-30.

Huber, O., & Huber, O. W. (2003). Detectability of the negative event: Effect on the acceptance of pre- or post-event risk-defusing actions. *Acta Psychological, 113,* 1-21.

Huber, O., & Huber, O. W. (2008). Gambles vs. Quasi-realistic scenarios: Expectations to find probability and risk-defusing information. *Acta Psychologica, 127*, 222-36. doi:10.1016/j.actpsy.2007.05.00

Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review, 39*, 25.

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research, 40,* 366-371.

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall, Inc. Retrieved from https://www.princeton.edu/~kahneman/docs/attention_and_effort/Attention_lo_quality.pd

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*, 263-291.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74,* 657-690.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107,* 392-402.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39,* 98-114.

Kirchler, E., Hoelzl, E., & Huber, O. (2010). Risk-defusing in decisions by probability of detection of harm and promotion and prevention focus. *The Journal of Socio-Economics, 39,* 595-600.

Knight, F. H. (1921). *Risk, Uncertainty and Profit.* New York, NY: Hart, Schaffner, & Marx.

Koehler, D. J., Brenner, L., & Griffin, D. (2002) The calibration of expert judgment: Heuristics and biases beyond the laboratory. In Gilovich, T., Griffin, D., & Kahneman, D. (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 686-715). Cambridge, UK: Cambridge University Press.

Kop, N., & Euwema, M. C. (2001). Occupational stress and the use of force by Dutch police officers. *Criminal Justice and Behavior, 28,* 631-652.

Krebs, J. (1977). Optimal foraging: Theory and experiment. *Nature*, *268,* 583-584.

Kurzban, R. (2016). The sense of effort. *Current Opinion in Psychology, 7*, 67-70. doi:10.1016/j.copsyc.2015.08.00

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral Brain Science, 26,* 661-679.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist, 57,* 705-717.

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research, 24*, 199-205. doi:10.1016/j.cogbrainres.2005.01.01

Lorist, M. M., & Faber, L. G. (2011). Consideration of the influence of mental fatigue on controlled and automatic cognitive processes and related neuromodulatory effects. In Ackerman, P. L. (Ed.) *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications* (pp. 105-126). Washington, DC, US: American Psychological Association.

Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-experience gap in risky choice limited to rare events? *PLoSone, 6,* 1-7.

Murphy, A. H., & Winkler, R. L. (1977). Reliability of subjective probability forecasts of precipitation and temperature. *Journal of the Royal Statistical Society: Series C, 26,* 41-47.

National Institutes of Health, National Institute of Dental and Craniofacial Research. (2016). *Seal Out Tooth Decay* (NIH Publication No. 16-489). Retrieved from https://www.nidcr.nih.gov/OralHealth/Topics/ToothDecay/SealOutToothDecay.htm

Neale, M. A., Huber, V. L., & Northcraft, G. B. (1987). The framing of negotiations: Contextual versus task frames. *Organizational Behavior and Human Decision Processes, 39,* 228-241.

Nishiyama, R. (2014). Response effort discounts the subjective value of rewards. *Behavioural Processes, 107,* 175-177.

Ostaszewski, P., Bąbel, P., & Swebodziński, B. (2013). Physical and cognitive effort discounting of hypothetical monetary rewards. *Japanese Psychological Research, 55*, 329-337. doi:10.1111/jpr.1201

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111-163.

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes, 106,* 168-179.

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23,* 1-14. doi:10.1002/bdm.68

Razon, S., Hutchinson, J., & Tenenbaum, G. (2012). Effort perception. In Tenenbaum, G., Eklund, R., & Kamata, A. (Eds.) *Measurement in Sport and Exercise Psychology (pp.* 265-275). Champaign, IL: Human Kinetics.

Ricci, J. A., Chee, E., Lorandeau, A. L., & Berger, J. (2007). Fatigue in the U.S. Workforce: Prevalence and implications for lost productive work time. *Journal of Occupational and Environmental Medicine, 49,* 1-10. doi:10.1097/01.jom.0000249782.60321.2

Rogers, A. E. (2008). The effects of fatigue and sleepiness on nurse performance and patient safety. In R. G. Hughes (Ed.), Patient safety and quality: An evidence-based handbook for nurses. Rockville, MD: Agency for Healthcare Research and Quality.

Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 201-211.

Sigurdsson, S. O., Taylor, M. A., & Wirth, O. (2013). Discounting the value of safety: Effects of perceived risk and effort. *Journal of Safety Research, 46,* 127-34. doi:10.1016/j.jsr.2013.04.00

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63,* 129-138.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28,* 1-39.

Smith, M. R., Marcora, S. M., & Coutts, A. J. (2015). Mental fatigue impairs intermittent running performance. *Medicine and Science in Sports and Exercise, 47,* 1682-1690.

Swaen, G. M. H., van Amelsvoort, L. G. P. M., Bültmann, U., & Kant, I. J. (2003). Fatigue as a risk factor for being injured in an occupational accident: Results from the maastricht cohort study. *Occupational and Environmental Medicine, 60,* 88-92. doi:10.1136/oem.60.suppl_1.i8

Sugiwaka, H., & Okouchi, H. (2004). Reformative self-control and discounting of reward value by delay or effort. *Japanese Psychological Research, 45,* 1-9.

Tversky, A., & Kahneman, D. (1985). The framing of decisions and the psychology of choice. In Environmental Impact assessment, technology assessment, and risk analysis (pp. 107-129). New York, NY: Springer.

Unity Game Engine (Version 5.4) [Computer Software]. San Francisco, CA: Unity.

U.S. Food & Drug Administration. (2015). *About Dental Amalgam Fillings*. Retrieved from http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DentalProducts/DentalAmalgam/ucm171094.htm

van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: Effects on perseveration and planning. *Acta Psychologica, 113,* 45-65.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* (2nd ed.). Princeton, NJ: Princeton University Press.

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11,* 192-196.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E., & Rushworth, M. F. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks: The Official Journal of the International Neural Network Society, 19,* 1302-1314.

Wang, X. T. (2004). Self-framing of risky choice. *Journal of Behavioral Decision Making, 17,* 1-16.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3,* 159-177.

Appendix A -

The Game Engagement Questionnaire

Please answer the following questions with respect to the game you just played.

1. I lost track of time.

1	2	3
Yes	*Sort of*	*No*

2. Things seemed to happen automatically.

1	2	3
Yes	*Sort of*	*No*

3. I felt different.

1	2	3
Yes	*Sort of*	*No*

4. I felt scared.

1	2	3
Yes	*Sort of*	*No*

5. The game felt real.

1	2	3
Yes	*Sort of*	*No*

6. If someone talked to me, I wouldn't hear them.

1	2	3
Yes	*Sort of*	*No*

7. I got wound up.

1	2	3
Yes	*Sort of*	*No*

8. Time seemed to kind of stand still or stop.

1	2	3
Yes	*Sort of*	*No*

9. I felt spaced out.

1	2	3
Yes	*Sort of*	*No*

10. I didn't answer when someone talked to me.

1	2	3
Yes	*Sort of*	*No*

11. I couldn't tell that I was getting tired.

1	2	3
Yes	*Sort of*	*No*

12. Playing seemed automatic.

1	2	3
Yes	*Sort of*	*No*

13. My thoughts went fast.

1	2	3
Yes	*Sort of*	*No*

14. I lost track of where I am.

1	2	3
Yes	*Sort of*	*No*

15. I played without thinking about how to play.

1	2	3
Yes	*Sort of*	*No*

16. Playing made me feel calm.

1	2	3
Yes	*Sort of*	*No*

17. I played longer than I meant to.

1	2	3
Yes	*Sort of*	*No*

18. I really got into the game.

1	2	3
Yes	*Sort of*	*No*

19. I felt like I just couldn't stop playing.

1	2	3
Yes	*Sort of*	*No*

Appendix B -

Table 8

Random effect structures considered when testing the effects of the within-subject variables rate of damage from enemy characters and time-on-task on risk mitigation strategies.

Random Effect Structures	AIC
intercept	647.9
intercept + hits-only rate of damage received	642.2
intercept + shielded rate of damage received	645.7
intercept + health pack-adjusted rate of damage received	643.1
intercept + hits-only rate of damage received + time-on-task	632.5
intercept + health pack-adjusted rate of damage received + time-on-task	631.7
intercept + Hits-only Rate of damage received × Time-on-task	636.3
intercept + Health pack-adjusted Rate of damage received × Time-on-task	634.4

Table 9

Random effect structures considered when testing the effects of the within-subject variables rate of damage from enemy characters and time-on-task on risk mitigation strategies.

Fixed Effect Structures	AIC
Hits-only Rate of damage received × Time-on-task + videogame experience + condition	629.6
Health pack-adjusted Rate of damage received × Time-on-task + videogame experience + condition	626.4
Hits-only Rate of damage received × Time-on-task + Perception × Time-on-task videogame experience + condition	629.9
Health pack-adjusted Rate of damage received × Time-on-task + Perception × Time-on-task videogame experience + condition	625.5