THE INFLUENCES OF INFORMATION ACQUISITION AND HEIGHTENED AROUSAL ON ADOLESCENT RISK TAKING

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By Gail M. Rosenbaum
May, 2017

Examining Committee Members:

Jason Chein, Advisory Chair, Department of Psychology
Laurence Steinberg, Examining Chair, Department of Psychology
Vinod Venkatraman, Fox School of Business, Department of Marketing and Supply Chain Management
Vinay Parikh, Department of Psychology
David Smith, Department of Psychology
Crystal Reeck, Fox School of Business, Department of Marketing and Supply Chain Management

ABSTRACT

Adolescents are known to take more risks than adults, which can be harmful to their health and well-being. Interventions aimed at reducing risk taking typically provide descriptions of the negative outcomes that may result from a risky choice, and have shown little evidence of actually preventing risk taking. This lack of efficacy may be due in part to differences between how adolescents process information about risk when it is described (e.g., in a classroom intervention) versus when it is experienced (e.g., when a teenager experiences the outcome of a risky choice). In the present work, I first summarize the Description-Experience (D-E) gap literature from the adult Judgment and Decision Making field, which makes the crucial distinction between choice behavior when information is acquired to descriptions relative to experience. Next, I relate work on the D-E gap to laboratory research on risk taking between adolescents and adults. A review of the developmental literature demonstrates that experience-based experimental paradigms are more likely to show heightened risk taking in adolescents relative to adults (Rosenbaum et al., Resubmitted), and is consistent with an affect-based explanation of risk taking.

In Experiment 1, I present a novel within-subjects D-E gap paradigm, which I test in a sample of young adults, and show individual differences in the degree of bias when participants make choices from description verses experience. Subsequently, in Experiment 2, I test cohorts of adolescents and adults in the within-subjects D-E gap paradigm. In this developmental experiment, I additionally measure eye tracking to better understand decision processing and changes in heart rate variability by task (description, experience) and age group. Results show that adolescents and adults take similar risks in DFD and DFE, but unlike adults, adolescents' choices in DFD do not adhere to prospect theory predictions. Further, in DFD, adolescents spend more time looking at probabilities than values, while adults show the opposite pattern. Conversely, in DFE, adolescents make choices consistent with underweighting rare outcomes, similar to adults. There is some evidence that adolescents show enhanced rare-outcome

underweighting relative to adults, even after controlling for sampling bias. Concurrently, adolescents show a higher change in LFHRV from baseline relative to adults during DFE, but not in DFD.

In sum, results are consistent with the idea that adolescents have trouble utilizing descriptive information, but are able to adapt choices readily based on information acquired through experience. Teens, relative to adults, may show enhanced biases toward risk taking when a rare outcome is unfavorable, a process that may be supported by higher affective arousal.

ACKNOWLEDGMENTS

My dissertation would not have been possible without my wonderful mentors, Jason Chein, Vinod Venkatraman and Larry Steinberg. This project was a team effort, and being part of this collaboration was inspiring. Jason, thank you for continually encouraging my research efforts. You helped me grow immensely as a scientist and as a writer, and I am truly grateful. Thank you also for cultivating a lab environment in which I enjoy both my work and my colleagues. Vinod, I am thankful that we met and were able to work together beginning early in my graduate career. Thank you for consistently providing me with constructive feedback. Your involvement was essential to my research – you guided me toward the D-E gap literature that is the basis for this dissertation, and provided the opportunity for me to learn new physiology techniques that will serve me well in my future career. Larry, although I did not see you often, I learned something new every time we had lab meeting or exchanged emails. I am thankful to have worked with you for the last five years. I am also thankful to Vinay Parikh for serving on my core committee and providing valuable insights on my work from the animal literature, Crystal Reeck for helping me get my feet wet in physio research, and both Crystal and David Smith for serving as readers on my defense committee. Additionally, I would like to thank Ryan Velasco and Eugene Olkhov for help collecting data for this project, and Anthony Resnick for help with analyses in MATLAB.

My interest in research was sparked long before graduate school, and I am thankful to my mentors who inspired me to pursue research and helped me in the graduate application process (and beyond). Peter Todd, taking your class my freshman year at Indiana University set the course of my career. Thank you for working with me as an undergraduate, even though I had no idea what I was doing. I am glad I still study decision making so I can see you at Society for Judgment and Decision Making meetings. Yuhong Jiang, thank you for mentoring me as a post-bacc RA, guiding me through publishing my first papers, encouraging me to apply to graduate school, and for simultaneously making sure I knew the state of the faculty job market. I appreciate your honesty, and your outlook on research and life. I am lucky that I still get to see you when I

visit my family in Minnesota, and hope we remain in contact for years to come. Thanks also to Khena Swallow, Yaakov Kareev, Judith Avrahami, and Cara Wellman, who each helped me grow as a researcher.

I would not have survived graduate school without my Cheinberg labmates, who are both brilliant, and some of the most fun people I have ever met, including Harry Wilmer, Karla Fettich, Grace Icenogle, Natasha Duell, Karol Silva, Morgan Botdorf, Jamie Patrianakos, Lauren Sherman, Jamie Samper, Kara Blacker, Alex Morrison, and of course my work wife, Ashley Smith. To my fantastic fellow BCS women, Ashley Drew and Kylie Alm, thank you for commiserating and celebrating with me every step of the way. To my first Philadelphia roommate and phenomenal friend, Jill Rabinowitz, thank you for being sticking with me through all of our adventures in this program and this city. And speaking of Philadelphia, thanks to La.Va Café for providing me with a second office for the past five years, and to my Thrive trainer Abbey for keeping me sane.

Thanks to my friends outside of Temple for at least pretending to be interested in my research, and for being patient when I was working on this dissertation and had trouble staying in touch.

Additionally, thank you to Ellen Gartzman and Alan Sklarsky for providing me with a quiet space to work, free laundry, delicious dinners, and moral support.

Finally, thank you to my family for your support throughout graduate school. I wouldn't have gotten through this dissertation without your constant encouragement. Steven and Em, now you can't make fun of me for having the lowest degree in the family – I'm surpassing both of you, haha! Mom and Dad, I am forever grateful to you! Thank you for instilling a strong work ethic in all three of us, and for showing us what it's like to love learning. Mom – I hope we get an A on this dissertation! Dad – thank you for actually reading some of my papers!

Eli, you are my #1 cheerleader, and have helped me through this dissertation more than you know. I am a better person when I'm with you, and my life is infinitely better with you in it.

TABLE OF CONTENTS

Page

ABSTRACT.. ii

ACKNOWLEDGMENTS...iv

LIST OF TABLES..viii

LIST OF FIGURES..ix

CHAPTER

1. INTRODUCTION..1

Description-Experience Gap Literature...3

Adolescent Risk Taking and the D-E Gap...8

A Brief Summary of Developmental Studies of Risky Choice through a Description-Experience 14

An Affect-based Model for Risky Choice in Adolescence.................................20

Alternative Explanations..23

Disentangling Prospect Theory Biases from Risk taking.....................................26

Summary and Motivation for Following Empirical Studies..................................28

2. EXPERIMENT 1: A WITHIN-SUBJECTS DESCRIPTION-EXPERIENCE GAP IN ADULTS...31

Methods...32

Results...39

Discussion...46

3. EXPERIMENT 2: TESTING THE D-E GAP IN ADOLESCENTS AND ADULTS.................49

Aims and Hypotheses... 50

Methods... 53

Behavioral Results... 66

Physiological Results…………………………………………………………………………82

Discussion……………………………………………………………………………………91

4. GENERAL DISCUSSION………………………………………………………………… 98

Relating Results from Experiments 1 and 2……………………………………………99

Results in Relation to Past Adolescent Risk Taking Literature

and Prospect Theory ………………………………………………………………100

Study Limitations and Future Directions………………………………………………104

REFERENCES CITED……………………………………………………………………………107

APPENDIX

COMPREHENSION QUESTIONS FOR D-E GAP PARADIGM…………………………………127

LIST OF TABLES

Table Page

1. Tasks discussed in qualitative analysis of adolescent vs. adult risk taking studies……………………………………………………………………….......... 17
2. List of gambles used in the study, classified by problem types……...……………...……......34
3. D-E gap by problem…………………………………………………...……………...……......41
4. Group-level D-E gap by problem type……………………………..……………...……...... 42
5. Racial composition of adolescent and adult samples………………..……………...……...... 62
6. Group-level D-E gap in adolescents and adults for single-domain problems………..……......68
7. Group-level D-E gap in adolescents and adults for mixed-domain problems………..……......69
8. Predicting DFD risk taking in adolescents and adults………..………..……………..……......77
9. Predicting risk taking across all DFE trials………..………..……………………...……......78
10. Predicting DFE risk taking on all-outcome trials………..………..…………………..……....... 79
11. Predicting risk taking from gaze time on each AOI………..………………………..……......86
12. Correlations between self-report measures and the D-E gap by age……………..……….......90
13. Summary of hypotheses and data……………..…………………………..………..……….....91

LIST OF FIGURES

Figure Page

1. Illustration of the Spectrum of Affect model……………………..……………………….……….. 23

2. Group-level D-E gap by problem type ……………………..…..…….……………………… 40

3. The within-subjects D-E gap measure …….……...……………………..…..……………….43

4. Correlation between within-subjects D-E Gap measure and

false alarm rate during Go/No-Go.………………….........…………………..…..………….. 45

5. Choices consistent with PT predictions after controlling for

sampling bias in DFE…………………………………………………………………..……… 46

6. Illustration of the D-E gap paradigm……………………………………………………………. 55

7. A schematic of the study design for Experiment 2………..……………………..…..…………58

8. Biometrics setup……………………..…..………………………………..…..…………………… 59

9. Within-subjects D-E gap measure by age group and problem type……………………….…. 71

10. Choices in single-domain problems in DFD and DFE consistent with PT……………………... 73

11. Choices in mixed-domain problems in DFD and DFE consistent with PT…………………….. 74

12. Odds ratios for risk taking in DFD by problem parameter……………………..…..…………… 77

13. Odds ratios for risk taking in DFE by problem parameter……………………..…..…………… 78

14. Odds ratios for risk taking in DFE by problem parameter for trials where

all outcomes were observed………………..…..…….……………………..…..…….……… 80

15. Sampling in DFE and shape estimation tasks by age group……………………..…..………… 81

16. Estimation error by distribution type and by age in the shapes estimation task………………. 81

17. Eye tracking results in DFD……………………..…..………………………………..…..………... 82

18. Single-domain eye tracking data……………………..…..…….………………………...…..….. 84

19. Gaze time by age group……………………..…..……………………………..…..……………… 85

20. Odds ratios predicting risk taking by gaze time……………………..…..……….……………… 87

21. Low-frequency heart rate variability by task and age……………………..…..………………… 88

22. Change in low-frequency heart rate variability from baseline 1 to baseline 2 by first task and age group……………………………………………………………………… 89

CHAPTER 1

INTRODUCTION

Although adolescents are at the peak of their physical health, there is a spike in morbidity and mortality during the teenage years, largely from preventable behaviors connected to risky decision making (e.g., automobile crashes, delinquency, unprotected sex, binge drinking, illicit drug use; Kann et al., 2012). Accordingly, substantial resources have been invested in trying to understand why adolescents engage in such excessive risk taking, and how these behaviors might be prevented (Reyna & Farley, 2006).

Many efforts to reduce adolescent risky decision making focus on instruction-based interventions, which have frequently proven unsuccessful or of limited efficacy (Steinberg, 2008). Perhaps the most obvious example of an ineffective intervention is the Drug Abuse Resistance Education (D.A.R.E.) program. D.A.R.E. is one of the most widely implemented drug use prevention programs, reaching students in 75% of US school districts (D.A.R.E. America). D.A.R.E., like many other prevention programs, focuses on giving students more knowledge about drugs, often through descriptions of the possible negative outcomes of drug use. Despite its broad adoption and intuitively appealing approach, several large-scale studies indicate that D.A.R.E. has little to no efficacy in preventing drug use in adolescents (Clayton, Cattarello, & Johnstone, 1996; Ennett, Tobler, Ringwalt, & Flewelling, 1994; West & O'Neal, 2004; Wysong & Wright, 1995). Similarly limited efficacy has been found for other classroom-based interventions. While there are occasionally promising results (e.g., with programs focused on reducing sexual risk taking through instruction; Downs et al., 2004; Griffin, Botvin, & Nichols, 2006; Jemmott III, Jemmott, & Fong, 2010; Kirby & Laris, 2006; Reyna & Mills, 2014), most approaches designed to diminish adolescent risk behavior through instruction have yielded disappointing results.

One plausible explanation for the limited efficacy of D.A.R.E. and other instruction-based interventions is that they do not adequately consider the importance of the informational sources that guide adolescents' decisions about risk. While information acquired through instruction (i.e.,

explicit *descriptions* of risk outcomes) may play a role in decision making, such information may be utilized differently than information acquired through the decision-maker's own personal *experience*. For instance, it is possible that adolescents are unable to or are biased away from using descriptive information when rendering their decisions, especially when information is available via experience.

In the present work, the remainder of Chapter 1 will comprise an introduction as well as a short review of decisions from description and experience, as they may relate to adolescents' heightened risk-taking propensity. A much more extensive version of this review is currently resubmitted for publication as Rosenbaum, Venkatraman, Steinberg, and Chein (Resubmitted). Chapter 2 will discuss Experiment 1, which presents a novel paradigm to test decisions from description and experience within-subjects. This experiment is currently in revision as Rosenbaum, Chein, and Venkatraman (In Revision). Next, Experiment 2 in Chapter 3 tests adolescents and adults in the decision making paradigm developed in Experiment 1. Finally, Chapter 4 will include a general discussion and conclusions.

To lay a foundation for the empirical work described herein, I first examine prior research that explores how decision making is affected by the nature of information that informs the decision, whether through description or experience, and consider how accounting for the disparate impacts of these alternate sources might inform the development of better prevention and intervention strategies. Specifically, I review relevant work from judgment and decision making (JDM) studies exploring how decision making varies after learning about risk through description alone, experience alone, and when descriptions accompany experience. This review focuses in particular on the Description-Experience (D-E) gap, a phenomenon that has been examined extensively in other recent reviews (see e.g., Camilleri & Newell, 2013; de Palma et al., 2014). In Experiment 1 (Chapter 2), I then present results from a novel within-subjects paradigm developed to study individual differences in the D-E gap. Experiment 1 assesses the value of this paradigm in a sample of adults.

Despite considerable research on the D-E gap in normal adults, very few studies have looked at its relevance to adolescent decision making (Defoe, Dubas, Figner, & Van Aken, 2015; Hartley & Somerville, 2015; c.f. Rakow & Rahim, 2010; van den Bos & Hertwig, 2017; van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012). To understand how adolescents' decision making may change as a function of described vs. experienced contexts, I begin by mining the emergent patterns from a rich body of literature that explores differences between teens and adults, across a diverse range of laboratory risk-taking paradigms in the latter portion of Chapter 1. In doing so, I re-evaluate the developmental risk-taking literature with respect to the nature of information used in the risky-choice tasks – that is, by contrasting the findings from studies of adolescents versus adults derived from tasks that explicitly describe risks to participants with those that require participants to learn about risk outcomes through accumulated task experiences. Through this review, I demonstrate the relevance of the way in which risk information is acquired in accounting for the presence or absence of age differences in risk behavior.

In Chapter 3, I present a multi-methodological study (Experiment 2), which was designed to directly test the impact of the source of decision information (description, experience) on laboratory risk taking in adolescents and adults. The study additionally uses physiological measures to expose variation in affective responses during choice behavior in these tasks. Finally, in Chapter 4, I discuss my results in the context of past literature, and implications of this study for both adolescent risk taking research and intervention efforts.

Description-Experience Gap Literature

Judgment and Decision Making (JDM) studies have extensively investigated decision making and risk taking in the context of varying modes of information acquisition. Historically, JDM studies have asked participants to make risky decisions based on explicit descriptions of the probabilities and associated costs (these paradigms have come to be referred to as Decisions from Description, or DFD, paradigms). In a traditional decision problem, participants may be given

a choice between gaining $3 for sure, and taking a risk with an 80% chance of gaining $4 but a 20% chance of earning $0. While this simple type of problem may not seem immediately relevant to the sort of decision making that takes place in daily life, one can equate the problem to a real-life choice in which taking a risk confers a high probability of a favorable outcome and a low probability of an unfavorable outcome. For example, most of the time, driving above the speed limit allows one to reach the destination sooner, but in rare cases, can lead to a speeding ticket (or worse).

Cumulative Prospect Theory (henceforth referred to as PT; Tversky & Kahneman, 1992), a highly influential model of decision making in DFD (de Palma et al., 2014), predicts that people will typically be risk-averse when faced with high probability favorable outcomes and low probability unfavorable outcomes, but risk-seeking when faced with low probability favorable outcomes and high probability unfavorable outcomes (Tversky & Kahneman, 1992). In the context of the aforementioned decision problem, PT predicts that participants will tend to overweight the rare likelihood of the unfavorable $0 outcome, even when the risky option also confers a favorable $4 gain most (80%) of the time, leading the majority to select the safe $3 alternative (Hertwig, Barron, Weber, & Erev, 2004).

Although PT has some shortcomings (e.g., it is limited in explaining behavior in certain types of framing problems; Kühberger & Tanner, 2009; Reyna, Chick, Corbin, & Hsia, 2014), it remains the predominant model in understanding DFD behavior. However, over the last 10 years, the field has begun to redesign some of its experimental paradigms in recognizing that real-world decision making is often based at least in part on information obtained from past experiences in the same or similar contexts. Accordingly, a new class of experimental paradigms has emerged – those involving Decisions from Experience (DFE). In DFE paradigms, which some have argued are more similar to real-world decision making than DFD (Hertwig et al., 2004), participants must actively acquire information about the possible outcomes by repeatedly sampling the alternatives and observing the decision outcomes (i.e., by accumulating personal experience). Surprisingly, in

DFE paradigms, the pattern of decision preferences is generally opposite to that predicted by PT: participants tend to make decisions consistent with the underweighting of rare outcomes (Hertwig et al., 2004). In the lottery example above, this would translate to choosing the risky option over the safe option (i.e., *under*weighting the chance of a $0 outcome and taking a risk in hopes of gaining $4, rather than taking the sure $3). Interestingly, this flipped pattern more closely resembles many behaviors observed in everyday decision making. To return to the speeding example, speeding is an exceptionally common behavior, which suggests that drivers tend to overweight the likelihood of the higher probability outcome (reaching the destination more quickly) and underweight the likelihood of the rare outcome (a speeding ticket; Camilleri & Newell, 2011a) – a pattern opposite to that predicted by PT.[i]

A large number of recent studies have explicitly demonstrated the difference in behavior that occurs between described and experienced decision problems (Camilleri & Newell, 2013a; de Palma et al., 2014). The disparate choices of those in a DFD condition compared to those in a DFE condition is known as the Description-Experience gap (D-E gap; Camilleri & Newell, 2013). While the D-E gap is robust, there is some debate about its cause and most appropriate interpretation. For instance, in some cases, participants do not experience (or sample) every possible outcome in the distribution, especially when one outcome is particularly rare (very low probability; Fox & Hadar, 2006; Hau, Pleskac, Kiefer, & Hertwig, 2008). Accordingly, some have suggested that the D-E gap can be fully accounted for by the "sampling error" that arises when one's experiences under-represent (or over-represent) the true underlying probabilities (Fox & Hadar, 2006; Hadar & Fox, 2009). Some experimental manipulations aimed at eliminating sampling error have, indeed, also eliminated the D-E gap (Camilleri & Newell, 2011a; Hadar & Fox, 2009; Rakow, Demes, & Newell, 2008). However, a number of other studies have found that

---

[i] It is important to note that sometimes this bias for underweighting rare events can only be considered as *underweighting* when it is considered relative to behavior in description-based choices, and some have argued that DFE actually makes estimates that more accurately reflect actual probabilities (de Palma et al., 2014; Hilbig & Glöckner, 2011; Ungemach, Chater, & Stewart, 2009).

the D-E gap remains significant, though diminished, after accounting for sampling error (Camilleri & Newell, 2011b; Hau, Pleskac, & Hertwig, 2010; Hau et al., 2008).

Researchers have also interpreted the D-E gap with respect to memory limitations (Camilleri & Newell, 2013a). In the DFE condition, the extended history of outcomes must be maintained and accessed from memory in order to guide subsequent decision making. Since recent information tends to be more easily accessed and retrieved, participants may exhibit a "recency bias" in which more recently presented outcomes have a relatively greater influence on decision making (Barron & Yechiam, 2009; Hertwig et al., 2004; Rakow et al., 2008). However, the D-E gap has been found to persist (albeit in a potentially reduced form) even when the order of outcomes is experimentally controlled to account for recency effects, and when participants are given an exact record of the entire sampling history (so that this history is fully accessible at the time of decision making; Camilleri & Newell, 2011a; Gottlieb, Weiss, & Chapman, 2007; Hau et al., 2010). Thus, while both sampling error and recency bias likely contribute to the D-E gap, the D-E gap is observed even when these factors are taken into account.

Although the JDM field has focused considerable attention on the distinction between decisions made from description (DFD) and those made from experience (DFE), there are many instances in life when people have access to both descriptive information and information derived from prior experiences. For example, a doctor may acquire information about the likelihood of a patient having a disease, given a certain pattern of symptoms, both from a textbook description and from her experience seeing patients with similar symptoms. To assess how decision making might vary based on the availability of both descriptive and experienced information, a few studies have deployed a paradigm in which participants are asked to make repeated decisions from description either with, or without, experiential feedback regarding the outcomes of those choices. Jessup, Bishara and Busemeyer (2008), for instance, compared patterns of choice when participants were asked to make repeated decisions from description with feedback (i.e., the outcome of the chosen option was revealed) vs. without any feedback. They found that

participants who received feedback underweighted rare outcomes relative to those in the no-feedback group. That is, those who received feedback behaved the way participants in a standard DFE paradigm (experience only) tend to behave, even though they were also provided with full descriptive information. A study by Lejarraga and Gonzalez (2011) built upon the study by Jessup and colleagues (2008) by examining behavior in experience-only (with feedback), description-only, and description plus experience (repeated feedback with descriptions) conditions in a lottery game. They found that participants in the description plus experience condition behaved more like those in the experience-only condition than those in the description-only condition. Put together, these studies indicate that in paradigms with repeated decisions from description with feedback (i.e., experience), participants tend to weight the experienced information more heavily over time, leading to underweighting of rare outcomes (Yechiam, Barron, & Erev, 2005). Some have even found that this pattern can be further exaggerated when information about the foregone outcome is also provided. (When feedback on both the chosen and foregone outcomes is provided, this is referred to as "full feedback," as opposed to "partial feedback" when only the chosen outcome is shown; Yechiam & Busemeyer, 2006.)

Other relevant work has explored the relative timing of descriptive and experienced information and the differential impact on choice behavior. In naturalistic settings, descriptive information may not always precede experience. One study addressed this issue by varying, relative to the opportunity to engage in experience sampling, the timing of descriptive information in the form of a "warning" about the possibility of a rare negative outcome (Barron, Leider, & Stack, 2008). Participants were given a warning about the possibility of a $15 loss (which occurred with a probability of only .001, so it was rarely experienced) either "early," before experience sampling began, or "late," after half of their sampling decisions had been made. Those who received an early warning took fewer risks compared to those who received a late warning, and the late warning was not enough to prevent participants from underweighting the rare negative outcome. This continued underweighting of the rare event after a late warning was

even more pronounced when participants received full feedback after every sampling decision, perhaps again due to the emphasis on the rarity of the unlikely outcome.

Putting the findings together, there seems to be something psychologically different about repeatedly experiencing outcomes from decision alternatives, and this leads to different decision making compared to circumstances in which decision alternatives and their consequences are merely described. The ability of experience-based information to outweigh description in experimental paradigms, and the parallels between experimental experience-based paradigms and naturalistic contexts that involve repeated decisions provide key insights that can inform understanding of risk behavior during adolescence, a point I will address extensively in the upcoming section.

Adolescent Risk Taking and the D-E Gap

As discussed previously, teens are known to take more risks than adults in many real world contexts. While there is a large literature exploring the specific developmental mechanisms that might explain this phenomenon, and a growing literature on the D-E gap in adults, there has, for the most part, been a disconnect between the JDM literature and examinations (both empirical and theoretical) of age differences in risk taking (Hartley & Somerville, 2015). Specifically, the question of whether descriptions and experiences differentially affect *adolescent* decision making has gone almost entirely unexplored. Accordingly, I next turn my focus to a rich literature surrounding adolescent decision making, but re-examined through the lens of prior work on the adult D-E gap. In particular, I examine how adolescent decision making varies when the information that guides it has been acquired via descriptions versus experience. In this reevaluation of the adolescent risk-taking literature, I explore whether the prevailing influence of experience over description is observed, and perhaps amplified, among teens as it has been among adults.

Based on both experimental and epidemiological data, there is reason to believe that adolescents may actually exhibit larger behavioral differences in DFE relative to DFD than adults.

For example, adolescents have more difficulty than adults in updating their beliefs after being given explicit instruction about risk likelihood (Moutsiana et al., 2013). Moreover, teens engage in heightened risk taking even in situations where they clearly possess descriptive knowledge of the risks involved, and even when that descriptive knowledge overestimates the true likelihood of negative outcomes (an effect often attributed to a cost-benefit analysis between very likely rewards and unlikely consequences; Reyna & Farley, 2006). Perhaps the impact of ones' own experiences, even in the face of factual information, is more salient during adolescence and thus overshadows descriptions, making these described risk probabilities more difficult to learn (Albert & Steinberg, 2011; similar to results in studies investigating repeated DFD, Jessup et al., 2008).

To date, three studies directly or indirectly test D-E gap in adolescents. In the first, (Rakow & Rahim, 2010), three experiments tested the D-E gap across different age cohorts: 1) children vs. adults, 2) children vs. adolescents, and 3) younger adolescents vs. older adolescents. Due to the wide age range of participants in the study, the authors used a child-friendly DFE task, in which participants drew colored cubes from boxes rather than sampling numbers from "Point Machines" or "Money Machines," as the task is generally implemented in adult-only samples. Each color cube represented a different point value. Participants were required to sample from each box 10 times (with replacement) before making a final decision about which box they preferred. They also completed the DFD paradigm, where the values and numbers of each color cube were explicitly described as a frequency (e.g., this box has 9 red cubes worth 10 points each and 1 blue cube worth 0 points). In the child vs. adult and child vs. adolescent experiments, DFD and DFE conditions were presented between subjects, while in the younger vs. older adolescent experiment, DFD and DFE were both presented to the same participants in counterbalanced order.

The authors found a D-E gap in adolescents and in children, although the size and significance of the gap varied greatly across problems and across the three experiments (Rakow

& Rahim, 2010). However, the study found no significant difference in the overall D-E gap when directly comparing younger versus older groups, for any of their developmental comparisons.

When interpreting these results, it is important to consider some limitations of the study. Foremost, in the experiment that compared adults to children, the overall D-E gap was not significant among the adults, which is noteworthy given that most experiments using similar paradigms find a robust D-E gap in adult samples. The null result may have been due to the use of a modified paradigm requiring fixed samples from each alternative (which also led to a null result in Hadar & Fox, 2009). Additionally, whereas my present interest is in directly contrasting a D-E gap in adolescents to that in adults, the methods employed in Rakow and Rahim's (2010) study precluded a direct comparison of behavior between the adolescent and adult participants, as these two age groups were never directly paired in any of the experiments, and the authors employed a different set of decision problems within each of their three experiments. Moreover, the adult D-E gap was examined in a between-subjects design, whereas that for adolescents was tested using a within-subjects design. Thus, although important for having demonstrated, for the first time, the presence of a D-E gap in an adolescent sample, Rakow and Rahim's (2010) investigation provides no direct evidence regarding how the use of descriptive and experienced information might differ between adolescence and adulthood.

A more recent study by van Duijvenvoorde and colleagues (van Duijvenvoorde, Jansen, Bredman, & Huizenga, 2012) provides further insight into how adolescents might differentially respond to descriptive and experience-based information, and provides a direct comparison to an adult group. This study did not employ standard DFD or DFE paradigms, but rather, had participants play a modified Iowa Gambling Task (IGT). In the task, participants were asked to choose which of four different machines they wished to play. Two of the machines were "good" and provided subjects with smaller gains coupled with smaller and/or less frequent losses, such that the expected value (EV) of both machines was positive. The other two machines were "bad" and yielded higher gains coupled with larger and/or more frequent losses, giving them negative

EVs. In an *informed* condition, participants were shown the true underlying probability distributions in a graphic form, making this condition akin to a DFD paradigm. In the *uninformed* condition, participants needed to learn the probability distributions by trying each machine, similar to a DFE paradigm. This study included participants aged 7 to 29, and many of the analyses focused on differences across the spectrum of development rather than on a contrast between specific age groups (e.g., adolescents vs. adults). However, careful examination of the figures in the paper suggests that in the informed (DFD-like) condition, adolescents performed more similarly to adults than in the uninformed (DFE-like) condition (van Duijvenvoorde et al., 2012). In thinking about how risk is defined in behavioral economics – the pursuit of the alternative with the highest outcome variance (i.e., largest variance around the expected value; Weber, Shafir, & Blais, 2004) – choices from the "bad" machine with the highest outcome variance are of particular interest (Schonberg, Fox, & Poldrack, 2011). In the last 60 trials, adolescents chose this specifically riskier machine less often than chance only in the informed, but not in the uninformed, condition. This pattern is consistent with the notion that age-related differences may be less apparent in tasks that emphasize descriptive information, than in those in which participants' choices are more influenced by prior (or accumulating) experience.

Last, a very recent study tested both DFD and DFE in participants ages 9 to 22 (van den Bos & Hertwig, 2017).[ii] Like van Duijvenvoorde et al. (2012), this study used continuous age measures without directly contrasting adolescent to adult behavior. Interestingly, the authors found a quadratic relationship between risk-taking and age in DFD in the gain domain, such that adolescents showed more risk taking than younger or older individuals. In loss problems, there was a linear decrease in risk taking across age groups, with adolescents displaying intermediate risk taking between children and adults.

[ii] This study was published after data collection for Experiments 1 and 2 was completed, and after resubmission of Rosenbaum et al., (Resubmitted).

The van den Bos and Hertwig (2017) study did not report risk taking behaviors within DFE by age group (their focus was on the search process in DFE by age, which will be discussed later). However, the authors reported data on D-E gap in the supplemental materials, and showed a quadratic pattern in the D-E gap by age when the rare event was unfavorable, translating to heightened risk taking in adolescents in this type of problem. This pattern persisted, even after accounting for sampling bias. This heightened underweighting of rare events in DFE maps onto teens' dangerous risk taking behaviors in the real world. For instance, when adolescents choose to use drugs, despite the rare possibility of overdose (a rare unfavorable outcome), they are behaving as if they underweight the overdose outcome. The fact that laboratory behavior in DFE mirrors the risky behaviors that interventions attempt to attenuate hints that DFE (and differences in choice between DFE and DFD; i.e., the D-E gap) may be a fruitful way to study adolescent risk taking in the laboratory.

In developing a theoretical account of the D-E gap among adolescents, it is important to bear in mind the possible contributions of sampling bias and recency effects; two factors that have been shown to affect the size of the D-E gap in adults (as discussed previously). Related evidence suggests that adolescents might be more prone to both of these effects than are adults. Adolescents are known to be more impulsive than adults, a factor thought to contribute to their heightened risk taking (Casey, 2015; Luna, Garver, Urban, Lazar, & Sweeney, 2004; Shulman et al., 2016). Such heightened impulsivity could lead adolescents to render decisions after fewer samples in a DFE sampling paradigm, leading to higher sampling error. Indeed, van den Bos and Hertwig (2017) found lower sampling in adolescents relative to adults during DFE. Although the direct relationship between sampling volume and the D-E gap was not reported in that study, the D-E gap was higher in adolescents even in the subset of problems where participants experienced the rare outcome. This result suggests there may be a more pervasive underlying rare-underweighting bias underlying DFE in adolescents beyond that induced by sampling error.

Likewise, developmental differences in working memory are important to consider with respect to the D-E gap, since DFE tasks require participants to hold previously experienced outcomes in mind in order to inform their final decisions (Camilleri & Newell, 2013a), a process that is affected by memory capacity (Kareev, 1995, 2000). Accordingly, a relatively immature capacity for working memory could increase bias towards only the more recently presented outcomes that are still accessible in memory. Although some studies indicate near adult-like working memory and brain function in mid-adolescence (Crone, Wendelken, Donohue, van Leijenhorst, & Bunge, 2006), others provide evidence of a more protracted period of development extending into adulthood (Andre, Picchioni, Zhang, & Toulopoulou, 2015; Bunge & Wright, 2007; Jenkins, Myerson, Hale, & Fry, 1999; Luciana et al., 2005). Where links between the developmental status of working memory and decision making has been more directly explored, the data suggest that the role of working memory is small. Crone and van der Molen (2004), for example, showed that working memory span was not significantly related to performance on the Iowa Gambling Task, an experience-based task. Likewise, in Rakow and Rahim (2010)'s developmental study, a fourth experiment (not discussed earlier) contrasted adolescent and adult participants' memory of a sampling distribution and found that performance on the memory test did not differ significantly between adolescents and adults. Further, van den Bos and Hertwig (2017) found linear improvement in working memory tasks between childhood and adulthood (although in examining supplementary figures, the difference seems to be driven by differences between children and adults), but this difference was not related to risk taking in either DFD or DFE. So, while developmental differences in both sampling and recency effects should be explored systematically in future studies, these factors do not appear to be critical to the exposition below.

A Brief Summary of Developmental Studies of Risky Choice through a Description-Experience Lens

As noted above, very little research has been conducted in developmental samples using the traditional DFD and DFE paradigms that form the backbone of work on the D-E gap. There is, however, a relatively large literature comparing adolescents' propensity for risky decision making to that of adults. In this section I revisit this corpus of work while taking into account the distinction between decisions made from description and those made from experience. Specifically, I consider whether the observation of age differences in a given study depends on the type (description or experience) of paradigm that was used.

A recent meta-analysis (Defoe et al., 2015) provides some initial grounding on this issue. The authors examined laboratory risk-taking from childhood to adulthood and determined that, as expected, adolescent risk-taking in laboratory tasks is overall significantly elevated relative to adult risk-taking, despite a few individual studies that do not show this age-related difference. The authors also found that this age difference in risk taking is accentuated in tasks providing immediate, relative to delayed, feedback about decision outcomes, although this moderator was only marginally significant. In their investigation, Defoe and colleagues (2015) also acknowledged the importance of the description versus experience distinction. They discussed how experience-based tasks are thought to be more affective, or arousing, in nature (a point also made by Figner, Mackinlay, Wilkening, & Weber, 2009) and noted that this could be important because heightened affect is often a contributor to adolescent risk taking (Steinberg, 2008). In experience-based tasks, heightened affect is likely induced by the tendency to rely more on a "gut" feeling rather than a cognitive calculation (Defoe et al., 2015). Conversely, Defoe and colleagues suggested that description-based tasks are more likely to be associated with non-affective or "cold" contexts, which call upon more deliberative processes (e.g., explicit calculation of expected values; Defoe et al., 2015; Figner et al., 2009). This difference in affect may make description-based tasks less likely to evince developmental differences. Upon consideration of the distinction between

description- and experience-based tasks, Defoe and colleagues (2105) made an effort to code studies along these dimensions (though few details are given with respect to how they went about classifying individual tasks). Unfortunately, after applying their eligibility criteria, the authors concluded that there were too few studies in each category (description, experience) to run a moderator analysis.

As discussed in further detail below, I find that task format may indeed be a crucial factor in determining whether or not a given study finds age differences in risk taking. Here, I briefly review how task format relates to the findings in studies that compare adolescent to adult risky decision making. In doing this, I am able to consider several studies that could not be considered in Defoe et al. (2015)'s formal meta-analysis due to a lack of adequate effect size information. I additionally identified several studies that may have been overlooked by Defoe and colleagues for other reasons, and several informative studies that have been published since that meta-analysis was completed.

To support this work, I identified all experiments that compared risk taking in adolescents (ages 11-18) to adults (ages 18 and older),[iii] and classified them according to whether the task used in the comparison was description- or experience-based. I classified any task that presented full descriptive information to a participant, either graphically or verbally, as a description-based task. Conversely, I classified tasks that did not present descriptive information, but rather required participants to learn about potential decision outcomes through experience in the task as experience-based. *Table 1* lists studies included in the resulting analysis.

For the purposes of the present assessment of the literature, if the authors reported that *any* of the adolescent groups took significantly more risks than the adult group in the sample, I classified the study as identifying a developmental difference, but was careful to indicate any caveats in the *Table 1* notes. Further, some studies found a developmental difference under

---

[iii] Some of the studies include 18-year-olds in the adolescent group, while others include 18-year-olds in the adult group. See *Table 1* for age ranges in all included studies.

certain risk conditions but not others (e.g., a difference when rewards were very high, but not when they were moderate or low; e.g., Reyna et al., 2011). If the authors reported a significant developmental difference in risk taking in any condition in the study, I indicated that a difference was present, but I noted variation in condition effects as a caveat to the result.

In total, I found 38 published studies, including 41 experiments and comparisons relevant to my analysis. Nineteen of the studies featured experiments using only a description-based task and 15 featured experiments using only an experience-based task. Additionally, the study I previously discussed by van Duijvenvoorde and colleagues (2012) included two tasks: one experience-based, and one with experience accompanied by description. A further study (Figner et al., 2009) also included two different tasks: one description-based task, and one in which description was accompanied by repeated choices (i.e., experience). One study used a task involving both descriptions and experience (van Duijvenvoorde et al., 2015). Finally, one study used both a DFD and DFE task (van den Bos & Hertwig, 2017). Thus, I identified 17 experience-based experiments, 21 description-based experiments, and 3 experiments using description plus experience tasks.

As is apparent from *Table 1*, an intriguing pattern emerges in comparing the age differences obtained across description and experience tasks: 15 out of the 17 experience studies showed an age-dependent difference in risk taking, such that adolescents took more risks than adults, at least for certain comparisons. In contrast, only 8 of the 21 description studies reported an age-dependent difference. In other words, experience-based tasks are over 12 times more likely[iv] to show age differences in risk taking relative to description-based tasks. Of the three D+E studies, one of them showed a clear age-related difference in the expected direction, while in the other two, no difference was apparent. This overall pattern of choice suggests that DFE paradigms may be more likely to reveal age differences in risk taking than DFD paradigms.

[iv] The odds of showing heightened risk taking in adolescents in DFE are 15/2 = 7.5; in DFD, 8/13 = .62; odds ratio = 12.19.

Table 1

Tasks Discussed in Qualitative Analysis of Adolescent vs. Adult Risk Taking Studies

First author	Year	Task	Description or Experience	Adolescent Ages	Adult Ages	Experience - more risk taking in teens than adults?	Description - more risk taking in teens than adults?
Aïte	2012	Soochow Gambling Task (Modified IGT)	Experience	12-13	18-32	Yes	
Beitz	2014	IGT	Experience	11-16	17-59	Yes	
Cassotti	2011	IGT	Experience	12-14	18-28	Yes	
Cauffman	2010	Iowa Gambling Task (Play or Pass version)	Experience	10-17	18-30	Yes[b]	
Crone	2003	Modified Iowa Gambling Task	Experience	12-13, 15-16	18-25	Yes	
Crone	2004	Hungry Donkey Task (Modified IGT)	Experience	13-15	18-25	Yes	
Ernst	2003	Iowa Gambling Task	Experience	12-14	21-44	No	
Overman	2004	Iowa Gambling Task	Experience	11-18	17-23	Yes	
Van Duijvenvoorde	2012	Modified IGT, Uninformed and Informed	Experience (Uninformed), D+E (Informed)	11-13, 12-14, 14-17	18-29	Yes[c]	No (D+E)
Bjork	2007	Chicken	Experience	12-17	23-33	No	
Gardner	2005	Chicken	Experience	13-16	18-22, 24+	Yes	
Chein	2011	Stoplight	Experience	14-18	19-22, 24-29	Yes[c]	
Kim-Spoon	2015	Stoplight	Experience	17-20	31-51	Yes	
Steinberg	2008	Stoplight	Experience	10-17	18-30	Yes[b]	
Braams	2015	BART	Experience	8-27 (Longitudinal data)		Yes	
Mitchell	2008	2BIT (Modified BART)	Experience	14-17	35-55	Yes	

Table 1, Continued

Figner	2009	Hot and Cold Columbia Card Tasks	D+E (Hot), Description (Cold)	14-16, 17-19	20+	Yes (D+E)	No
Van Duijvenvoorde	2015	Hot Columbia Card Task	Experience	16-19	25-34	No (D+E)	
van den Bos	2017	DFD and DFE	Description, Experience	9-22 (Groups not defined)		Yes	Yes
Barkley-Levenson	2013	Mixed Gambles Task	Description	13-17	25-30		No
Barkley-Levenson	2014	Mixed Gambles Task	Description	13-17	25-30		No
Burnett	2010	Probabilistic Gambling Task	Description	12-15.5, 15.5-18	25-35		Yes[b]
Engelmann	2012	Probabilistic Gambling Task	Description	12-14, 15-17	18-45		Yes[c]
Eshel	2007	Wheel of Fortune	Description	9-17	20-40		No[d]
Habib	2012	Wheel of Fortune	Description	13-15	18-24		No
Habib	2015	Wheel of Fortune	Description	13-15	18-23		No
Haddad	2014	Wheel of Fortune	Description	11-18	20-38		Yes[c]
Keulers	2011	Play-Pass Gambling Task	Description	12-13, 16-17	20-21		No
Harbaugh	2002	Risky-Safe Gambling Task	Description	14-20[a]	21-65		No
Reyna	2011	Framing Spinner Task	Description	14-17	18-22		Yes[c]
Paulsen	2011	Nonsymbolic Economic Decision Making Task	Description	15-16	18-32		Yes[c]
Paulsen	2012	Nonsymbolic Economic Decision Making Task	Description	14-15	18-36		No
Galván	2012	Cups Task	Description	14-17	18-21		No
Levin	2014	Cups Task	Description	12-13, 14-15, 16-17	Parents (M = 45.5)		Yes

Table 1, Continued

Van Leijenhorst	2008	Cake Gambling Task	Description	14-15, 17-18	25-30	No
Van Leijenhorst	2010	Cake Gambling Task	Description	12-14, 16-17	19-26	Yes[b,c]
Tymula	2012	Risky/Ambiguous Lottery Game	Description	12-17	30-50	No
Blankenstein	2016	Risky/Ambiguous Lottery Game	Description	10-12, 14-16	17-20, 21-25	No

Notes. [a]In this study, the mean age for the adolescent group was unusually high (19.6). [b]Only one of the adolescent groups (or a subset of the adolescent age range, in Cauffman et al., 2010 and Steinberg et al., 2008) differed from the adult group. [c]The developmental difference was only found in some conditions (e.g., only some problem types, only in social context, etc.). [d]In this study, a portion of the discussion surrounds the lack of a developmental difference, and indeed, the main effect of age on risk taking was not significant. However, in one of the problem types, there was a significant negative correlation between age and risk taking.

An Affect-based Model for Risky Choice in Adolescence

The present assessment of the relevant risky decision-making literature indicates that experience- and description-based laboratory tasks may be differentially effective in exposing age-related differences in risky decision making. In my earlier discussion, I briefly mentioned one possible explanation for this difference – namely, that experience-based decision contexts are more affectively arousing, or "hot", in nature than are description-based contexts, which may be inherently "cold" (unless some further manipulation of affect is used; Defoe et al., 2015; Figner et al., 2009).

Interestingly, despite widespread use of the term affect in accounts of adolescents' heightened risk taking, this term is rarely, if ever, explicitly defined. While emotion research has not reached a consensus on the definition of affect (Scherer, 2005), a useful framework comes from Duncan and Barrett (2007), who define the psychological construct as, "a basic, psychologically primitive state that can be described by two psychological properties: hedonic valence (pleasure/displeasure) and arousal (activation/sleepy)" (p. 1185). While not explicitly stated, discussions of affect as it relates to adolescent risk taking usually refer to elevated arousal or activation (independent of valence), and this is how I define affect in the discussion that follows.

In order to understand the importance of the distinction between affective and non-affective contexts in adolescence, it is informative to turn to a popular theoretical account of adolescent risky decision making. The dual systems model (as well as several similar models; see Shulman, Smith et al., 2016) suggests that adolescent risk taking stems from an imbalance between the development of affective brain circuitry that matures earlier, and shows heightened responsivity during adolescence, and a cognitive control system that undergoes a more prolonged and gradual period of maturation (Steinberg, 2008). Within this model, teens are more likely to take risks in affective contexts because their affective systems are more reactive than those of adults, and because adolescents' cognitive control system is still immature. In cold (i.e.,

non-affective) situations, adolescents make similar decisions to adults because adolescents' affective systems are less likely to be engaged (or more effectively regulated by cognitive control), and hence less likely to "interfere" with decision outcomes. Thus, research studies guided by this model often try to distinguish between risk taking in affective vs. non-affective laboratory task environments (Albert & Steinberg, 2011; Figner et al., 2009).

Importantly, there is not a clear consensus in the literature about what constitutes an affective vs. a non-affective task. As documented in *Table 1*, the IGT reveals a developmental difference in risk taking in 7 out of 8 studies. Consistent with this pattern, many have argued that the IGT, an experience-based task, relies on affective processing (Bechara, Damasio, Tranel, & Damasio, 1997; Damasio, 1996). Evidence for this position comes, for instance, from participants' heightened skin conductance responses (SCRs) before selecting from bad decks, an affective response that appears before they have developed explicit knowledge about which decks are good and bad (Bechara et al., 1997; Wagar & Dixon, 2006).[v]

Others have made similar claims about experience-based tasks being more affective in nature than descriptive tasks. For example, Ludvig and Spetch (2011) suggested that their repeated choice (experience) paradigm induced a "hot" decision-making context because it provided feedback to participants after every decision. In the adolescent decision-making literature, Figner et al. (2009) explicitly named the version of the CCT that gave feedback the "hot" version, and contrasted this version with a "cold" CCT, which did not provide feedback every round. Following suit, Defoe et al. (2015) extended this argument and claimed that dynamic ("experiential") tasks are by nature more "emotional" because they require learning over time and possibly incline participants to respond more based on feeling than on facts.

---

[v] It should be noted that the Somatic Marker Hypothesis (Bechara et al., 1997), which supports the idea that IGT behavior is driven by emotion, has been criticized. Specifically, some have shown that healthy participants possess explicit knowledge of the IGT deck contingencies, and have questioned the link between SCRs and risk behavior (for an extensive discussion, see Dunn, Dalgleish, & Lawrence, 2006).

The dual systems model suggests that the mechanism driving adolescents' greater risk taking is heightened activity in affective neural structures (Steinberg, 2008), which is consistent with the arousing decision-making context (Defoe et al, 2015; Figner et al., 2009). This account further explains why it is possible to increase or reduce developmental differences by making description-based tasks more affectively arousing, or experience-based tasks less so. As noted earlier, a social context manipulation can evoke age differences when description-based tasks are used – a pattern that arises when participants are observed (Smith, Chein, & Steinberg, 2014) or advised (Haddad et al., 2014) by others. Of course, the peer manipulation may not be the only way to induce heightened affect in an otherwise cold task. For example, I discussed earlier a study by Burnett and colleagues (2010) showing an age difference in risk taking in a descriptive probabilistic gambling task that required participants to rate their emotional responses after receiving decision feedback. It is possible that both decision feedback (a dimension previously shown to promote developmental differences; Defoe et al., 2015) and requiring an emotional rating promoted age differences in a task that would not have shown such a difference otherwise.

Keeping results such as these in mind, I propose a basic model, the Spectrum of Affect model, that integrates recent hypotheses about affective arousal during risk taking (Defoe et al., 2015; Figner et al., 2009; Hartley & Somerville, 2015) with the ideas about DFD and DFE explicated in greater depth above. Specifically, I propose that there may be a spectrum of affective arousal (*Figure 1*) along which decision-making tasks can be arranged. Tasks that are description-based, without any added affective manipulations, tend to be toward the "cold" end of the spectrum, but can be made "hotter" with additional manipulations. Tasks based on experience are usually relatively hotter by nature because of task features such as feedback, and can also be made hotter with additional manipulations. At a certain point on this spectrum, affect is high enough for a task to elicit a detectable age difference (in the case of experience-based tasks, this may occur even in the absence of further manipulation of affect).

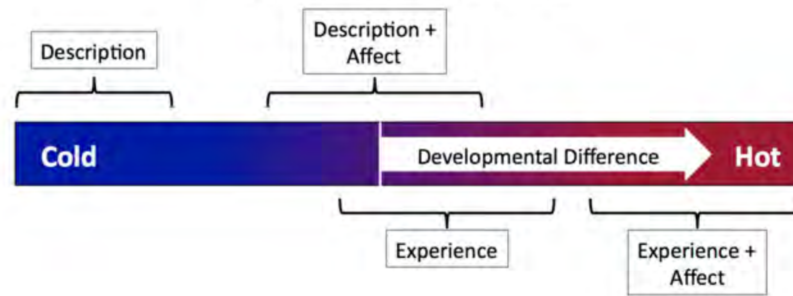


Figure 1. Illustration of the Spectrum of Affect model. This model suggests that a paradigm must reach a minimum threshold of affective arousal to evince a developmental difference. Experience-based paradigms are inherently more affectively arousing than description-based paradigms, but both can be made affectively "hotter" by adding additional affective manipulations (e.g., a manipulation of social context).

Alternative Explanations

I find an affect-based explanation to be compelling in explaining the disparate findings in description-based tasks relative to experience-based tasks. However, it is important to address possible confounding variables that might account for task-based differences in risk taking.

Expected Value

One potential confounding factor is adolescents' ability, relative to adults', to compute Expected Value (EV). The Iowa Gambling Task (IGT), and particularly the net score measure often used to conceptualize risk taking, has been criticized for its inability to disentangle EV sensitivity from risk taking, as the "bad decks" have lower EV than the "good decks" (Schonberg et al., 2011). Further, van Duijvenvoorde and colleagues (2015) arrived at a similar conclusion based on evidence from the hot CCT (which involves both description and experience): adolescents were less sensitive to EV than adults. Therefore, one might relate adolescents' higher risk taking on this task (relative to adults) to their lower sensitivity to EV, or alternatively, to their inability to compute EV. However, in description-based studies, developmental differences in sensitivity to EV are inconsistent (if they are reported at all). In their description-based study, Barkley-Levenson and Galvan (2014) found that adolescents were *more* sensitive to EV than adults, while Burnett and colleagues (2010) found the opposite pattern, and several studies have

reported no difference (Barkley-Levenson et al., 2013; Galván & McGlennen, 2012; Keulers et al., 2011). It is additionally worth noting that in many DFD studies, EV is equated in available alternatives (Haddad et al., 2014; Paulsen et al., 2012, 2011; Reyna et al., 2011). In other DFD studies expected value and risk are confounded, such that the riskier option (i.e., the alternative with the higher outcome variability) also has a lower EV (Eshel et al., 2007; Habib et al., 2012, 2015). Future studies should clarify the relative impact of EV sensitivity on developmental differences in risk.

Risk Taking Versus Sensation Seeking

While the present review focused on the *behavioral outcomes* (i.e., the risky behaviors observed in the lab or real-world) reported across studies, ideally experiments be able to more directly assess the *psychological states* that motivate, or drive, these behaviors (Shulman, Smith et al., 2016). This distinction is important because meaningful developmental differences in the drive to pursue risks may be present even when no behavioral differences emerge, such as when the opportunity to enact the behavior is limited by the experimental task or by real-world constraints on behavior. Dual systems explanations suggest that adolescents' greater inclination toward reward-seeking and sensation seeking may underlie their risk seeking. Unfortunately, most of the present studies confound risk taking with reward- and sensation-seeking, such that participants must take a risk to receive higher reward (Defoe et al., 2015). How these constructs relate to decisions from description vs. experience remains to be studied. It is possible, for example, that the arousing context of an experience-based decision heightens adolescents' drive to seek rewards, while the less arousing description-based context has the opposite effect. In this way, what emerges as differences in risk behavior are really driven by differences in reward sensitivity.

Attention/Task Engagement/Impulsivity

A further alternative account to the framework proposed here is one that is based on differences in attention and task engagement. For example, one could argue that experience-

based tasks are simply more exciting relative to description-based tasks, leading to differential levels of engagement across these tasks among adolescents, and accounting for differences in risk taking (i.e., more risk taking on more exciting tasks). Indeed, it is plausible that merely having heightened attention to a risk-taking task might lead to more risk behavior. However, the opposite prediction is also plausible, as higher affect is linked to impulsivity, which is often associated with reduced attention (e.g., individuals with ADHD, who have low attention, often show higher task-based arousal; Bjork & Pardini, 2014; Elkins, McGue, & Iacono, 2007). Future studies can help delineate the precise roles of arousal vs. attention/task engagement by careful examination of behavioral results (e.g., decision latency, sampling behavior, etc.), or through the use of eye tracking, neurophysiological measures, and self-reports of emotional states.

Relatedly, the relationship between impulsivity itself and DFD vs. DFE remains to be characterized. Importantly, impulsivity is not a uniform construct. In the laboratory, tasks measuring impulsivity can typically be dissociated into two (or more; Caswell, Bond, Duka, & Morgan, 2015) broad categories, including impulsive choice (placing higher value on outcomes that occur sooner to those that would be materialized in the future) and impulsive action (failure to inhibit a motor response, often a proponent response; Kim & Lee, 2011). Both impulsive choice and action are more pronounced in adolescents relative to adults in laboratory tests of impulsivity, and are thought to contribute to heightened risk taking in adolescents (Humphrey & Dumontheil, 2016; Olson, Hooper, Collins, & Luciana, 2007; Steinberg et al., 2008, 2009). If impulsivity is a mechanism of adolescent risk taking, and adolescents show higher risk taking in DFE than DFD, then impulsivity may also be related to the D-E gap, or be related differentially to DFE and DFD. The relation between the D-E gap, age, and measures of impulsive action and choice should be tested in future studies.

Ambiguity vs. Experience

The JDM literature also includes a related (but different) type of paradigm, known as decisions from ambiguity. In decisions from ambiguity, participants are asked to make between a

sure alternative and an ambiguous alternative for which a subset of the possible outcomes were undefined (results from the description trials, but not the ambiguous trials, were included in our earlier tally of study findings). Adults are known to be ambiguity averse (for a review, see Camerer & Weber, 1992) – they typically opt for the sure rather than the ambiguous alternative, even against their economic interests. In contrast, three recent studies have demonstrated that adolescents are more likely than adults to gamble on trials characterized by ambiguity (Blankenstein et al., 2016; Tymula et al., 2012; van den Bos & Hertwig, 2017). Indeed, just as in the ambiguous condition, real life decisions are often made when the full range of possible decision outcomes and their precise probabilities are unknown. While not an experience task per se, the ambiguity condition is analogous to DFE in that DFE participants often do not sample the full range of possible outcomes and thus render their choices under ambiguous probabilities (although DFE, is more accurately classified as a decision under *uncertainty*, where it is unclear whether all outcomes have been observed; Fox & Hadar, 2006; Hertwig et al., 2004). Following this evidence, one may theorize that adolescents' heightened risk taking is not driven by a tolerance for risk, but rather by tolerance for ambiguity, or their acceptance of situations in which all outcomes are not known or cannot be known (van den Bos & Hertwig, 2017). Indeed, van den Bos and Hertwig (2017) showed that adolescents are both more tolerant to uncertainty and to ambiguity. In sum, it seems that decisions under ambiguity and uncertainty (DFE) may be similarly relevant to adolescent risk taking. The nature of the relationship between choices from uncertainty and ambiguity is an interesting topic for examination in future studies.

Disentangling Prospect Theory Biases from Risk Taking

The Spectrum of Affect model makes predictions about risk taking in adolescents relative to adults, in DFD relative to DFE. Specifically, the model proposes that adolescents will take more risks in DFE relative to adults due to heightened affective arousal, but that risk taking by age group will be similar in DFD. Although the JDM/D-E gap literature motivated the distinction between DFD and DFE and my subsequent review of the developmental literature, it is imperative

to point out that the Spectrum of Affect model itself was driven primarily by the developmental literature on risk taking, and that the traditional D-E gap paradigm deviates from most paradigms used in the developmental literature. In particular, the standard D-E gap paradigm indexes relative biases in weighting outcomes by decision context (overweighting rare outcomes in DFD, and underweighting rare outcomes in DFE). Crucially, these paradigms contain some decision problems with rare favorable outcomes, and some with rare unfavorable outcomes. According to predicted weighting patterns, participants should make a *safe* choice when the rare outcome is favorable in DFE or when it is unfavorable in DFD.

Developmental DFE risk taking tasks, in contrast to D-E gap paradigms, typically only contain problems with *unfavorable* rare outcomes (if a rare outcome exists at all; often risky alternatives involve two equally probable outcomes). For example, in the Iowa Gambling task (IGT), the most commonly employed DFE task to contrast adolescent to adult decisions, risk taking behavior (in *Table 1*, quantified as a lower net score, or more draws from bad decks than the good decks) is associated with seeking a more common and higher reward relative to that available in the good decks, with the possibility of an unlikely but larger loss (Bechara, Damasio, Damasio, & Anderson, 1994). According to the predicted pattern, these types of problems in DFE (i.e., with a rare unfavorable outcome) would bias a participant towards risky rather than safe choices. Following this logic, a more pronounced rare-underweighting bias in adolescents (consistent with data in van den Bos and Hertwig, 2017) would produce the pattern of higher risk taking in adolescents relative to adults in DFE observed in *Table 1*. Consistent with this pattern, adolescents' dangerous risk taking (e.g., drug use) involves a rare unfavorable event (e.g., overdose).

It is unclear whether this systematic rare-underweighting bias would extend to DFE problems with a *favorable* rare outcome, showing *lower* risk taking in adolescents versus adults in these problems. An intriguing possibility is that rare outcome underweighting itself could be the mechanism of adolescent risk taking, and this behavioral pattern may be accompanied by

heightened affective arousal. In other words, if rare underweighting is the mechanism of heightened risk taking in adolescents, there may not be an overall difference in risk taking if a paradigm systematically tests problems with both favorable and unfavorable rare events.

Although there is little developmental data systematically exploring the weighting of rare outcomes in DFE, several developmental studies of risk taking have shown immature DFD weighting patterns in adolescents (Engelmann et al., 2012; Harbaugh et al., 2002; Tymula et al., 2012). This suggests that, while adolescents and adults might have similar mean risk-taking tendencies, decision *processes* during DFD could differ developmentally; in other words, the particular risks that adolescents and adults decide to take may differ systematically. An age difference in DFD processing is interesting from an intervention standpoint. If adolescents do not yet show adult-like biases in outcome weighting, a described intervention message (e.g., using drugs can lead to overdose) might not be perceived or processed in the same way across development.

In sum, the relationship between weighting of rare outcomes in both DFE and DFD, and in the D-E gap pattern, depends critically on problem parameters (i.e., the favorability of the rare outcome). This point is not necessarily inconsistent with predictions from the Spectrum of Affect model, but is not directly acknowledged by the model. Two separate predictions can be made based on the developmental and D-E gap literatures: 1. that adolescents take more risks in DFE regardless of the problem type, and 2. that adolescents' heightened risk taking in laboratory paradigms is more likely when rare outcomes are unfavorable, due to a pronounced rare-outcome underweighting bias. Neither of these patterns is inconsistent with the prior review, and either could be accompanied by heightened arousal in DFE.

Summary and Motivation for Following Empirical Studies

JDM studies of the D-E gap clearly show that variability in how information about risk is acquired can influence subsequent decision preferences (Hertwig et al., 2004), such that experience-based tasks lead to underweighting of rare outcomes (Camilleri & Newell, 2013a;

Hertwig et al., 2004), an influence that often overpowers that of described information (Jessup et al., 2008; Lejarraga & Gonzalez, 2011). Further, the patterns observed in DFE are, in many cases, more closely aligned with decision making and cognitive processes observed in daily life than those in DFD paradigms (Hertwig et al., 2004).

The distinction between description and experience is especially helpful in making sense of the adolescent risk taking literature. While description and experience tasks are rarely examined within the same sample and in the same study of adolescent decision making (Defoe et al., 2015; c.f. Rakow & Rahim, 2010; van den Bos & Hertwig, 2017; van Duijvenvoorde et al., 2012), my analysis of the literature examining developmental differences in risk taking suggests that experience-based tasks do, substantially more often, show developmental differences than do description-based tasks. This pattern is aligned with the idea that teens often make similar decisions to adults in situations that are deliberative and emotionally cold, but make riskier decisions than adults in affective situations more accurately simulated in experience-based tasks (Defoe et al., 2015; Figner et al., 2009).

The distinction between description and experience is informative in understanding the existing literature on adolescent decision making. However, one shortcoming of the current summary of the literature is that it involves the comparison of a wide variety of tasks with varying definitions of risk taking (also a limitation of the meta-analysis completed by Defoe and colleagues, 2015). As research on developmental differences in risk taking progresses, it will be important to directly empirically test the Spectrum of Affect model, using both description- and experience-based tasks within the same study, and adding other manipulations that may increase affective arousal (e.g., offering rewarding feedback). To understand whether affect is, indeed, an important factor in whether a task shows differences between adolescent and adult decision making, measures of task-evoked affective arousal must also serve as a manipulation check. This argument may at first seem circular (i.e., when not showing heightened adolescent risk taking, one can always claim that the task was not “affective enough”). However, using

physiological tools often used in the emotion literature (e.g., skin conductance, heart rate measurements, pupilometry, EEG; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2008) will help to clarify situations in which teens show heightened affective arousal. Additionally, self-reported emotional engagement may be a useful tool (e.g., Burnett et al., 2010).

Future research is necessary to better understand whether and under what circumstances teens' decisions based on described versus experienced information differ. I present two experiments in the following chapters of this dissertation. Experiment 1 (Chapter 2) tests a novel within-subjects version of the D-E gap paradigm in adults, which allows for measurement of each individual's bias in decision making based on problem format (rather than quantifying the D-E gap on a purely group level, as is the norm in the JDM literature). Experiment 2 (Chapter 3) uses this within-subjects paradigm to test the D-E gap in adolescents and adults, and to test many of the predictions raised in this present discussion. In Experiment 2, I additionally collect measures of physiological arousal to test the hypothesis that differences in experience-based choice (risk taking, or overall rare outcome underweighting) will coincide with heightened physiological arousal in adolescents.

CHAPTER 2

EXPERIMENT 1:

A WITHIN-SUBJECTS DESCRIPTION-EXPERIENCE GAP IN ADULTS

Despite progress in understanding the Description-Experience (D-E) gap within the adult Judgment and Decision Making (JDM) literature, as detailed in the review above, an important limitation is that DFD and DFE behavior are almost always contrasted between groups of subjects (c.f. Camilleri & Newell, 2009; for within-subjects studies using modifications of the traditional DFD and DFE paradigms, see also Abdellaoui, L'Haridon, & Paraschiv, 2011; Ludvig & Spetch, 2011; Rakow & Rahim, 2010, Experiment 3). Critically, both anecdotal evidence and the scientific literature show that there is considerable individual-to-individual variation in decision-making behavior (Kahneman & Tversky, 1979). Indeed, even when experiments yield a pronounced D-E gap, there is clear choice variability within both the DFD and the DFE groups themselves. Nevertheless, the standard between-subjects design cannot reveal whether the majority of individuals would actually reverse (or shift) their preferences according to presentation format, or whether only a few individuals drive the group-wise effect. Further, between-subjects designs cannot speak to the role of sampling bias across individuals; perhaps sampling bias can explain the D-E gap in some individuals, but not in others. In the broader developmental context motivating the present work, a traditional between-subjects design cannot address potential differences in DFD relative to DFE choice behavior in adolescents and adults, or how sampling bias may differentially influence decision making across development. Crucially, there is currently no paradigm for exploring individual differences in the D-E gap systematically.

The present experiment sought to develop and test a paradigm indexing the existence of a D-E gap *within* individual adult participants, for later use testing the Spectrum of Affect model in a developmental sample. To this end, I employed a design in which participants completed the same decisions from description and experience in counterbalanced blocks. I systematically varied the problem types to determine whether certain problem parameters affect the magnitude

of the gap, and to see whether choice patterns would be similar to those observed in previous between-subjects studies.

In characterizing individual differences in the D-E gap, I propose a novel metric that incorporates behavior in both description and experience conditions into a single index based on a subject's preference reversals across the two decision formats. By using this index as an individual differences variable, I subsequently explored whether the within-subjects gap variable varied systematically with an index of motor impulsivity. Last, I show that, for participants with a large D-E gap, choice behavior cannot be explained by sampling error alone.

Methods

Participants

Participants included 91 undergraduate students at Temple University (63 female, mean age = 20.71, SD = 5.28)[vi]. All participants received course credit, plus a small cash bonus (approximately $5) based on the outcome of one of their decisions drawn randomly from either the description or experience condition.

Materials

Decision Problems

The 18 problems completed by participants are listed in *Table 2*. Problems were analyzed individually, but were also organized into six problem *types* (A-F) in order to explore the specific task dimensions that influence the D-E gap. Across problem types, I varied the valence of outcomes (gain/loss), the favorability of the rare outcome (favorable vs. unfavorable), and the certainty of the alternatives (sure-risky/risky-risky). Within each problem type, I also varied the relative expected value of alternatives (equal/unequal). *Table 2* displays the various parameters of each problem and problem type. Importantly, the prior study that systematically varied problem

---

[vi] Participants were collected in two separate samples using the same methods, the first with 46 participants and the second with 45 participants. The size of the D-E gap at the group level and of the individuals did not significantly differ across the samples, so they were combined into a single sample for all subsequent analyses.

types in this way (Camilleri & Newell, 2013b) used a between-subjects design, involved repeated consequential decisions (rather than sampling), and required participants to make 40 choices with full feedback (including feedback about the foregone outcome) provided after each choice. The present study tests whether the results extend to a within-subjects free sampling paradigm without decision feedback.

Problems 1-12 (types A-D) formed the core problems in this study and included one risky (probability < 1) option and one sure option. The risky option always included a rare outcome (i.e. an outcome that occurred on less than or equal to 20% of choices for that alternative). These problems are similar to those often employed in D-E gap studies (Camilleri & Newell, 2013b). In fact, several of the problems used (problems 1, 4, 7, and 10) were taken directly from Hertwig et al. (2004)'s seminal DFE study. Based on the use of similar or identical problems in the past literature, the "expected" preference is defined as the alternative that would be favored under the assumption that rare events are relatively underweighted in DFE (and relatively overweighted in DFD, in accordance with PT). For example, in problem 7, participants are expected to "underweight" the rare 0 outcome in the risk, but overweight the 4 outcome, making the possibility of winning the larger sum more attractive relative to the sure 3 outcome when deciding based on experience. Alternatively, in problem 10, participants making decisions from experience are expected to underweight the rare 32 outcome, overweight the likelihood of a 0 outcome, and therefore prefer the sure outcome of 3 relative to the risky alternative.

In addition to these core problems, participants completed six supplemental problems (13-18, types E and F), each with two risky alternatives. Consistent with prior literature (Lejarraga, 2010), the D-E gap did not reach significance in these problems, so they are only briefly discussed.

Table 2

List of gambles used in the study, classified by problem types

Problem Type	Description	Problem Number	Option 1	Option 2	Option w/greater EV
A	Loss Framing Undesirable Rare Outcome (large loss)	1	-32, .1	-3, 1	2
		2	-40, .2	-8, 1	-
		3	-34, .2	-7, 1	1
B	Loss Framing Desirable Rare Outcome (0)	4	-4, .8	-3, 1	2
		5	-10, .9	-9, 1	-
		6	-14, .8	-12, 1	1
C	Gain Framing Undesirable Rare Outcome (0)	7	4, .8	3, 1	1
		8	10, .9	9, 1	-
		9	16, .8	13, 1	2
D	Gain Framing Desirable Rare Outcome (large win)	10	32, .1	3, 1	1
		11	35, .2	7, 1	-
		12	29, .2	6,1	2
E	Gain Framing Two probabilistic options	13	5, .2	3, .25	1
		14	15, .4	10, .6	-
		15	28, .25	9, .8	2
F	Loss Framing Two probabilistic options	16	-6, .2	-4, .25	2
		17	-3, .8	-12, .2	-
		18	-12, .5	-16, .4	1

Tasks

All experimental tasks were programmed in E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Participants completed the same 18 problems in description and experience conditions, with the order counterbalanced across participants, and a 4.5-minute continuous performance filler task in between formats (based on the Go/No-Go task employed in Tottenham, Hare, & Casey, 2011). In all problems, the position of the risky and safe option (on the left or right side of the screen) was counterbalanced. Additionally, problem types were semi-counterbalanced: participants were presented with 3 blocks of 6 problems. Each block contained one randomly drawn problem from each of the problem types listed in *Table 2*.

Problems were framed in terms of points to be won or lost from an endowed 50 points (which allowed for losses). All decisions were incentive compatible. Participants were told that

their winnings would be translated to money at the end of the experiment. At the end of the study, one of the participant's decisions was randomly selected and resolved according to the underlying probabilities, and the points were converted to money. The exact conversion equivalencies were unknown to the participants when making decisions, and were only disclosed at the end of the experiment.

Decisions from description. In decisions from description, participants saw two "Point Machines" and were told that they were "Describing" point machines. The machines, labeled 1 and 2, were shown along with explicit information about their respective probability distributions and corresponding outcomes. For those completing the study without eye-tracking, this information was presented in textual form; e.g., "40% chance to win 4 points, 20% chance to win 0 points [Neither win nor lose]," or "100% chance to win 3 points"). For the eye-tracking group, the possible outcomes were presented within a schematic grid that represented outcome likelihoods. The presentation format of the point machines was the only procedural difference between the two groups. For each problem, participants were asked to choose one of the two machines. A box around the chosen option was highlighted to indicate the participant's selection, and the next problem appeared 4s later. Due to a computer error, one participant only completed two problems in the DFD block and thus was excluded from all analyses.

Decisions from experience. The experience point machines were termed "Sampling" machines. Participants "sampled" or tried each machine until they felt confident that they knew which one they wanted to play. In each trial, two machines, labeled 1 and 2, were presented as in the description condition, but without any overt information about the outcome values or probabilities. Participants were told to press "1" to try machine 1 and "2" to try machine 2. When they felt they had acquired enough information to choose between the two machines, they pressed "D" to make a final decision. After pressing "D" they pressed "1" or "2" to choose machine 1 or 2, respectively. As in the description condition, the box around the machine was highlighted to indicate the selection, and the next problem appeared 4s later.

Go/No-Go

Between DFD and DFE blocks, participants completed a Go/No-Go task using cars as stimuli (e.g., “Go” for SUVs, “No-Go” for 4-door sedans). Stimuli were presented for 500ms each. Go and No-Go stimuli changed across 4 blocks of 30 trials, and No-Go stimuli presented on 30% of trials. Although the primary purpose of this task was to distract participants from problem parameters in DFD relative to DFE, false alarm rate (i.e., rate of No-Go errors) from this task was computed a measure of motor impulsivity.

Procedure

All participants provided informed consent before beginning the experiment. The order of description and experience tasks was counterbalanced with Go/No-Go in the middle of the two. For each task, participants were asked to read the instructions carefully, as they would be tested on the instructions at the end of the task session. After their second decision task (either description or experience), participants were asked to respond to three comprehension questions to ensure that participants were paying attention to the experiment. All of these questions and the multiple-choice answers are listed in the Appendix. Four participants who answered two or more of the three comprehension questions incorrectly were excluded from analyses, as it was unclear whether they were paying attention or understood experiment instructions. Participants were also asked whether they believed that the problems in description and experience conditions were a) exactly the same; b) similar, but not exactly the same; or c) completely different. Only 17 participants (19.54% of participants who answered the comprehension questions correctly) reported that the problems were exactly the same, and the difference in individual subject D-E gap variable between these participants and the remainder of participants did not reach significance, $t(83) = .49$, $p = .63$, $r^2 = .003$. Therefore, these participants are included in all subsequent analyses.

The computer then randomly selected the specific game (trial) upon which the participant’s bonus was based. Participants were informed about the outcome for that game

(estimated using a computer program based on underlying probabilities), and received a bonus payment based on a translation of the outcome from points to dollars (each point equaled $0.10, so the 50-point endowment was equivalent to $5). Bonuses were rounded up to the nearest quarter. On average, participants earned a bonus of $4.92 (*SD* = $1.00).

Individual Subject D-E Gap Calculation

A goal of the present study was to develop a summary statistic that could quantify the magnitude of an individual subject's D-E gap, based on how he or she responded to the description and experience problems. My starting point for characterizing the individual subjects D-E gap was determining the number of preference reversals evinced by each participant across the problem set. For a given problem, a participant could choose the same option in both DFD and DFE (no switch), or could choose different options across the two conditions, indicative of a preference reversal. The direction of the preference reversal is important. Based on the prior D-E gap literature, one would expect preference reversals to reflect a systematic overweighting of rare outcomes in DFD (consistent with PT), and a systematic underweighting of rare outcomes in DFE (opposite to PT). Switches consistent with these predicted biases could thus be treated as "expected switches," while switches in the opposite direction could be characterized as "opposite switches." In order to determine the systematicity of switching, and hence, the size of an individual's D-E gap, I used the following formula:

$$\frac{\#\text{ Expected Switches} - \#\text{ Opposite Switches}}{\#\text{ Problems}}$$

Calculated in this fashion the gap can range from +1 (a participant who switches in the expected direction on every problem) to -1 (a participant who switches in the opposite direction on every problem). A gap higher than 0 suggests that the individual exhibits more expected switches than opposite switches; consistent with someone who behaves in the way predicted by past between-subjects D-E gap studies. Meanwhile, a gap value of 0 signifies random switching or no switching, and a gap lower than 0 is reflective of someone who, unexpectedly, makes more

opposite switches than expected switches. Though the measure cannot distinguish between someone who makes consistent choices in all trials (no switch) and someone who randomly switches (6 expected and 6 opposite switches), I believe that this is not as critical as capturing individuals who show a consistent switching pattern in the expected direction. Only the core problems (1-12) were used in the computation of this gap measure since those problems were the only problems for which a strong directional prediction could be made based on prior studies looking at the D-E gap between subjects. Further, if a participant did not sample from each alternative at least once before making a decision, that problem was excluded from the calculation.

Data Analysis Specifications

Types of D-E gap. In the results section below, I report analyses on the D-E gap at two levels: the *group-level* and the *within-subjects* D-E gap. The group-level D-E gap refers to the number or proportion of choices in each problem predicted by past literature (i.e., underweighting rare outcomes in DFE, overweighting rare outcomes in DFD) across the entire sample. The within-subjects D-E gap refers to the individual differences metric detailed above.

Additional exclusion criteria. In specific data analyses, I needed to exclude additional participants for various reasons. There were several instances where participants did not sample each alternative at least once in a given trial; these trials were excluded from group-level D-E gap analyses. Additionally 9 participants sampled fewer than 2 times on average per problem; these participants were excluded from the within-subjects D-E gap because they were likely either not paying attention to instructions or did not understand them. There were additionally 2 participants who sampled on average greater than 30 times per problem; these participants were excluded at the within-subjects level as they were extreme outliers in sampling.

Results

Overall Performance

Sampling Behavior

Across all problems (1-18), participants sampled a mean of 9.61 times per problem (*SD* = 6.84) and 4.80 times per alternative (*SD* = 3.42). This is slightly lower than the mean number of samples in several other studies, in which sampling typically ranges from 11 to 19 samples (see Hau, Pleskac, & Hertwig, 2010, Table 1). However, as will become evident in my discussion of the group-level D-E gap, my results are consistent with most other studies of the D-E gap. The average number of samples[vii] was higher in participants who completed DFD first relative to those who completed DFE first (D-First: *Mdn* = 11.78; E-First: *Mdn* = 5.39, $U = 498.00$, $z = 2.0$, $p < .05$, $r = -.23$). Additionally, in the risky-safe problems, participants sampled more from the risky alternative (*Mdn* = 4.50) than the safe alternative (*Mdn* = 3.79), $z = -3.25$, $p = .001$, $r = -.38$. As is expected when sampling is minimal, sampling error was pronounced. Participants only experienced the rare outcome in 47.09% of problems on average.

Group-level D-E Gap

Table 3 shows the results for each of the problems in the study. To test whether a group-level D-E gap was present (using all participants' decisions from both DFD and DFE problems), I compared the total number of choices made from each option in problems 1-12 (types A through D) for the DFD vs. DFE conditions. Results are reported in terms of proportion of choices consistent with *overweighting* rare outcomes; in other words, the proportion of choices consistent with PT predictions. This emphasizes the accuracy of PT predictions for DFD, and the deviation from PT in DFE.

Overall choice patterns across these problems were very different for the two conditions, and more specifically, indicated underweighting of rare outcomes in DFE, but overweighting of

---

[vii] Because the number of samples overall and within risky or safe alternatives was not normally distributed, tests on these values were non-parametric.

rare outcomes in DFD, resulting in a highly significant group-level D-E gap (proportion of choices consistent with overweighting rare outcomes in DFE: $M = .32$; DFD: $M = .59$; $X^2 = 140.60$, $p < .001$, $OR = 3.05$)[viii]. To explore problem-specific patterns, I also tested for the presence of a group-level D-E gap in each individual problem. A difference was obtained in the direction predicted by past literature on 11 of the 12 problems (types A-D, Table 2). Based on a chi square test, the effect was statistically significant for 9 of the problems, and approaching significance for one further problem. One problem showed no evidence of a gap (problem 11). Consistent with the results of earlier studies using problems in which both options are “risky” (outcome probabilities < 1), the group-level D-E gap did not reach significance in any of other problems (13-18, problem types E and F; although the gap in problem 13 was marginally significant).

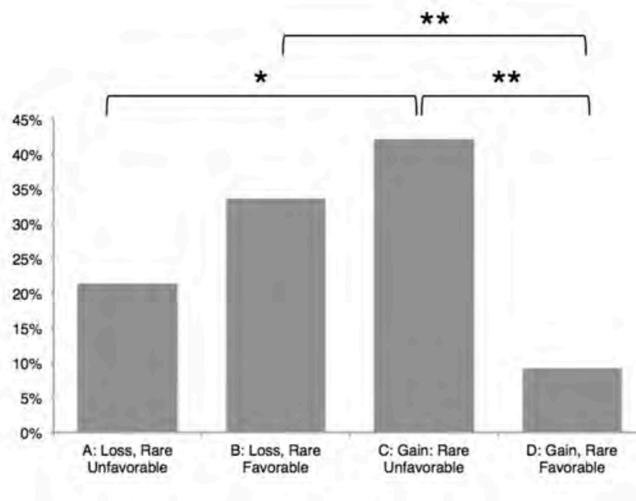


Figure 2. Group-level D-E gap by problem type. $^*p < .05$; $^{**}p < .01$

[viii] I also tested to see whether a traditional between-subjects D-E gap was significant in our study population, as the group-level D-E gap includes all participants regardless of task order. Using only the responses from each participant’s first task (i.e. DFD responses from those who completed DFD first, and DFE responses from those who completed DFE paradigm first) there was a highly significant D-E gap (proportion of choices consistent with overweighting in DFE: $M = .30$; DFD: $M = .60$; $X^2 = 90.29$, $p < .001$, OR = 3.58) in this sample, consistent with previous studies.

Table 3

D-E gap by problem

Problem Number	Problem Type	Option 1	Option 2	PT Prediction (Option consistent w/ rare overweighting)	DFD % Choosing PT Prediction	DFE % Choosing PT Prediction	Difference (D-E Gap)
1		-32, .1	-3, 1		56.25%	27.50%	28.75%***
2	A	-40, .2	-8, 1	2	52.56%	38.46%	14.10%†
3		-34, .2	-7, 1		50.62%	30.86%	19.75%*
4		-4, .8	-3, 1		64.20%	24.69%	39.51%***
5	B	-10, .9	-9, 1	1	64.20%	27.16%	37.04%***
6		-14, .8	-12, 1		69.51%	35.37%	34.15%***
7		4, .8	3, 1		72.50%	27.50%	45.00%***
8	C	10, .9	9, 1	2	76.25%	47.50%	28.75%***
9		16, .8	13, 1		79.27%	28.05%	51.22%***
10		32, .1	3, 1		43.21%	28.40%	14.81%*
11	D	35, .2	7, 1	1	35.44%	35.44%	0.00%
12		29, .2	6,1		38.46%	30.77%	7.69%
13		5, .2	3, .25		72.15%	58.23%	13.92%†
14	E	15, .4	10, .6	N/A (choices listed are for option 1)	34.18%	37.97%	3.80%
15		28, .25	9, .8		45.57%	34.18%	11.39%
16		-6, .2	-4, .25		45.68%	38.27%	7.41%
17	F	-3, .8	-12, .2	N/A (choices listed are for option 1)	54.43%	44.30%	10.13%
18		-12, .5	-16, .4		60.49%	59.26%	1.23%

Note. *$p < .05$; **$p < .01$, ***$p < .001$; †$p < .1$

Problem Parameters

To gain further traction in understanding the factors that drive the group-level D-E gap I examined how the magnitude of the gap varied by problem type. *Table 4* shows the percent of choices for Option 1 in each problem type. *Figure 2* displays the size of the gap (the difference between underweighting decisions in DFD compared to DFE) in each problem type. An omnibus ANOVA entering valence (gain, loss) and rare-outcome favorability (rare favorable, rare unfavorable) showed that there were significant differences in the group-level D-E gap by problem type ($F(3, 11) = 11.63$, $p < .003$, $\eta^2 = .81$). More specifically, there was a significant valence by rare-outcome favorability interaction ($F(1,11) = 30.07$, $p < .001$, $\eta^2 = .79$), such that in loss problems (problem types A and B), the gap was largest when the rare outcome was favorable,

but in gain problems (types C and D), the gap was largest when the rare outcome was unfavorable.

Post-hoc Tukey HSD tests showed that the gap was significantly larger in both problem types B (M = .37) and C (M = .42) compared to problem type D (M = .08; $p < .008$ for B vs. D; $p < .003$ for C vs. D). The gap in problem type C was also larger than that in type A (M = .21, $p < .05$). None of the other pairwise comparisons were significant (ps > .1). To summarize, the size of the gap in type A falls in between that of problem types B/C (large gap) and type D (virtually no gap). This pattern of results is consistent with Camilleri and Newell's (2013b) finding that the D-E gap is more pronounced in loss problems for which the rare outcome is desirable (problem type B, where the rare outcome is 0), and in gain problems for which the rare outcome is undesirable (problem type C, where the rare outcome was also 0 but in this case is unfavorable compared to a gain).

Table 4

Group-level D-E gap by problem type

Problem Type	Problem Description	DFD % Overweighting Rare Outcomes	DFE % Overweighting Rare Outcomes	Difference (D-E gap)
A	Loss Rare Outcome Unfavorable	54.55%	33.06%	21.49%
B	Loss Rare Outcome Favorable	68.03%	34.43%	33.61%
C	Gain Rare Outcome Unfavorable	72.73%	30.58%	42.15%
D	Gain Rare Outcome Favorable	69.49%	60.17%	9.32%

***$p < .001$

Additionally, I examined whether expected value (EV) of the risky compared to the safe option influenced the size of the D-E gap. A chi-square test based on the frequency of decisions consistent with overweighting rare probabilities in DFE compared to DFD when the risky option was higher, lower, or the same EV as the safe option did not reach significance, $X^2(2) = 0.07$, $p = .79$, $V = .01$. Hence, the evidence suggests that relative EV does not influence the magnitude or direction of the D-E gap.

Task Order

When the sample was divided according to the order in which the DFD and DFE tasks were completed, participants who started with DFD problems (*DFD-First*, $N = 44$) and those who started with DFE problems (DF*E-First*, $N = 46$) were found to exhibit similar choice patterns. There was a significant group-level D-E gap in problems types A-D for both the DFD-First and DFE-First groups, (DFD-First: $X^2 = 64.63$, $p < .001$, $V = .26$; DFE-First: $X^2 = 75.60$, $p < .001$, $V = .28$). Further, a *t*-test on the size of the gap across these 12 problems showed that the gap did not differ between DFD-First ($M = .26$) compared to the DFE-first group ($M = .27$; $t(11) = .32$, $p = .76$, $r^2 = .01$).

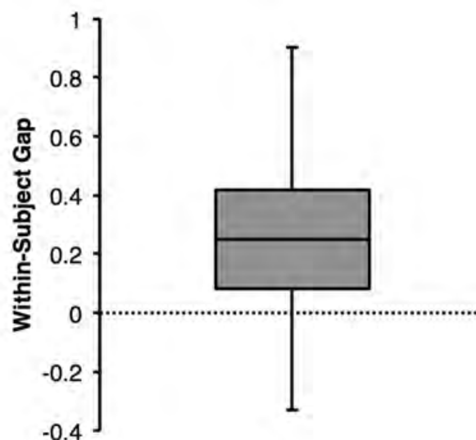


Figure 3. The within-subjects D-E gap measure.

The Within-Subjects D-E Gap

Figure 3 displays a boxplot characterizing the within-subjects D-E gap variable across the entire sample. The gap ranged from -.33 to .9, with a mean of .25 (SD = .24). A one-sample *t*-test indicated that, as expected based on the systematic biases described in the literature, the average D-E gap was significantly higher than 0, $t(73) = 9.02$, $p < .001$, $r^2 = .51$. As mentioned when I detailed the calculation of this D-E gap measure, the variable only accounts for directionality of switching, not for total number of switches. However, the within-subjects D-E gap is significantly correlated with the total number of preference reversals (in either direction) in the core problems, $r = .41$, $p < .001$. Interestingly, the D-E gap variable was not significantly correlated with the number of preference reversals in problem types E and F (for which I had no specific predictions about the direction of participants selections; $r = -.05$, $p = .65$), suggesting that the variable is capturing something unique to problems with one risky and one safe alternative. Additionally, the within-subjects D-E gap variable did not significantly vary by task order, $t(72) = 0.29$, $p = .77$, $r^2 = .001$, but was marginally significantly correlated with the number of samples taken ($r = -.20$, $p < .09$), such that those who took more samples had a smaller D-E gap.

D-E Gap Variable and Behavioral Impulsivity

A next step was to examine individual variability in the size of the within-subjects D-E gap, and to explore the relationship of this variable to other indices. *Figure 3* provides a visual representation of all participants' preference reversals, showing clear variability in the number of preference reversals exhibited across participants.

I next sought to evaluate whether this variability across participants can be explained by other experimental measures. In particular, all participants completed Go/No-Go, which served not only as a filler task, but also as a measure of impulsivity; assessed as false alarm rate in the task (i.e., failure to inhibit a motor response). Indeed, false alarm rate was positively correlated with the D-E gap, $r = .29$, $p < .02$, as displayed in *Figure 4.*

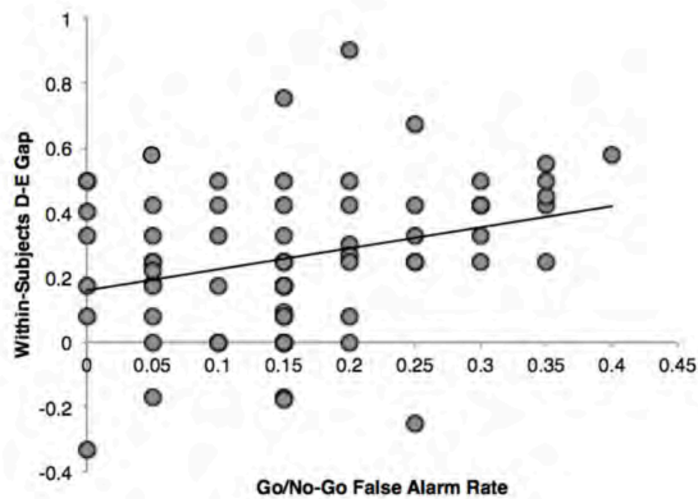


Figure 4. Correlation between within-subjects D-E Gap measure and false alarm rate during Go/No-Go.

Modeling Sampling Error at the Within-Subjects Level

Some have suggested that sampling bias can fully explain DFE deviation from PT (e.g., Fox & Hadar, 2006; Hadar & Fox, 2009). As mentioned above, there is some evidence that sampling is related to this measure. It is therefore possible that more sampling leads to a more accurate sampling of the distribution, and that with accurate information about the distribution, participants decide in accordance with PT in DFE. To examine this possibility, I computed new PT predictions (Venkatraman, Payne, Bettman, Luce, & Huettel, 2009) based on the outcomes participants actually observed in DFE. (Note that in DFE, participants should decide *counter* to PT, in that they display underweighting, rather than overweighting, of rare events.) Next, I divided the sample into those with a high D-E gap, (.25 and above, *N* = 35) and those with a low D-E gap (below .25, *N* = 25), and ran an ANOVA entering D-E gap size, (low, high) and task (DFE, DFD) as indipendent variables, and the proportion of modeled PT-consistent choices as a dependent variable, controlling for number of samples taken in DFE[ix]. As is displayed in *Figure 5,* there was

[ix] Note that only participants who observed all outcomes at least once could be included in this analysis.

a significant D-E gap size by task interaction ($F(1,57) = 7.52$, $p = .008$), such that even when controlling for sampling bias, PT does not adequately account for experience-based decisions of those with a high D-E gap. In contrast, the D-E gap is fully explained by sampling bias in those with a low D-E gap (i.e., PT explains DFD and DFE similarly well in these participants). *Figure 5* displays the estimated marginal means for decisions consistent with PT predictions by gap size (low or high).

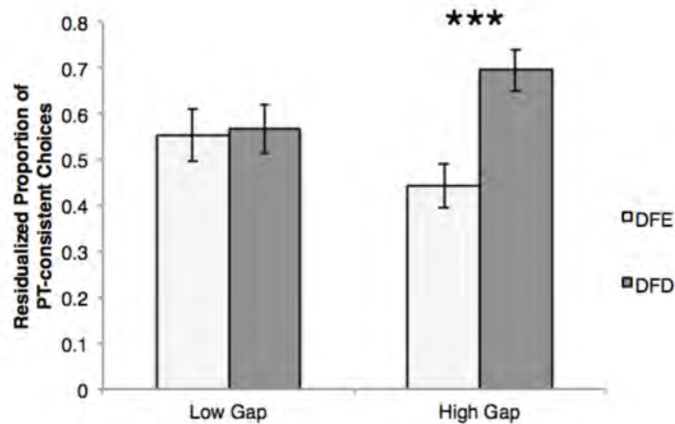


Figure 5. Choices consistent with PT predictions after controlling for sampling bias in DFE. The analysis also controls for average number of samples taken in DFE. ***$p < .001$

Discussion

Judgment and decision making research in the past decade has begun to focus on the way in which participants learn about a decision problem, and how the source of information influences decision preferences (Camilleri & Newell, 2013a; Hertwig et al., 2004). However, most such experiments have used between-subjects designs (c.f. Camilleri & Newell, 2009), which cannot adequately characterize how a given individual responds to changes in the mode of presentation.

In the present study, I replicate earlier work showing that choice patterns depend on the mode of presentation (DFD vs. DFE), and extend those results by showing that this group-wise pattern derives from a within-subjects paradigm. While the gap was slightly larger when DFD was

presented before DFE, both task orders produced a significant group-level D-E gap. Additionally, the group-level D-E gap in the present within-subjects paradigm was most reliable in gain problems with a rare unfavorable outcome (when the rare outcome was 0) and loss problems with a rare favorable outcome (also when the rare outcome was 0, but 0 is favorable relative to a loss). This result provides further corroboration for the validity of the proposed paradigm, and shows consistency between Camilleri and Newell's (2013b) full-feedback repeated choice paradigm and the present free sampling paradigm. This result also supports the idea that, when sampling is free, slightly different experience-based paradigms produce similar choice biases and may rely on similar cognitive processes (Camilleri & Newell, 2011b; Gonzalez & Dutt, 2011).

The present study also contributes a potentially valuable tool to the literature on the D-E gap, by specifying the derivation of an individual subject D-E gap index that incorporates how a given participant responds to both described and experienced information. As expected, this within-subjects D-E gap was significant in the direction of underweighting rare outcomes in DFE, but overweighting them in DFD, regardless of whether participants completed DFD or DFE first. Mean number of samples was marginally negatively correlated with the D-E gap variable, suggesting that sampling bias may play a role in the D-E gap, but sampling bias could only explain the D-E gap for those individuals with a low D-E gap. For those with a high D-E gap, participants still chose as if they underweighted rare events when they observed them during sampling, while at the same time, deciding in accordance with PT in DFD. Crucially, this result suggests that learning about probabilities from experience may involve mechanisms that are inherently different than those involved in decisions based on descriptions (at least for individuals with a high D-E gap). Further, motor impulsivity was significantly positively correlated with the D-E gap, suggesting that the two constructs may share mechanisms, or that heightened impulsivity may itself be a mechanism of the D-E gap.

While there are many exciting aspects to this study, it is important to point out some limitations. Foremost, since the sampling paradigm did not require a minimum number of samples

from each option in the DFE condition, many participants did not experience the rare outcome, and hence, could not have formed an accurate representation of the underlying probability distributions (i.e., there was substantial sampling error). However, the aim of this study was to evince as much inter-individual variability as possible so that this variability could be used to evaluate the meaningfulness of the individual gap. Indeed, if given the opportunity to sample, a person who is willing to make a decision based only on one sample likely possesses different traits than does a person who pursues more information before rendering a decision. This paradigm is also analogous to experiences in daily life – in many risk situations, it is unlikely for a person to experience the rare, but costly, outcome (although they might have an intuition about its potential to occur). Thus, how people decide in the absence of such experience, even when they know a rare outcome might be possible, is important in applying this work to real-world decisions. However, future studies exploring the properties of the within-subjects D-E gap obtained from a paradigm requiring participants to take a specified number of samples from each option, or when the outcomes of sampling are experimentally fixed to optimally represent the true underlying probability distributions, could be informative.

Importantly, the work in Experiment 1 provides a crucial building block for understanding risk taking as a function of information source, and therefore, for testing the Spectrum of Affect model. Experiment 2 will use the within-subjects D-E paradigm in adolescents and adults to index DFD and DFE as they relate to affective arousal.

CHAPTER 3

EXPERMENT 2: TESTING THE D-E GAP IN ADOLESCENTS AND ADULTS

Experiment 2 employed the D-E gap paradigm presented in Experiment 1 in samples of adolescents and adults as a test of the Spectrum of Affect model. In using a version of the traditional D-E gap paradigm in adolescents and adults, Experiment 2 was similar to a new study by van den Bos and Hertwig (2017), and the present results can serve as a replication and extension of that study (although there are some differences in the tasks and results, which I address in the General Discussion). One important extension in Experiment 2 involves my use of physiological methods to understand processes underlying decision behavior by age group. Beyond behavioral and survey measures of risk taking, I employed eye tracking methods during Decisions from Description (DFD), and direct measures of physiological arousal throughout the tasks (skin conductance and heart rate variability, although results focus on heart rate variability). Such process-tracing methods have become increasingly popular as researchers acknowledge that there is more to decision making than the input (a decision problem or task) and output (the choice; Payne & Venkatraman, 2011; Weber & Johnson, 2009). Specifically, eye tracking is an important method in JDM research, as it can provide insights into the particular information that is sought in decision making. Eye tracking research has shown that participants do not allocate their attention to outcomes according to their contingent probabilities, as one might hypothesize. Rather, participants focus *more* on rare or extreme outcomes in DFD, potentially biasing them towards overweighting those outcomes in choice (Glöckner, Fiedler, Hochman, Ayal, & Hilbig, 2012; Weber & Johnson, 2009). Importantly, eye tracking can help to address situations in which two individuals or groups make the same choices, but due to different psychological processes. Therefore, eye tracking is an exciting way to address age differences in risk taking in the laboratory, particularly in DFD where behavioral age differences in risk taking are often not observed (Rosenbaum et al., Resubmitted). It is possible that eye tracking patterns will diverge by age group, even in the absence of behavioral differences between age groups.

Process-tracing methods can also aid in testing the Spectrum of Affect model, which assumes that heightened affective processes motivate teens' risk taking (and additionally, the hypothesis that risk taking is limited to choice contexts with rare unfavorable outcomes, accompanied by heightened arousal). Indeed, the JDM literature has seen an increasing trend in studying affective processes as they relate to risky choice (Slovic, Finucane, Peters, & MacGregor, 2007; Weber & Johnson, 2009). As discussed in Chapter 1, the adolescent risk taking literature also relies on affective explanations for adolescent risk taking (e.g., that affective structures like the ventral striatum are hyperactive in adolescent vs. adult decision making, leading to increased risk taking; e.g., Shulman, Smith et al., 2015). However, affective arousal is rarely, if ever, directly measured in studies of adolescent risk taking using physiological methods that traditionally index affect, like skin conductance, heart rate or pupil dilation (Payne & Venkatraman, 2011). Instead, most of the developmental cognitive neuroscience literature uses fMRI, which can also be used to infer affective arousal (Knutson, Katovich, & Suri, 2014), but is more expensive and less direct than traditional measures.

In the present study, I tested the within-subjects D-E gap in cross-sectional samples of adolescents and adults, and concurrently employed eye tracking and arousal measures to better understand decision processing by age. More specifically, Experiment 2 was designed to address the following specific aims:

Aims and Hypotheses

Primary Aim 1:
Age Differences in Information Processing and Decision Making as a Function of Information Source

The within-subjects D-E gap paradigm from Experiment 1 was used to test the hypotheses that adolescents show similar risk taking to adults in DFD (*Hypothesis 1.a.*), but take more risks in DFE (*Hypothesis 1.b.*), stemming from the Spectrum of Affect model and the developmental risk taking literature.

I additionally tested the hypothesis that most adolescents do not yet exhibit PT-consistent weighting preferences in DFD (*Hypothesis 1.c.*), based on prior studies estimating immature weighting functions in adolescents relative to adults in DFD (Engelmann et al., 2012; Harbaugh et al., 2002; Tymula et al., 2012). Moreover, I propose that adolescents will show different eye gaze patterns in DFD than adults when rendering their decisions, perhaps revealing an uneven distribution of gaze time across probabilities and values (*Hypothesis 1.d.*).

I expect a different pattern of results in DFE with respect to outcome weighting and development. When information about a risk is acquired through experience, I predict both adolescents and adults will make choices consistent with underweighting of rare outcomes, but that this rare-underweighting tendency will be exaggerated in adolescents (*Hypothesis 1.e.*). If adolescents show less PT-consistent behavior in DFD, but higher rare-outcome underweighting than adults in DFE, then the overall size of the D-E gap in adolescents may be either similar to or slightly smaller than that in adults (*Hypothesis 1.f.*).[x]

I additionally predict that age differences in DFE choices will be accompanied by a tendency for adolescents to take fewer samples than adults (*Hypothesis 1.g.*), due in part to heightened impulsivity (see Aim 3 for more on this idea). I additionally hypothesize that age differences are *not* driven by an inability to properly remember and estimate probabilities from sampling in adolescents (*Hypothesis 1.h.*), and attempt to rule out this possibility using a novel sampling and estimation paradigm analogous to DFE where participants sample shapes from jars and report their estimation of the contents of each jar.

Primary Aim 2:

Impact of Affective Arousal on Decisions from Description and Experience

A key component of the Spectrum of Affect model is that differences in physiological arousal will accompany age differences in decision-making behavior. To investigate this

---

[x] Note that this hypothesis differs from van den Bos and Hertwig's (2017) study, which found a higher D-E gap in adolescents relative to adults. I will return to this point in the discussion.

possibility, I collected physiological data during task completion. I hypothesize that adolescents will show signs of heightened physiological arousal relative to adults in DFE (*Hypothesis 2.a.*) due to learning from feedback during sampling, but that arousal levels will be similar in DFD (*Hypothesis 2.b.*).

I also attempted to manipulate the degree of affect in the problems themselves by varying the types of decision problems presented to participants. In addition to testing the traditional set of single-domain decision problems (i.e., the gain-only and loss-only problems used in Experiment 1; Rosenbaum, Chein, & Venkatraman, In Revision) participants also completed a new set of mixed gain-loss problems (Brooks & Zank, 2005; Lichtenstein, 1965; Loomes, 2010; Venkatraman, Huettel, Chuah, Payne, & Chee, 2011; Venkatraman et al., 2009) thought to increase affective arousal (Hochman & Yechiam, 2010; Yechiam & Telpaz, 2011; although whether this manipulation heightens arousal in an adolescent sample also remains untested).

Specific hypotheses based on the problem type manipulation include: *2.c.* adolescents will show the highest degree of underweighting rare outcomes in mixed-domain (relative to single-domain) problems during DFE; *2.d.* both adolescents and adults will show higher physiological arousal during mixed-domain blocks relative to single-domain blocks of DFE and DFD; and *2.e.* physiological arousal will be highest in adolescents within the DFE mixed-domain block.

Exploratory Aim: Individual Differences in Risk Taking as a function of Age Group

The degree of risk taking in adolescence varies greatly by individual. The description- and experience-based decision making paradigm that I developed for use in this study allows calculation of each participant's D-E gap (as described in Chapter 2). I will correlate the within-subjects D-E gap measure with responses to a motor impulsivity task (Go/No-Go), an impulsive choice task (Delay Discounting) and self-report questionnaires. I will additionally explore whether this measure is related to task-based impulsivity, decision making styles, or self-reported real-world risk taking propensities, across development. Based on results from Experiment 1, I hypothesize that false alarm rate in Go/No-Go will correlate positively with the D-E gap measure

(*Hypothesis 3*). While there is reason to believe that motor impulsivity may correlate with the D-E gap measure, it is unclear whether impulsive choice, as indexed by the Delay Discounting task will also correlate. Therefore, this analysis, along with the analyses described next, is exploratory.

I will run three sets of correlations between self-report questionnaires and the D-E gap measure by age group. The first set of correlations will include widely-used measures from the Judgment and Decision Making (JDM) field (Generalized Decision Making Scale, Scott & Bruce, 1995; Maximization-Regret scales, Schwartz et al., 2002; Rational-Experiential Inventory Epstein, Pacini, Denes-raj, & Heier, 1996). The second set of correlations will include measures often used in developmental literature to understand risk taking mechanisms in adolescents (Sensation seeking scale, Zuckerman, Eysenck, & Eysenck, 1978; Barratt Impulsivity scale, Patton, Stanford, & Barratt, 1995a; Zuckerman Disinhibition Scale, Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993; Future Orientation scale, Steinberg et al., 2009). Last, I will assess the relationship between the D-E gap and self-reports of real-world risk taking as indexed by an adaptation of the Benthin Risk Perception Scale (Benthin, Slovic, & Severson, 1993; Steinberg et al., 2009). Because these analyses are underpowered, I do not make any direct hypotheses about possible results and do not expect to find large effects. These analyses are run mainly to aid in future hypothesis generation. With that said, it is possible that D-E gap behavior (i.e., underweighting rare events in DFE and overweighting them in DFD) will be positively correlated with self-reported measures of impulsivity, faith in intuition and real-world risk taking, while the gap will be negatively correlated with maximization and regret scores.

Methods

Participants

Thirty-one adolescents (ages 13-17, *M* = 15.33, *SD* = 1.30, 21 female) and 33 adults (ages 25-36, *M* = 29.39, *SD* = 4.02, 15 female) participated. Adolescents were recruited largely through outreach efforts at Philadelphia middle and high schools. Several adolescents and the entire adult sample were recruited via flyers posted at community centers, coffee shops, and

around Temple University, Craigslist ads, and word of mouth. Attempts were made to recruit demographically similar samples of adolescents and adults; the similarities and differences of the final samples are discussed further below.

One adolescent and 3 adults were excluded from all analyses presented below, as they did not respond correctly to comprehension questions about task instructions (listed in the Appendix). Therefore, the final sample included 30 adolescents and 30 adults.

Behavioral Tasks

D-E Gap Paradigm

Age-differences in the D-E gap were assessed using a paradigm I developed, which was described in the methods section of Experiment 1 (Rosenbaum et al., In Revision) with a few modifications. This paradigm indexes within-subject differences in the same decisions made on the basis of description and experience. In this paradigm, participants are presented with a series of problems requiring a choice between two alternative "point-machines". In counterbalanced blocks, decisions are based on explicit *descriptive* information regarding the values and outcome probabilities assigned to each machine (*Figure 6.A.*), or on prior *experience* obtained by sampling outcomes from those machines in the absence of any prior or overt information about the internal values or probabilities (*Figure 6.B.*). Each of the problems used in the task is presented in each condition (description [DFD], experience [DFE]), in pseudo-random order. In the present study, problems always included one risky option (a machine that produces a likely outcome paired with a rare outcome) and one sure option (a machine with a single fixed outcome). The experiment in Chapter 2 included six problems in which both alternatives were risky. These risk-risk problems were excluded in the present experiment because they did not evince a reliable D-E gap.

Further, while most studies of the D-E gap (including that detailed in Chapter 2) use problems with outcomes limited to either gains or losses, I introduced 12 mixed problems, which included both gain and loss outcomes within a single problem (Brooks & Zank, 2005; Lichtenstein, 1965; Loomes, 2010; Venkatraman et al., 2011, 2009), as mixed problems have

been shown to evince heightened physiological arousal in adults (Hochman & Yechiam, 2010; Yechiam & Telpaz, 2011). In order to collect the cleanest physiological data possible, I presented all 12 of the single-domain problems (gain or loss only; problem types A-D in Chapter 2) in a block separately from the 12 mixed-domain problems in both DFD and DFE conditions. The block order (single-domain first, mixed-domain first) was counterbalanced across participants.

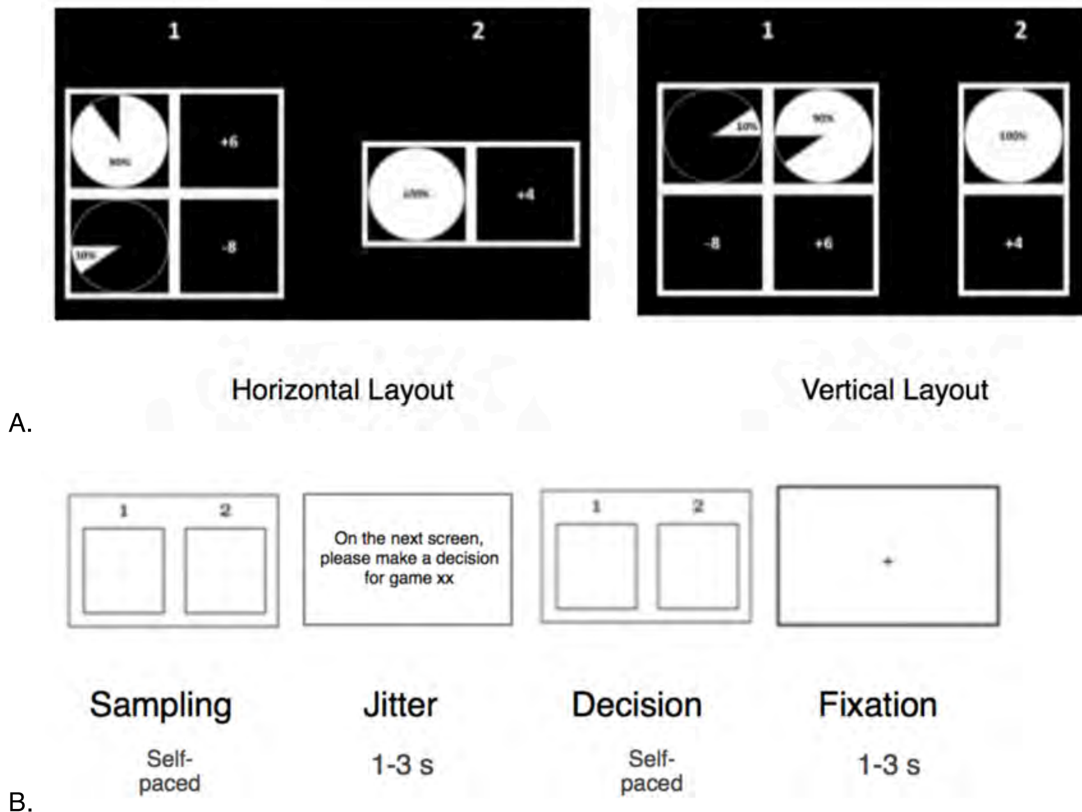


Figure 6. Illustration of the D-E gap paradigm. A. Decisions from description trials. Horizontal and vertical layouts were randomized for eye tracking. The side of the screen on which the risky and safe options appeared, and the position of the rare and common outcomes within the risky grid (top/bottom in horizontal layout, left/right in vertical layout) were also randomized. After each decision, a 1-3s fixation screen was presented. If decision period is slower than 8 s, extra time is added to 1-3 s fixation. B. Decisions from experience trials.

Based on prior D-E gap studies (Camilleri & Newell, 2013a; de Palma et al., 2014; Hertwig et al., 2004), an *overweighting* of rare outcomes is expected in the DFD condition and an *underweighting* of rare outcomes is expected in the DFE condition. An individual's D-E gap can

be calculated by identifying the problems for which the participant switches his or her choice in the expected direction across the DFD and DFE conditions (Rosenbaum et al., In Revision). I created a *single-domain gap* and a *mixed-domain gap*, as well as an *overall gap* including both single- and mixed-domain problems. In addition, as in Experiment 1, I modeled prospect theory (PT) predictions for the participants' experienced outcomes to account for sampling error.

Impulsive Action (Go/No-Go)

As in Experiment 1, participants completed a filler Go/No-Go task between DFD and DFE blocks. The task used pictures of cars as stimuli (e.g., "Go" for SUVs, "No-Go" for 4-door sedans), and each stimulus was presented for 500ms. Seventy percent of trials were "Go" trials and 30% of trials were "No-Go" trials. Once again, false alarm rate (i.e., rate of No-Go errors) was computed a measure of motor impulsivity.

Impulsive Choice (Delay Discounting)

Participants completed a delay discounting task (O'Brien, Albert, Chein, & Steinberg, 2011; Weigard, Chein, Albert, Smith, & Steinberg, 2014) in which they made hypothetical choices between receiving a given amount now and $1000 at some time delay. The first immediate amount offered was $200, $500, or $1000 (chosen randomly). The immediate reward was then adjusted based on the participant's response. For example, if the participant chose the $1000 reward at a delay, the next offer was mid-way between the prior immediate reward and $1000, and so on, until an indifference point was determined. Delay periods included one day, one week, one month, six months, and one year. Mean indifference point across these time periods was computed as a measure of impulsive choice.

Shapes Estimation Task

To better understand the mechanisms of DFE sampling, participants completed a shapes estimation task, in which they drew shapes (circles, stars, triangles, diamonds) from two bins and used a slider to report their estimation of the distributions within each bin. On each trial, participants were presented with two unmarked bins, each containing shapes, and sampled freely

until they felt confident reporting the relative distributions of shapes in each bin. The shape distributions in the bins paralleled those in the DFE point machines: one of the bins only contained one shape (analogous to safe point machines) and one bin contained a rare shape and a common shape (analogous to risky point machines). Participants each completed 4 trials of this task: 2 where the rare shape comprised 20% of the distribution and 2 with a rare shape that comprised 10% of the distribution, in random order.

Self-Report Measures

In addition to completing specific behavioral tasks, participants also provided demographic information and completed a battery of self-report survey measures. These included an adaptation of the Benthin Risk Perception Scale including risk perception and risk taking measures related to 9 problem behaviors (Benthin et al., 1993; Steinberg et al., 2009), the Generalized Decision Making Styles Inventory (Scott & Bruce, 1995), the Maximizing and Regret scales (Schwartz et al., 2002), the sensation seeking scale (Zuckerman et al., 1978), two measures of impulsivity (Patton et al., 1995; Zuckerman et al., 1993), and the Future Orientation scale (Steinberg et al., 2009).

Physiological Measures

Eye Tracking

In order to better understand decision processes in DFD and whether they vary between single- and mixed-domain blocks across development, a Tobii T60XL eye-tracking system was used to acquire information about eye gaze. The environment (lighting, sound) was controlled across participants, and stimulus positioning (the position of the risky vs. safe option on the left vs. right; in DFD, the position of the point values and probabilities on the left vs. right and top vs. bottom) was randomized to account for biases in gaze patterns[xi].

[xi] Due to a coding error in the first 8 adolescent participants and the first adult participant, all vertical displays showed the common value and probability on the left, and horizontal displays always showed the common value on the top. However, gaze data did not differ between these

Biometric Data

Participants were affixed with BIOPAC (MP150) BioNomadix wireless physiology devices for collecting skin conductance (or electrodermal activity; EDA) and heart rate data. These measures were then used to index the autonomic processes and affective arousal evoked by task manipulations. Measures were sampled at 1 kHz. I used BIOPAC disposable dry electrodes (EL509) with BIOPAC isotonic 0.05 molar NaCl gel (GEL101) to collect skin conductance data. For heart rate data, I used Carolina Exercise ECG Electrodes with BIOPAC 0.85 molar NaCl gel (GEL 100).

Procedure

Figure 7 displays a schematic of the study design. Upon arrival at the lab, consent was obtained from participating adults, and parental consent and child assent was obtained for all adolescent participants.

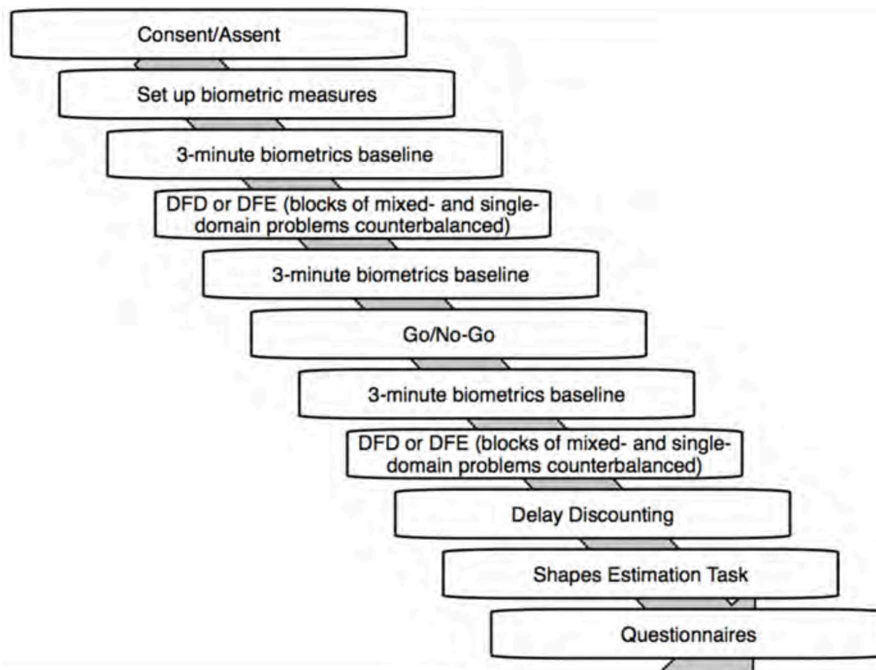


Figure 7. A schematic of the study design for Experiment 2.

---

participants and the rest of the sample after the error was corrected, so all data were included in eye tracking analyses.

Biometrics Setup

Figure 8 shows pictures of biometrics setup. Caffeine intake is thought to affect HRV measures (Boucsein et al., 2012). I therefore asked participants to self-report caffeine intake for the day of the study.

Heart Rate (Electrocardiogram; ECG). Three electrodes were affixed to the participant: One connected to a positive lead just below the left ribcage, one connected to a negative lead above the participant's right collarbone, and one connected to a ground lead above the participant's left collarbone (see *Figure 8.A.*). Participants wore a belt around their waste with a wireless transmitter that was connected to the leads.

Skin conductance (Electrodermal activity; EDA). Participants were asked to wash their hands with pH-balanced soap. Subsequently, participants indicated which hand they preferred to use for task responses, and electrodes were placed on the distal phalanges of the index and middle fingers on the opposite hand (see *Figure 8.B.*). The leads were connected to a transmitter strapped to the participant's wrist. Participants were asked to place their arm on an arm rest and to try not to move their arm or hand.

Eye Tracking

Participants completed all tasks on the TOBII monitor, allowing eye tracking data to be collected during DFD. E-Prime interfaces with the TOBII Eye Tracker to calibrate the eye tracker and collect data during the tasks.

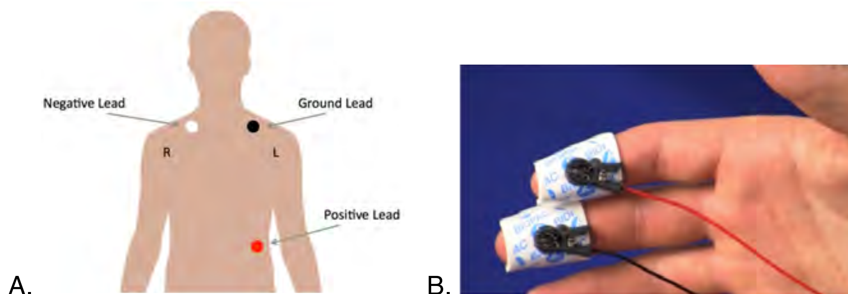


Figure 8. Biometrics setup. Placement of electrodes for A. ECG and B. EDA data collection.

Biometrics baseline

I collected three, three-minute biometrics baseline measurements, before and after the first decision-making task (DFD or DFE), and before the second-decision making task (*Figure 7*). During these periods, participants were instructed to sit quietly while looking at a gray fixation cross on a black background.

Tasks

DFD and DFE were administered in a counterbalanced fashion, with the round of go/no-go task performance occurring in between the two, and delay discounting and the shapes estimation task always occurring after both decision tasks. The experimenter verbally instructed participants, and participants were asked if they had any questions before beginning each experiment. Additionally, participants were asked a series of questions about the instructions (listed in the Appendix) at the end of the task period to ensure understanding. After the tasks, EDA and ECG electrodes were removed.

Questionnaires

Participants completed the questionnaires described above via computer, and interfaced to surveymonkey.com. The experimenter left the room during this period to ensure participants' privacy.

Data Analysis Methods and Considerations

Before parametric tests were performed on behavioral and physiological data, I checked for and removed outliers, defined as any value greater than 3 times the interquartile range from the age group mean, for any measure.

Behavioral Data

As in Chapter 2, I computed a *group-level* and *within-subjects* D-E gap within adolescent and adult groups using Excel and SPSS (Version 24.0). To more thoroughly explore behavioral patterns across adolescent and adult groups in DFD and DFE, I ran generalized linear mixed-effects logistic regressions using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R

(Version 3.3.2; R Core Team, 2016). In behavioral analyses, individual trials were excluded if the participant did not sample at least once from each alternative in DFE (only 1.11% of all trials; maximum of 4/24 trials per participant, average of 0.26 trials per participant).

Demographic and Self-Report Data

Gender. One concern in data collection was making the adolescent and adult groups as similar as possible. Despite attempts to recruit even samples of males and females, the final adolescent group included 21 females and only 9 males (1 additional male was excluded in all analyses), while the final adult group had 15 females and 15 males (3 additional males were excluded). The difference in gender distribution did not reach significance based on a chi-square test ($X^2 = 2.50$, $p = .11$, $V = .20$), and our D-E gap within-subjects measure did not differ by gender within adolescents or adults (ps > .95), so we did not consider gender further.

Socioeconomic status (SES). Because experimental tasks required participants to make monetary decisions, overall SES of the age groups was an important consideration. In the demographic questionnaire, participants were asked to report each of their parents' highest educational attainment as a measure of SES (coded as: 1 = Did not finish high school, 2 = High school diploma/GED, 3 = Some college, 4 = 2-year college degree, 5 = 4-year degree, 6 = Graduate degree). Nine adolescents either indicated they did not know their parents' education levels or did not wish to answer the questions. One adult did not answer the maternal education question, and two adults did not provide paternal education. In the existing data, a *t*-test did not show a difference between parental education levels in adolescent and adult groups (maternal: $t(48) = 1.21$, $p = .23$; paternal: $t(44) = 1.10$, $p = .28$). A different way to quantify SES is to compare adolescents' parents' education to the adult groups' own educational attainment, as it is unlikely that adults ages 25-36 are still limited by their parents' SES. This analysis revealed that the adult group had significantly higher education than the mean of adolescents' parental education $t(50) = 3.04$, $p = .004$. To investigate whether SES was related to task behavior, I ran correlations between the within-subjects D-E gap measure (the main behavioral measure of interest) and the SES measure. This correlation

did not reach significance within either the adolescent or adult group (*p*s > .2), so I did not consider SES further.

Ethnicity and Race. Ethnic composition of the adolescent and adult samples was quite similar (Adolescents: 1 Hispanic, 26 Non-Hispanic, 3 did not answer; Adults: 3 Hispanic, 26 Non-Hispanic, 1 did not answer). However, the two age groups significantly differed in race (see *Table 5*; $X^2 = 15.89$, $p = .003$, $V = .52$). To investigate the possibility that race impacted behavioral results, I ran an ANOVA investigating the size of the D-E gap measure (across all decision problems) by race using the three most common races (African-American, Asian and Caucasian). In adolescents, the ANOVA did not reach significance ($F < 1$). In adults the ANOVA was marginally significant ($F(2,26) = 2.72$, $p = .09$), with Asians exhibiting a lower gap ($M = .04$) than African Americans ($M = .26$) and Caucasians ($M = .28$). Therefore, I cannot entirely rule out the possibility that the difference in racial composition of the samples influenced the results. Whether the D-E gap systematically differs by race should be tested in future research.

Table 5

Racial composition of adolescent and adult samples

	African-American	Asian	Caucasian	Pacific-Islander	Other
Adolescents	11	14	4	0	1
Adults	5	5	17	1	2

Eye Tracking in DFD

Eye tracking data could not be collected for one adolescent and one adult who each needed to wear thick glasses throughout the experiment.

Data were analyzed using in-house scripts in MATLAB, and then transferred to SPSS for further analysis. For three adolescents and one adult, fewer than 50% of all eye tracking data acquisitions were valid, so they were excluded from all analyses. One further adolescent had enough valid fixations to be included in analyses of mixed-domain problems, but validity decreased below 50% for single-domain trials due to excessive movement. This participant was

only excluded from single-domain analyses. Overall, the difference in validity in adolescents (M = 74.07%, SD = 25.47%) and adults (M = 85.21%, SD = 17.08%) was marginally significant ($t(56)$ = 1.95, p = .06). After removing participants with validity under 50%, there was not significant evidence of an age difference in validity ($M_{adolescents}$ = 83.20%, SD = 11.50%; M_{adults} = 87.89%, SD = 9.25%; $t(51)$ = 1.67, p = .10, r^2 = .05).

Eye tracking data (in the form of proportions of valid looking time per decision trial) were positively skewed and were therefore square-root transformed. One adolescent participant was an outlier even after transformation and was excluded from analyses. Analyses focused on the proportion of gaze time allotted to probabilities and values DFD across age and domain.

Biometrics

Biometrics data were lost for one adolescent participant due to experimenter error. For an adult participant, biometrics data could not be collected due to health issues. Any missing data within individual measures is mentioned below.

Heart rate variability (HRV). Data were exported to Kubios HRV Version 2.2 (Tarvainen, Niskanen, Lipponen, Ranta-aho, & Karjalainen, 2013), run in MATLAB (Version R2012a). Kubios was designed to follow guidelines recommended by the Task Force of The European Society of Cardiology and the North American Society for Pacing and Electrophysiology (1996) for HRV analysis. I first manually corrected all heart rate time series using Kubios. In some participants on a small number of trials, there were artifacts that could not be manually corrected (on average, less than 0.2% of total beats per participant, and most of the artifacts did not occur during task or baseline blocks). For these artifacts, I applied a cubic spline interpolation at 4 Hz.

Frequency-based analyses are often used to understand autonomic activity from ECG signal, and a power spectrum based on ECG signal can be divided into very low frequency (VLF, .003-.04 Hz), low frequency (LF, .04-.15 Hz), and high frequency (HF, .15-.4 Hz) bands (Tarvainen, 2014; Task Force, 1996). My analyses filter out VLF rhythms using a detrending method with smoothness priors (Tarvainen, Ranta-aho, & Karjalainen, 2002), as VLF rhythms are not thought to

be relevant to sympathetic arousal and can distort analyses of LF and HF rhythms. Power in HF bands is commonly thought to reflect parasympathetic nervous system activity (e.g., respiratory sinus arrhythmia; Berntson et al., 1997; Tarvainen, 2014). The interpretation of LF rhythms is controversial (Berntson et al., 1997; Reyes del Paso, Langewitz, Mulder, van Roon, & Duschek, 2013), but LFHRV signal is often regarded as an index of sympathetic nervous system activity, or autonomic arousal in short (2-5 minute) recordings (Kiyono, Hayano, & Watanabe, 2017; Malliani, Pagani, Lombardi, & Cerutti, 1991; Task Force, 1996). Therefore, my HRV analyses focus on power in the LF band, extracted using parametric autoregressive (AR) modeling (Mailhes & Castanié, 2013) built into the Kubios HRV package (Tarvainen, 2014). I extracted LFHRV power in normalized units, and computed the change from the most recently collected baseline periods and for single-domain and mixed-domain blocks of DFD and DFE.

Two adolescents were missing data from a baseline period and were therefore excluded from analyses of the tasks that directly followed that baseline (i.e., if the baseline before DFD was missing, a change from baseline for DFD could not be computed). For one adult, HRV data could not be recorded, and for two other adults there were too many artifacts to correct. With these exclusions and the overall exclusions for biometrics, the final sample included 28 adolescents (and for some analyses 27 or 26) and 26 adults.

HRV and caffeine intake. While caffeine is known to affect some HRV measures (Sondermeijer, van Marle, Kamen, & Krum, 2002), its effect on LFHRV is not fully clear (Hibino, Moritani, Kawada, & Fushiki, 1997), and if an effect exists, may be absent in those who habitually drink caffeine (Rauh, Burkert, Seipmann, & Mueck-Weymann, 2006). To rule out caffeine's possible impact on HRV results, I asked participants to report caffeine intake on the day of the experiment, and whether this was more, less or the same as their usual intake. I estimated total caffeine intake in mg based on their reports using guidelines from Mayo Clinic (Mayo Clinic, 2014) and correlated this value with LFHRV power at the first baseline (this baseline period is most similar to past studies that have tested for effects of caffeine on HRV, and it occurs before any task could influence HRV,

Rauh et al., 2006; Sondermeijer et al., 2002). This correlation did not reach significance in the full sample (p = .39 in the total sample) or in either age group individually (p = .13 and p = .58 in adolescent and adult groups, respectively). Further, only one adolescent and one adult indicated that they had drank more caffeine than normal on the day of the study. Rather than showing increased LFHRV, the adult who drank more caffeine than usual was in the *lowest* quartile of LF power within the adult group, but was still not an outlier. The adolescent's LFHRV during the first baseline period was within one standard deviation of the mean for the adolescent group. Therefore, both participants were included in the final sample.

Skin conductance. Analyses were run using PsPM Version 3.1 (Bach & Friston, 2013) in MATLAB R2012a, and data were exported to SPSS for further analysis. I used PsPM's recommended settings to run skin conductance level (SCL) analyses, first applying a Butterworth bandpass filter with cut off frequencies of 0.0159 Hz and 5 Hz and downsampling to 10Hz. I extracted SCL for baseline periods and for each block (Single, Mixed) within each task (DFD, DFE). Task-based results are reported as the change from the most recently measured baseline period.

In analyzing the skin conductance data, I focused on skin conductance level (SCL) change from the most recent baseline by task and problem type, similar to the LFHRV analyses. A repeated measures ANOVA examining task (DFD, DFE) and domain (single, mixed) by age group (adolescent, adult) showed a main effect of task such that there was a higher change from baseline in DFD relative to DFE ($F(1,51) = 4.94$, $p = .03$). However, all of the mean change scores were negative, suggesting that SCL is higher in the baseline than during the tasks. None of the other main effects or interactions reached significance (ps > .3). It is unclear why this would be the case. Additionally, I attempted to hand-score event-related skin-conductance responses using two different sets of guidelines, and found too few responses to analyze for most participants. Due to the peculiar nature of the skin conductance data (and several attempts to use alternative analysis methods, yielding similar results), along with my confidence in the validity of the HRV data, my results and discussion focus on LFHRV as an index for autonomic arousal.

Behavioral Results

Sampling Behavior in DFE

The mean number of samples across all problems was 13.61 for adults, while the mean for adolescents was 11.03. The behavioral difference in the number of samples was significant, $t(58) = 2.14$, $p = .03$, $r^2 = .07$[xii]. Additionally, all participants took significantly more samples from the risky alternative than the safe alternative ($F(1,58) = 10.35$, $p = .002$, $\eta_p^2 = .15$) but there was no interaction between age and sampling alternative ($F < 1$). While adults took more samples than adolescents, sampling patterns did not differ qualitatively between the two groups: there were no age differences in the number of times they alternated between the two alternatives during sampling, $t(58) = .25$, $p = .80$, $r^2 = .001$. Also, despite higher average sampling rates in adults relative to teens, the mean sampling error (i.e., the difference between the observed outcomes and the underlying distribution) did not significantly differ by age ($M_{adolescents} = 13.78\%$, $M_{adults} = 12.85\%$, $t(58) = 1.13$, $p = .27$, $r^2 = .02$). Adults experienced all outcomes (i.e., both the common and rare outcomes) during sampling on 62.08% of trials, where adolescents experienced all outcomes on 58.33% of trials. A chi-square test suggested that the number of trials during which all outcomes were observed did not differ significantly by age ($X^2 = 2.11$, $p = .15$, $OR = 1.17$).

None of the sampling behavior metrics described above (mean number of samples, samples in the risky vs. safe alternatives, switching behavior, sampling error) significantly differed between the single- and mixed-domains in the adolescent or adult groups ($ps > .1$).

DFD Response Time

Response time (RT) data for DFD problems were positively skewed for both adolescents and adults, and within single- and mixed-domain problems, so I performed natural log

---

[xii] Note that in Experiment 1, distributions of sampling rates were nonnormal, so nonparametric tests were used for relevant analyses. In Experiment 2, sampling rates were normally distributed in both adolescent and adult samples so standard parametric tests were used.

transformations on the data. RTs did not significantly differ between single-and mixed-domain problems within either age group (ps > .27). Additionally, RTs did not differ by age group within either block (ps > .53).

Overall Risk Taking in DFD and DFE

I ran a repeated-measures ANOVA on risk taking behavior, using task (DFD, DFE) as a within-subjects measure and age group (adolescents, adults) as a between-subjects measure. Counter to my hypotheses, there was no significant interaction between task and age group and no significant main effect of age group. The main effect of task was marginally significant ($F(1,58) = 3.12$, $p = .08$, $\eta_p^2 = .05$), in the direction of higher risk taking during DFE ($M = 48.2\%$) relative to DFD ($M = 44.6\%$). Additionally, the proportion of risks taken in DFD across the whole sample was significantly lower than chance (50%; $t(59) = 3.02$, $p = .004$, $r^2 = .13$), suggesting general risk-averse behavior, whereas in DFE, risk taking was not significantly different from chance ($t(59) = 1.09$, $p = .28$, $r^2 = .02$).

Group-Level D-E Gap

Single-Domain Problems

To reiterate my hypotheses as they relate to Prospect Theory (PT), I predicted underweighting of rare events in DFE based on past literature, counter to PT, which predicts overweighting of rare events, and which has been shown to better correspond with DFD behavior. I predict that adolescents will not have formed PT-consistent biases, but that they will exhibit heightened rare-outcome underweighting relative to adults. Results in this section are reported in terms of *overweighting* rare events, consistent with past DFD studies and inconsistent with my DFE hypotheses, to emphasize the deviation from PT and the contrast between DFD and DFE choices. Thus, in DFE I predicted that fewer than 50% of choices would be consistent with overweighting of rare events and in DFD, I predicted that greater than 50% of choices would be consistent with rare-event overweighting.

Table 6

Group-level D-E gap in adolescents and adults for single-domain problems

Problem Number	Problem Type	Option 1 (Risky)	Option 2 (Safe)	PT Prediction (Consistent w/ rare overweighting)	Adolescents			Adults		
					DFD % Choosing PT Prediction	DFE % Choosing PT Prediction	Difference (D-E Gap)	DFD % Choosing PT Prediction	DFE % Choosing PT Prediction	Difference (D-E Gap)
1 S		-32, .1	-3, 1		60.00%	40.00%	20.00%	55.17%	48.28%	6.90%
2 S	A	-40, .2	-8, 1	Option 2	58.62%	51.72%	6.90%	66.67%	53.33%	13.33%
3 S		-34, .2	-7, 1		46.67%	43.33%	3.33%	57.14%	42.86%	14.29%
4 S		-4, .8	-3, 1		48.28%	37.93%	10.34%	65.52%	37.93%	27.59%*
5 S	B	-10, .9	-9, 1	Option 1	42.86%	14.29%	28.57%*	50.00%	13.33%	36.67%**
6 S		-14, .8	-12, 1		56.67%	36.67%	20.00%	63.33%	36.67%	26.67%*
7 S		4, .8	3, 1		53.33%	43.33%	10.00%	60.00%	40.00%	20.00%
8 S	C	10, .9	9, 1	Option 2	56.67%	10.00%	46.67%***	53.33%	50.00%	3.33%
9 S		16, .8	13, 1		56.67%	50.00%	6.67%	80.00%	33.33%	46.67%***
10 S		32, .1	3, 1		33.33%	30.00%	3.33%	56.67%	40.00%	16.67%
11 S	D	35, .2	7, 1	Option 1	30.00%	40.00%	-10.00%	43.33%	16.67%	26.67%*
12 S		29, .2	6,1		36.67%	36.67%	0.00%	43.33%	43.33%	0.00%

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

The adult patterns observed were very similar to those in Experiment 1: adult participants made choices consistent with overweighting of the rare outcome on 38% of DFE choices, and on 58% of DFD choices ($X^2 = 28.37$, $p < .001$, $OR = 2.25$). In DFD, adolescent choices generally corroborated expectations with respect to development. Namely, adolescent participants made choices consistent with overweighting of rare event in 49% of DFD problems, suggesting that prospect theory did not predict their DFD choices. Further, adolescent and adult DFD choices were significantly different ($X^2 = 6.45$, $p = .01$, $OR = 1.46$). In contrast, during DFE, adolescents made choices consistent with overweighting of rare events in 36% of DFE problems. There was a significant group-level D-E gap within the adolescent group ($X^2 = 10.64$, $p < .001$, $OR = 1.64$), but this was clearly driven by underweighting rare outcomes (i.e., choices inconsistent with PT) in DFE. Adolescents' DFE choices did not differ from the choice patterns observed in adults ($X^2 = 0.15$, $p = .70$, $OR = 1.07$).

Table 6 displays choices within each age group for each single-domain decision problem. Note that because the sample size is smaller within the adult and adolescent groups than the sample described in Chapter 2, I had less power to detect significant differences in decision

making from DFD to DFE at the problem level. In the adult group, there was a D-E gap in the predicted direction on 11 of the 12 problems, with 5 reaching significance, and one problem with no D-E gap at all (problem 12). In adolescents, 10 of the 12 problems showed D-E gaps in the hypothesized direction, but these task-based differences were generally weaker, with only two problems reaching significance. One was in the opposite direction to my prediction (problem 11), and another showed no D-E gap (problem 12).

Table 7

Group-level D-E gap in adolescents and adults for mixed-domain problems

Problem Number	Problem Type	Option 1 (Risky)	Option 2 (Safe)	PT Prediction (Consistent w/ rare overweighting)	Adolescents			Adults		
					DFD % Choosing PT Prediction	DFE % Choosing PT Prediction	Difference (D-E Gap)	DFD % Choosing PT Prediction	DFE % Choosing PT Prediction	Difference (D-E Gap)
1 M		3, .8; -32	-2, 1		55.17%	44.83%	10.34%	66.67%	43.33%	23.33%†
2 M	A	2, .9;-43	-1, 1	Option 2	66.67%	30.00%	36.67%**	66.67%	36.67%	30.00%*
3 M		4, .9; -37	-1, 1		37.93%	44.83%	-6.90%	51.72%	34.48%	17.24%
4 M		23, .2; -6	-3, 1		48.15%	48.15%	0.00%	56.67%	66.67%	-10.00%
5 M	B	35, .1; -5	-4, 1	Option 1	53.33%	36.67%	16.67%	83.33%	33.33%	50.00%***
6 M		24, .2; -7	-2, 1		30.00%	26.67%	3.33%	40.00%	50.00%	-10.00%
7 M		6, .9; -8	4, 1		56.67%	26.67%	30.00%*	66.67%	50.00%	16.67%
8 M	C	5, .9; -10	3, 1	Option 2	40.00%	36.67%	3.33%	66.67%	30.00%	36.67%**
9 M		8, .8; -9	5, 1		72.41%	58.62%	13.79%	86.67%	30.00%	56.67%***
10 M		42, .1; -1	2, 1		46.67%	36.67%	10.00%	56.67%	40.00%	16.67%
11 M	D	39, .2; -2	4, 1	Option 1	43.33%	43.33%	0.00%	50.00%	43.33%	6.67%
12 M		44, .2; -3	3, 1		37.93%	24.14%	13.79%	53.33%	56.67%	-3.33%

Note. *$p < .05$, **$p < .01$, ***$p < .001$, †$p < .1$

Mixed-Domain Problems

As in the single-domain problems, both age groups showed highly significant overall D-E gaps (Adults: Choices consistent with rare-outcome overweighting in DFE: 43% vs. in DFD: 62%, $X^2 = 26.59$, $p < .001$, $OR = 2.18$; Adolescents: Choices consistent with rare-outcome overweighting in DFE: 39% vs. in DFD: 49%, $X^2 = 8.77$, $p = .01$, $OR = 1.57$). Once again, at the group level, choices in adolescent and adult groups significantly differed within DFD ($X^2 = 11.90$, $p < .00\ 1$, $OR = 1.68$), but not within DFE ($X^2 = 1.47$, $p = .23$, $OR = 1.22$).

Problem-level results were similar across single- and mixed-domain problems in adults (*Table 7*). Within each age group, 9 out of 12 problems showed D-E gaps in the predicted direction.

Task Order

Participants completed DFD and DFE in counterbalanced order. It is possible that task order influenced the size of the D-E gap. Chi-square tests suggested that within each domain type (single and mixed), within each age group and within each task order (DFD-first, DFE-first), there was a significant D-E gap (all ps $< .05$). I additionally ran paired-samples t-tests between task-order groups (DFD-first vs. DFE-first) on the size of the gap within each domain type and age group. In adolescents, the D-E gap did not significantly vary based on the first task in the single-domain ($t(11) = .39$, $p = .70$, $r^2 = .01$) or in the mixed-domain ($t(11) = .05$, $p = .96$, $r^2 < .001$) problems. In contrast, the adult DFD-first group showed a significantly higher D-E gap than the DFE-first group in both single- ($t(11) = 2.25$, $p = .05$, $r^2 = .32$) and mixed-domain ($t(11) = 2.69$, $p = .02$, $r^2 = .40$) problems.

Within-Subject D-E Gap

In adults, the within-subjects single-domain D-E gap variable was similar to that shown in Experiment 1, with a mean of .21 and a standard deviation of .29 (*Figure 9*). After removing two outliers with D-E gaps of -.50, the mean was .25 with a standard deviation of .22. Consistent with the data at the group level, the within-subjects D-E gap was lower in the adolescent sample than in the adult sample, with a mean of .12 and a standard deviation of .24. Both the adolescent and adult D-E gaps were significantly greater than 0 (adolescents: $t(29) = 2.79$, $p = .009$, $r^2 = .21$; adults: $t(27) = 6.06$, $p < .001$, $r^2 = .58$). However, the difference between the adolescent and adult gaps was significant, $t(56) = 2.18$, $p = .03$, $r^2 = .08$. The gap did not correlate with the mean number of samples taken in adolescents ($r = .16$, $p < .49$), but this correlation was highly significant in adults ($r = -.59$, $p < .001$), such that those who took more samples had a lower D-E gap.

The mixed-domain within-subjects D-E gap was similar to the single-domain gap in both adolescents ($M = .11$, $SD = .21$) and adults ($M = .19$, $SD = .28$; after removing one outlier), but the difference in the mixed-domain D-E gap by age group was only marginally significant ($t(57) = 1.93$, $p = .06$, $r^2 = .06$). Both the adolescent and adult groups had D-E gaps greater than 0 (adolescents: $t(29) = 2.73$, $p = .01$, $r^2 = .20$; adults: $t(28) = 5.10$, $p < .001$, $r^2 = .48$). Once again, the mixed-domain gap correlated with the mean number of samples in adults ($r = -.62$, $p < .001$) but not in adolescents ($r = .32$, $p = .18$). The difference between the mixed- and single- domain gaps did not reach significance in either adolescents or adults (ps > .72).

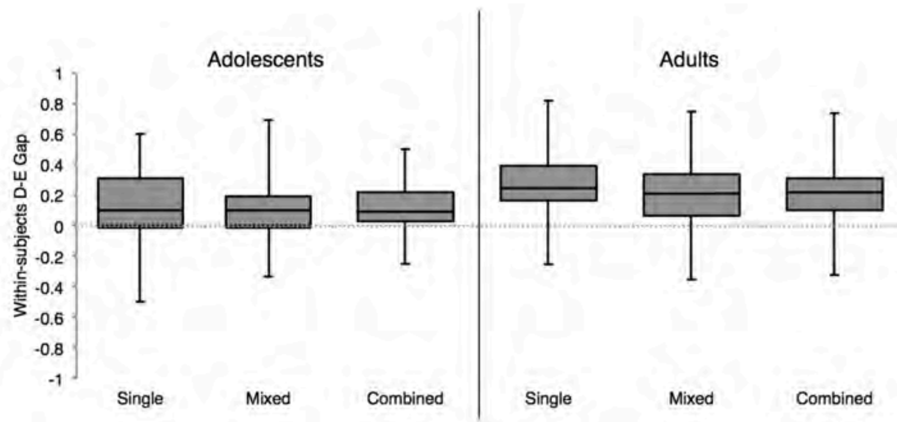


Figure 9. Within-subjects D-E gap measure by age group and problem type (single-domain, mixed-domain, single- and mixed-domain combined).

I also computed a within-subject D-E gap measure combining all 24 problems (from both single- and mixed-domain blocks). The mean combined gap in adults was .22 ($SD = .22$) and the mean combined gap in adolescents was .12 ($SD = .17$). Both were higher than 0 (adolescents: $t(29) = 3.65$, $p < .001$, $r^2 = .31$; adults: $t(28) = 5.60$, $p < .001$, $r^2 = .53$), but the difference in D-E gap by age group reached significance ($t(57) = 2.14$, $p = .04$, $r^2 = .07$). The gap measure correlated negatively with the average number of samples taken in adults ($r = -.65$, $p < .001$) but not in adolescents ($r = .33$, $p = .15$). In contrast to the group-level analysis above, where adults showed a larger D-E gap if they completed DFD first, the single-domain, mixed-domain and

combined gap measures, did not differ significantly by the first task completed within either age group (*p*s > .3).

Within-Subjects Gap and Sampling Error

Given that the D-E gap in single-domain problems, mixed-domain problems and across all problems correlated negatively with the mean number of samples in adults, it is possible that the gap (at least in adults) can be explained fully by sampling error (i.e., it is possible that more sampling led to a more accurate sampling of the distribution and therefore a smaller gap). As in Experiment 1, I explored the role of sampling error on the D-E gap within subjects, with the prediction that in participants with a high D-E gap, PT is a better predictor of DFD choices than DFE choices, even during trials where all outcomes (both common and rare) were encountered during sampling, and after controlling for sampling error. I further predicted an age by task type by D-E gap size interaction, such that for adolescents with a particularly high gap, PT would be an especially bad predictor of choice in DFE. The analyses that follow in this section (and some analyses in the following sections) are limited to trials during which all outcomes were observed, which I henceforth refer to as "all-outcome trials."

Single-domain problems. To examine the possible role of sampling error in the D-E gap, I modeled PT predictions (Venkatraman et al., 2009) for each single-domain problem based on the sampling outcomes that each participant actually experienced during all-outcome trials. I then divided the single-domain gap into high (.25 and larger, $N_{adolescents} = 10$, $N_{adults} = 17$) and low (less than .25, $N_{adolescents} = 20$, $N_{adults} = 13$) groups[xiii]. Next, I ran a repeated-measures ANOVA to examine whether the proportion of PT-predicted choices varied by task (DFD, DFE), size of the D-E gap (High, Low), or age group (adolescents, adults), controlling for the average number of samples. There was a significant task by gap size interaction ($F(1, 55) = 18.29$, $p < .001$, η_p^2 =

[xiii] Note that in this analysis, there are unequal numbers of adolescents and adults in each group because the mean D-E gap was different in the two groups, changing the definition of "High" and "Low" gap. Instead of defining the variable differently across the age groups, I chose to keep the definition consistent across the age groups.

.25). Interestingly, there was no evidence of a main effect or any interactions with age group (*p*s > .7), which suggests that adolescents with low gaps and those with high gaps indeed decided similarly to adults with low and high gaps, respectively.

Notably, the pattern of results in the present sample differed slightly from that in Experiment 1. As shown in *Figure 10*, for participants with a high D-E gap, PT was a good predictor of DFD behavior, and accounting for sampling error still did not close the D-E gap (i.e., PT still did not predict as many choices in DFE as DFD after accounting for sampling error in all-outcome trials). This result is similar to that seen in Experiment 1. In those with a low gap, the pattern was reversed: PT was *not* a good predictor of DFD behavior, and was better at predicting DFE behavior after controlling for sampling error. Conversely, in Experiment 1, those with a low gap exhibited similar PT-consistent behavior in DFD and DFE (i.e., there was no significant difference between DFD- and DFE-predicted behavior in the low gap group).

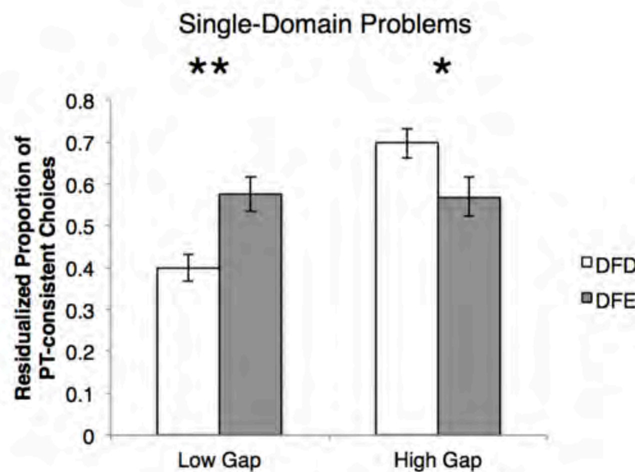


Figure 10. Choices in single-domain problems in DFD and DFE consistent with PT, controlling for sampling error in all-outcome trials, collapsed across age. *$p < .05$, **$p < .001$

Mixed-domain problems. I ran the same analysis described above within in the mixed domain. While the D-E gap measure was not significantly lower in the mixed relative to the single-domain problems, the distributions were shifted slightly lower in both adolescent and adult

groups. In the context of these problems, I defined a high gap as .15 and higher to get the most even spread between high and low gaps across age groups (High gap: $N_{adolescents}$ = 12, N_{adults} = 21; Low gap: $N_{adolescents}$ = 17, N_{adults} = 9; one adolescent was excluded for never experiencing all outcomes in mixed problems).

In the mixed-domain ANOVA, the 3-way interaction (task x gap size x age group) was marginally significant, $F(1, 54) = 3.71$, $p = .06$, $\eta_p^2 = .06$. Specifically, in adolescents, the pattern observed was similar to that in the single domain problems (although only trending significant in the high-gap group), while in adults, sampling error appeared to fully explain the D-E gap; (*Figure 11*). As would be expected, results from the combined D-E gap were intermediate to those observed in single- and mixed-domains, so they are not reported further. Instead, I use a different approach to look at behavioral patterns across all problems in the next section by age group and task. This analysis also did not require use of the high vs. low gap variable, which required using different sample sizes within adolescent and adult groups.

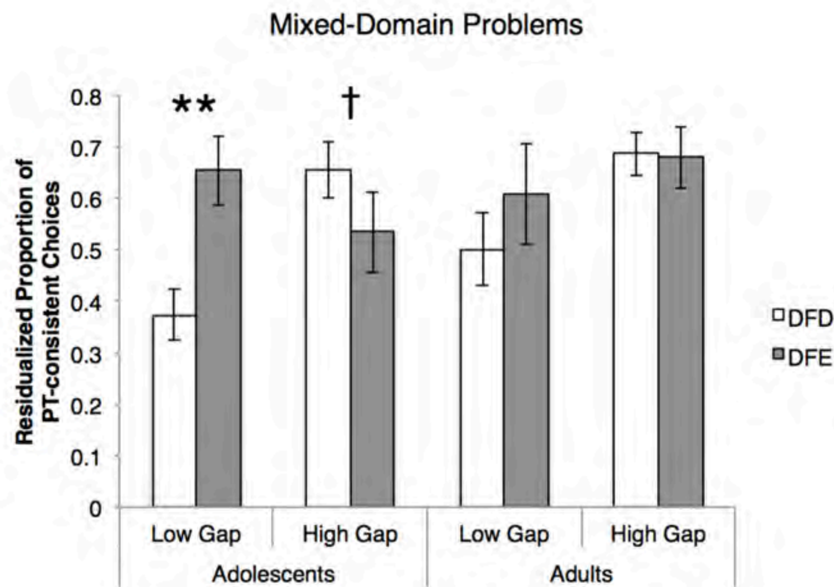


Figure 11. Choices in mixed-domain problems in DFD and DFE consistent with PT after controlling for sampling error in problems where all outcomes were observed. $^{**}p < .001$, † $p = .11$

Exploring DFD and DFE Choices Using Mixed-effects Logistic Regressions

In group- and participant-level D-E gap analyses above, it appears that adolescents tend to exhibit a smaller gap at the group level, and they seem to decide more like adults in DFE than DFD. In hopes of gaining a better understanding of decision patterns by age, I ran mixed-effects logistic regressions, treating participant as a random factor, to predict decision making in DFD and DFE. Because single and mixed problems produced similar results in the above behavioral analyses, and to allow for more power in the regression analyses, I combined data for single and mixed problems, but included a regressor testing whether choices systematically varied between single and mixed problems.

Within each task I ran a series of five regressions in adolescent and adult samples separately. The first four tests included only one predictor in each, and the fifth included all four predictors at once. The four predictors used were: 1. whether the problem was in the gain or loss domain (coded as Gain = 1, Loss = 0; in mixed problems, the EVs of all problems were in either the gain or loss domain, despite the risky option containing a possible outcome in the other domain); 2. whether the rare outcome was unfavorable relative to the common outcome in the risky alternative (coded as Rare Unfavorable =1, Rare Favorable = 0); 3. whether the problem was in the single or mixed-domain (coded as Single = 1, Mixed = 0); and 4. the difference in expected value between the risky and safe outcomes as a predictor of risk taking (a higher value translates to higher EV in the risky relative to the safe outcome). Results can be interpreted in terms of odds ratios (i.e., odds of risk taking in a problem), with a value of 1 indicating no effect of a variable. A value significantly higher than 1 indicates more risk taking at a higher value of that predictor (e.g., for the gain vs. loss predictor, an odds ratio higher than 1 indicates more risk taking in gain problems, which are coded as 1, than loss problems, which are coded as 0).

I ran these 5 regressions for both DFD and DFE. Within DFE, I ran two series of regressions: one with all trials included, and another only including all-outcome trials (i.e., trials

where sampling produced both the common and rare outcome in the risky alternative at least once).

DFD

Results for each of the 5 regressions in DFD by age group are shown in *Table 8. Figure 12* plots odds of risk taking by each of the predictors for adolescents (*Figure 12A*) and adults (*Figure 12B*) from Regression 5. In adults, the Rare Unfavorable, Gain and EV-Difference regressions were all significant predictors of risk taking on their own. Specifically, adults were less likely to take a risk when the rare outcome was unfavorable (vs. favorable). This finding is consistent with PT, which predicts overweighting of a rare outcome (and overweighting a rare unfavorable outcome in the risky alternative would decrease the odds of choosing that alternative). Additionally, adults took fewer risks in gain problems than loss problems, which is also consistent with PT, because losses are thought to "loom larger" than gains (Tversky & Kahneman, 1992), making risk taking more attractive when attempting to avoid a loss. Finally, adults were more likely to take a risk when the EV of the risky alternative is higher. The regression containing all moderators (Regression 5) showed that each of these variables explained unique variance in choice patterns. The single vs. mixed problem regression did not predict risk taking, nor did this predictor reach significance in the full regression.

In contrast, adolescents' risk behavior in DFD does not appear to vary systematically by any of the variables of interest. Interestingly, when all four of the regressors were entered at once (Regression 5) the difference between EVs of risky and safe alternatives reached significance ($p < .03$), such that adolescents are more likely to take a risk if the EV of the risky option is higher. Additionally, both the gain and rare unfavorable moderators were marginally significant in the full regression, (both ps $< .08$). The directionality of the gain predictor was the same as in adults: adolescents are less likely to take a risk in the gain domain relative to the loss domain. However, adolescents were slightly *more* likely to take a risk if the rare outcome was unfavorable, counter to PT predictions.

Table 8

Predicting DFD risk taking in adolescents and adults

	Models				
Adolescents	1	2	3	4	5
Rare Unfavorable	0.09 (0.15)				0.34 (0.19)†
Gain		-0.24 (0.15)			-0.28 (0.16)†
Single Domain			-0.05 (0.15)		0.12 (0.17)
Difference in EV				0.08 (0.06)	0.18 (0.08)*
χ^2	0.38	2.35	0.09	1.68	7.67†
df	1	1	1	1	4
Adults	1	2	3	4	5
Rare Unfavorable	-0.85 (0.16)***				-0.63 (0.19)***
Gain		-0.37 (0.15)*			-0.43 (0.16)**
Single Domain			0.05 (0.15)		0.22 (0.18)
Difference in EV				0.25 (0.06)***	0.16 (0.09)*
χ^2	30.35***	5.95*	0.09	17.12***	40.97***
df	1	1	1	1	4

Note. Beta values are listed with standard error in parentheses. **p* < .05, ***p* < .01, ****p* < .001, † *p* < .1.

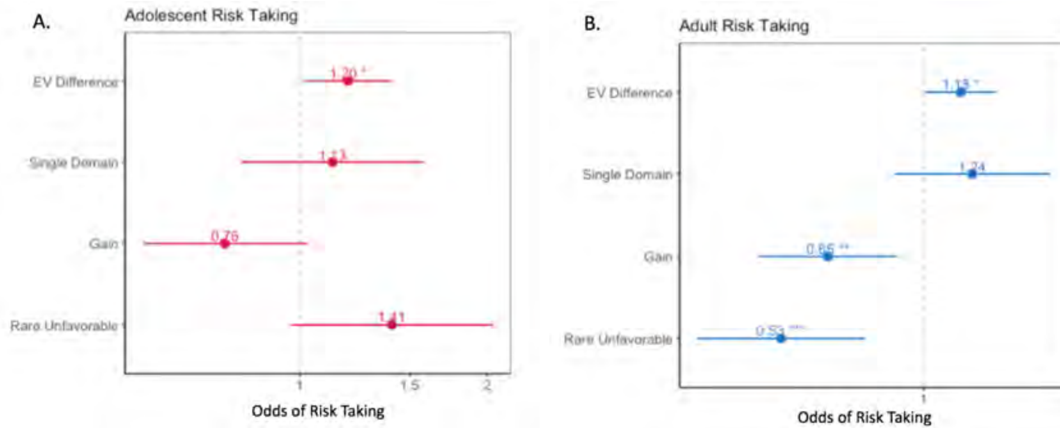


Figure 12. Odds ratios for risk taking in DFD by problem parameter (as calculated in Regression 5, *Table 8* in A. adolescents and B. adults. Error bars represent 95% confidence intervals. The value 1 on the x-axis represents no effect of the variable. A value significantly lower than 1 represents *lower* risk taking when the binary variable is set to 1 (i.e., lower risk taking in gain problems or when the rare outcome is unfavorable). **p* < .05, ****p* < .001

DFE

All DFE trials. In contrast to DFD where adolescent and adult choices diverge, choices were quite similar between the age groups in DFE. In both the simpler models and the overall model for both adolescents and adults, participants were more likely to take a risk when the rare outcome was unfavorable. This result is another demonstration of the D-E gap, as it suggests that participants are likely to underweight the rare outcome in DFE decision making (counter to PT predictions). Participants were also more likely to take a risk when the EV was higher in the risky

option than the safe option (*Table 9* and *Figure 13*). In adults, the model including only single vs. mixed domain as a predictor was significant (more risk taking in single domain) but this predictor was not significant as part of the larger model.

Table 9

Predicting risk taking across all DFE trials

	Models				
Adolescents	1	2	3	4	5
Rare Unfavorable	1.08 (0.16)***				1.78 (0.2)***
Gain		0.13 (0.15)			0 (0.19)
Single Domain			-0.06 (0.15)		0.35 (0.19)
Difference in EV				0.23 (0.02)***	0.28 (0.03)***
$\chi2$	48.46***	0.73	0.15	146.97***	241.74***
df	1	1	1	1	4
Adults	1	2	3	4	5
Rare Unfavorable	0.79 (0.15)***				1.27 (0.19)***
Gain		0.1 (0.15)			0.05 (0.18)
Single Domain			-0.5 (0.15)***		-0.25 (0.18)
Difference in EV				0.29 (0.03)***	0.29 (0.03)***
$\chi2$	26.95***	0.42	10.89***	187.5***	240.02***
df	1	1	1	1	4

Note. Beta values are listed with standard error in parentheses. ***$p < .001$

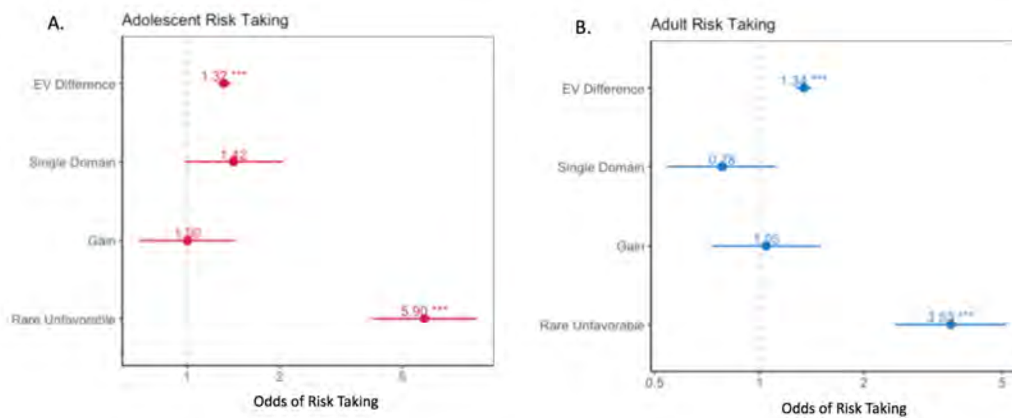


Figure 13. Odds ratios for risk taking in DFE by problem parameter (as calculated in Regression 5, *Table 9* in A. adolescents and B. adults. Error bars represent 95% confidence intervals. The value 1 on the x-axis represents no effect of the variable. A value significantly lower than 1 represents *lower* risk taking when the binary variable is set to 1 (i.e., lower risk taking in gain problems or when the rare outcome is unfavorable). ***$p < .001$

DFE all-outcomes trials. In the ANOVAs run in the prior section (using a binary high vs. low D-E gap predictor variable), I found that the D-E gap was diminished (but not eliminated) for trials where participants experienced all of a problem's possible outcomes. To address sampling bias differently, I ran the same regressions as in the prior section, but only on all-outcomes trials. Once again, both adolescents and adults were sensitive to EV, such that if the risky outcome had a higher EV it was more likely to be chosen (*Table 10, Figure 14*). When only the mixed vs. single domain variable was included as a predictor (Model 3), adults were more likely to take a risk in the mixed relative to the single domain, but this pattern was not seen in the full model. When the rare unfavorable variable was the sole predictor in the model (Model 1), both adolescents and adults were less likely to take a risk if the rare outcome was unfavorable, consistent with PT predictions. However, after controlling for all of the problem parameter variables in Model 5, adolescents, but not adults, were significantly more likely to take a risk when the rare outcome was unfavorable than when the rare outcome was favorable, even when they observed that rare outcome in sampling.

Table 10

Predicting DFE risk taking on all-outcome trials

	Models				
Adolescents	1	2	3	4	5
Rare Unfavorable	-0.49 (0.21)*				0.8 (0.29)**
Gain		0.14 (0.21)			-0.2 (0.23)
Single Domain			-0.11 (0.21)		0.32 (0.23)
Difference in EV				0.14 (0.02)***	0.2 (0.03)***
χ^2	5.6*	0.48	0.28	48.27***	57.60***
df	1	1	1	1	4
Adults	1	2	3	4	5
Rare Unfavorable	-0.85 (0.2)***				0.3 (0.27)
Gain		0.38 (0.2)†			0.03 (0.23)
Single Domain			-0.47 (0.2)*		-0.17 (0.23)
Difference in EV				0.21 (0.03)***	0.23 (0.04)***
χ^2	17.93***	3.75†	5.55*	83.63***	85.68***
df	1	1	1	1	4

Note. Beta values are listed with standard error in parentheses. *$p < .05$, **$p < .01$, ***$p < .001$

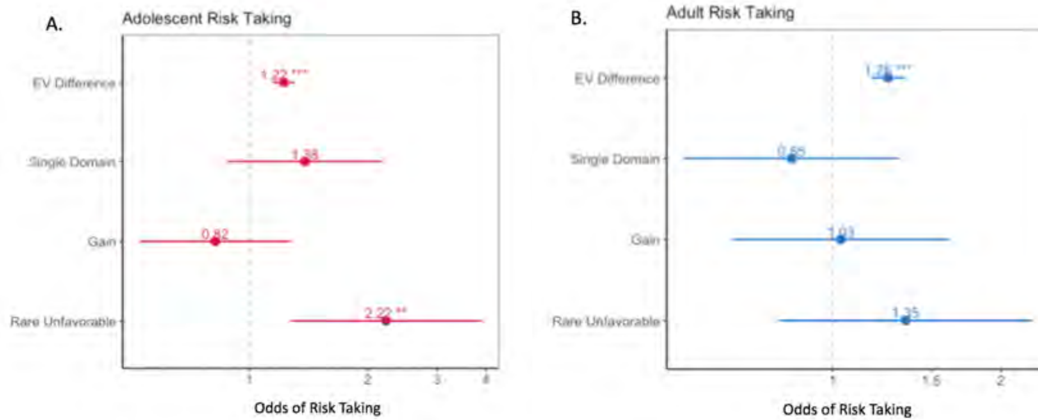


Figure 14. Odds ratios for risk taking in DFE by problem parameter for trials where all outcomes were observed (Regression 5, Table 10), A. in adolescents and B. in adults. $^{**}p < .01$, $^{***}p < .001$

Shapes Estimation Task

In order to ensure that differences in the DFE behavior were not merely due to adolescents' inability to estimate probabilities learned through sampling, I asked participants to sample shapes (rather than values) and to report probability estimates for the sampled distribution. I found, first of all, that participants' mean number of samples was significantly correlated across the DFE and shapes sampling tasks in both adolescents (r = .52, p = .003) and adults, (r = .47, p = 008). Sampling means by task and age group are displayed in *Figure 15*. A repeated-measures ANOVA showed a significant age group by task interaction (F(1,58) = 5.58, p = .02, η_p^2 = .09), and a main effect of task (more sampling in shapes than DFE, F(1,58) = 106.77, p < .001, η_p^2 = .65). Tukey LSD tests showed that adolescents sampled less than adults in DFE (p = .04), but that there was no significant age difference in sampling during the shapes estimation task (p = .20, although interestingly, adolescents sampled numerically more than adults in the latter case).

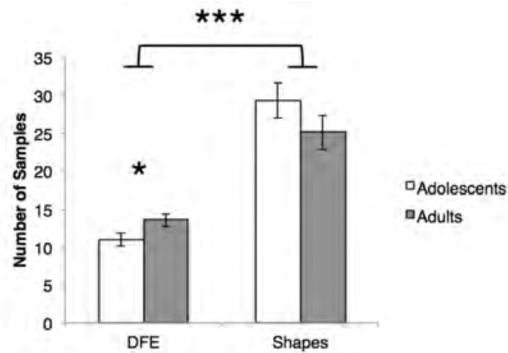


Figure 15. Sampling in DFE and Shapes Estimation tasks by age group. $*p < .05$, $***p < .001$

In the shapes task there were two types of trials: those with an 80:20 distribution of shapes and those with a 90:10 distribution (i.e., the rare shape either comprised 20% of the distribution or 10% of the distribution, respectively). To analyze the estimation data, I took the difference between the distributions participants observed through sampling and the participants' estimates of those distributions separately for the 80:20 distributions and the 90:10 distributions. A repeated-measures ANOVA did not show a significant age group by distribution type (80:20, 90:10) interaction ($F < 1$) or a significant main effect of distribution type. The main effect of age group was marginally significant ($F(1,58) = 3.41$, $p = .07$, $\eta_p^2 = .06$), such that adolescents were slightly *more* accurate at estimating the shape distributions than adults. Simple main effects analysis showed that within the 90:10 distribution, adolescents were marginally more accurate than adults ($p = .07$), but in the 80:20 distribution there was no age difference ($p = .29$; *Figure 16*).

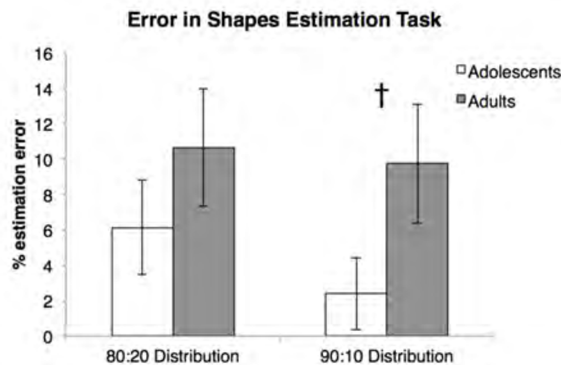


Figure 16. Estimation error by distribution type and by age in the shapes estimation task. †$p < .1$.

Physiological Results

DFD Eye Tracking Results

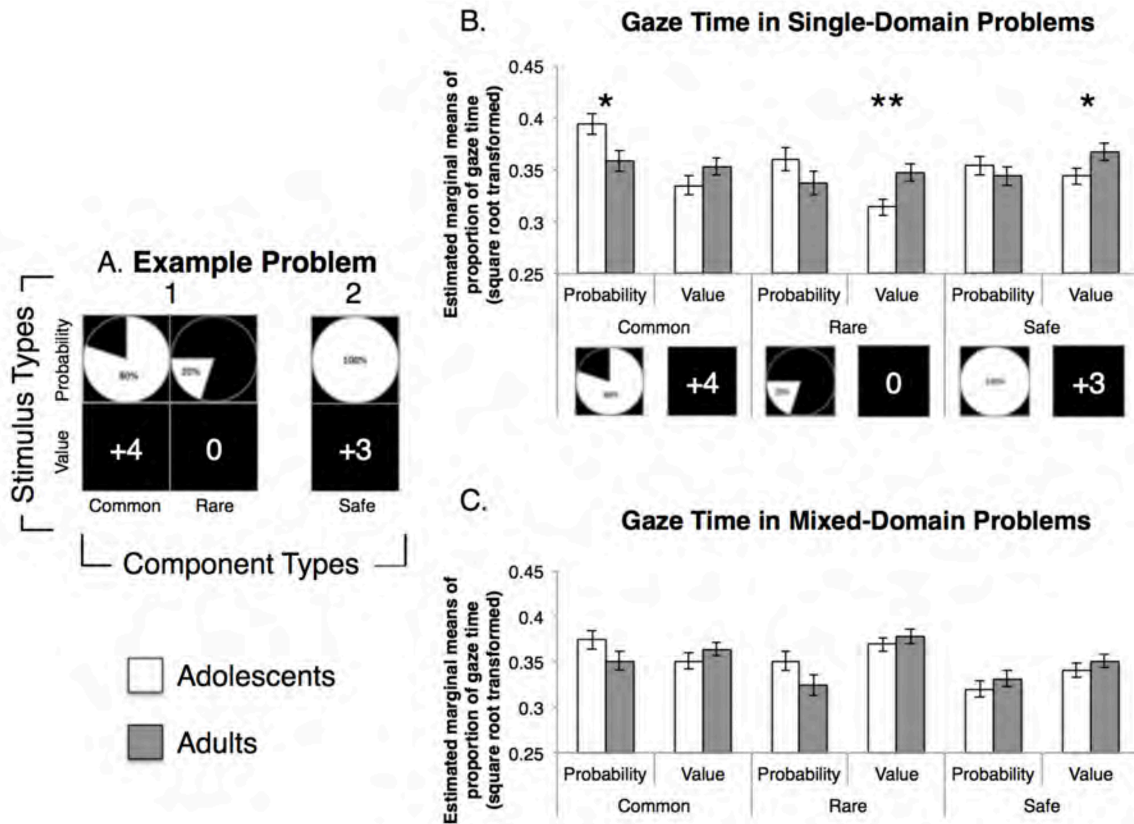


Figure 17. Eye tracking results in DFD. A. Example DFD problem, with Area of Interest (AOI) names defined. B. Gaze time results by AOI and age group for single-domain problems, with AOIs from example problem displayed under the corresponding results. C. Gaze time results for mixed-domain problems. Significance in B and C only shown for age contrasts; contrasts by stimulus and component type are further explored in *Figures 18 and 19.* $^{*}p < .05$, $^{**}p < .01$.

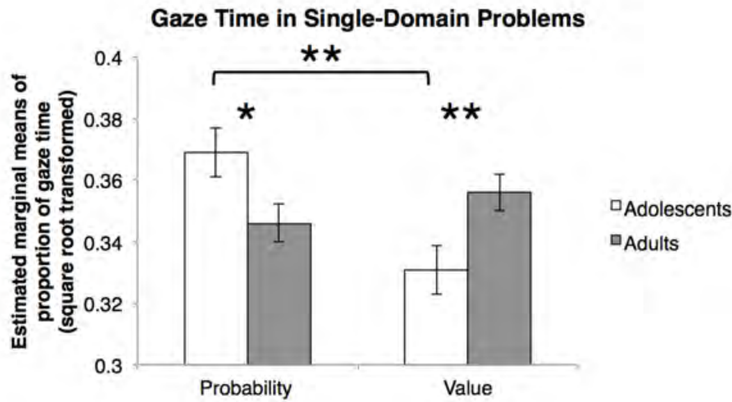
In DFD, each decision problem had six defined areas of interest (AOIs; *Figure 17.A.*) for eye tracking analysis. Analyses focused on the proportion of valid gaze time in each of those AOIs. Within the risky alternative, there was a *common probability* and *common outcome,* as well as a *rare probability* and *rare outcome*. Within the safe alternative there was a *safe probability* and a *safe outcome*. I capitalized on the common structure of the decision problems and, in a

series of ANOVAs, examined the impacts of "stimulus type" (coding the probability vs. value contrast) and "component type" (coding the common, rare and safe items).

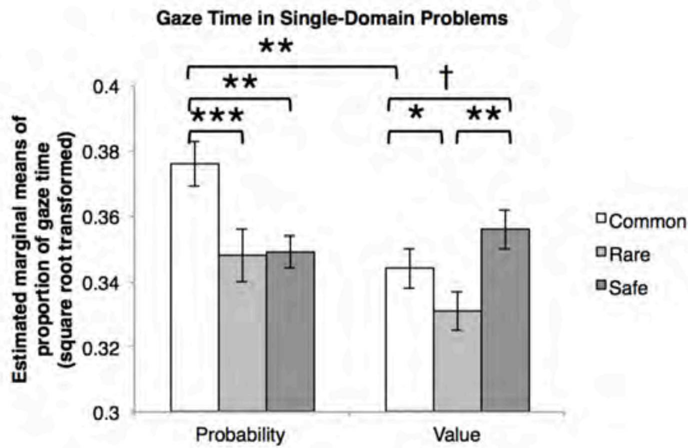
Initially a full 4-way ANOVA was tested, entering stimulus type (probability, value), component type (common, rare, safe) and domain (single, mixed) as within-subjects factors, and age group (adolescent, adult) as a between-subjects factor. There was a highly significant three-way stimulus type by component type by domain interaction ($F(2, 102) = 9.26$, $p < .001$, $\eta_p^2 = .15$). To aid in interpretation of this interaction, additional analyses were conducted separately for single and mixed problems, entering proportion of valid looking time for stimulus type and component type as within-subjects measures and age group as a between-subject measure. Results are displayed in *Figure 17*.

Single-Domain Eye Tracking Results

Within the single-domain problems, the three-way interaction between stimulus type, component type and age group did not reach significance ($F(2, 102) = 1.13$, $p = .33$, $\eta_p^2 = .02$). However, to thoroughly explore the possible effect of age group on DFD processing, I ran simple main effects tests on looking time as a function of age group, within all of the AOIs (*Figure 17.B.*). These tests revealed that adolescents looked significantly longer than adults at the common probability in the risky alternative ($p = .02$), while adults looked longer at the rare value ($p = .003$) and the safe value ($p = .05$). These effects likely contribute to a significant component type (probability, value) by age group interaction ($F(1, 51) = 7.60$, $p = .008$, $\eta_p^2 = .13$; *Figure 18.A.*), such that adolescents spent more time looking at probabilities overall than adults ($p = .05$), and adults spent more time looking at values than adolescents ($p = .01$). Additionally, adolescents looked significantly longer at probabilities than values ($p = .01$), while adults looked numerically (but not significantly) longer at values than probabilities. There was also a significant stimulus type by component type interaction ($F(2, 102) = 10.88$, $p < .001$, $\eta_p^2 = .18$). Within the probability types, participants across age groups looked most at common probabilities. Within value types, participants looked least at the rare values.



A.



B.

Figure 18. Single-domain eye tracking data. A. Gaze time by probability and value, separated by age. B. Gaze time by stimulus type and component type collapsed across age. $^{*}p < .05$, $^{**}p < .01$, † $p < .1$

Mixed-Domain Eye Tracking Results

In the mixed-domain ANOVA, there was a significant stimulus type by component type by age group interaction $F(2,104) = 3.19$, $p = .05$, $\eta_p^2 = .06$. However, none of the simple main effects of age in this interaction reached significance (ps > .13; *Figure 17.C.*), which reflects the fact that different pairwise comparisons within stimulus type or problem type reach significance depending on age group (*Figure 19.A.* presents the same data as *Figure 17.C.*, reorganized to emphasize differences within an age group). While both age groups looked longest at the rare probabilities relative to the other two probability AOIs (ps < .09), adolescents looked longer at the

rare probability than the safe probability (p = .03), but there was no difference between rare and safe probability AOI looking times in adults. Additionally, both age groups looked longest at the rare value relative to the safe and common values, but adults look significantly longer at the rare value than the rare probability ($p < .001$), and none of the other probability vs. value contrasts reached significance within either age group (ps > .1).

While there was a significant 3-way interaction, which diminishes the interpretability of lower-level interactions, the age group by stimulus type (probability, value) pattern, collapsed across component type (common, rare, safe), was similar to that observed in the single-domain problems. Adolescents looked (numerically) longer at probabilities and adults looked longer at values (*Figure 19.B.*). The simple main effect of stimulus type within the adult group suggested that they looked longer at values than probabilities (p = .01).

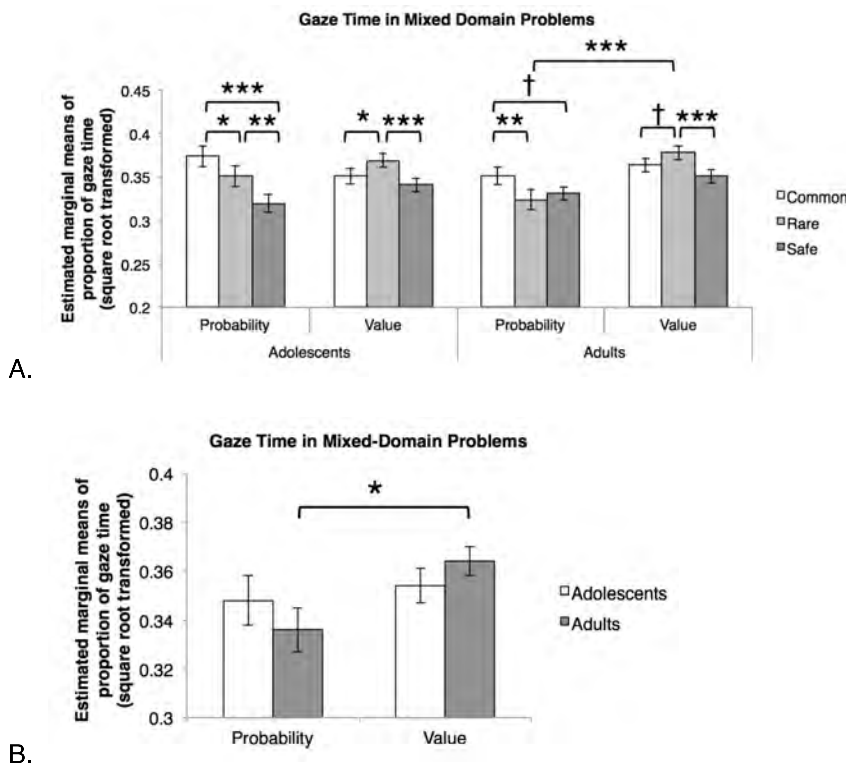


Figure 19. A. Gaze time by age group. These are the same data from *Figure 17.C.* reorganized to emphasize differences in gaze time by stimulus type and component type with an age group. B. Gaze time in mixed-domain problems by probability and value, separated by age. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$, † $p < .1$

Linking Eye Tracking to Behavioral Results

I also explored whether gaze time at each of the AOIs predicted choices, using mixed-effects logistic regressions. I first entered each of the six AOIs and their interactions with age group into a regression. There were several significant interactions with age group. In order to more thoroughly investigate these age differences, I ran separate models for adolescents and adults, which are reported in *Table 11*. Odds ratios for each of the AOI gaze time predictors are plotted in *Figure 20.* Interestingly, the only significant predictors of adult risk taking were gaze time within the safe probability and safe value AOIs, such that the more time they looked at these safe AOIs, the less likely they were to take a risk (i.e., the more likely they were to choose the safe option). In adolescents, gaze time for the safe probability was a significant predictor of risk taking in the same direction as adults (more looking at the safe probability predicts less risk taking). However, adolescents' gaze time at both the common value and probability also significantly predict higher risk taking.

Table 11

Predicting risk taking from gaze time on each AOI

	Adolescents	Adults
Common Probability	2.33 (0.88)**	1.36 (0.8)
Common Value	3.04 (0.97)**	0.53 (0.89)
Rare Probability	1.24 (0.85)	1.08 (0.77)
Rare Value	0.67 (0.94)	1.02 (0.87)
Safe Probability	-2.67 (0.89)**	-5.26 (0.97)***
Safe Value	-0.95 (1.06)	-3.69 (1.03)***
X2	36.36***	70.22***
df	6	6

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

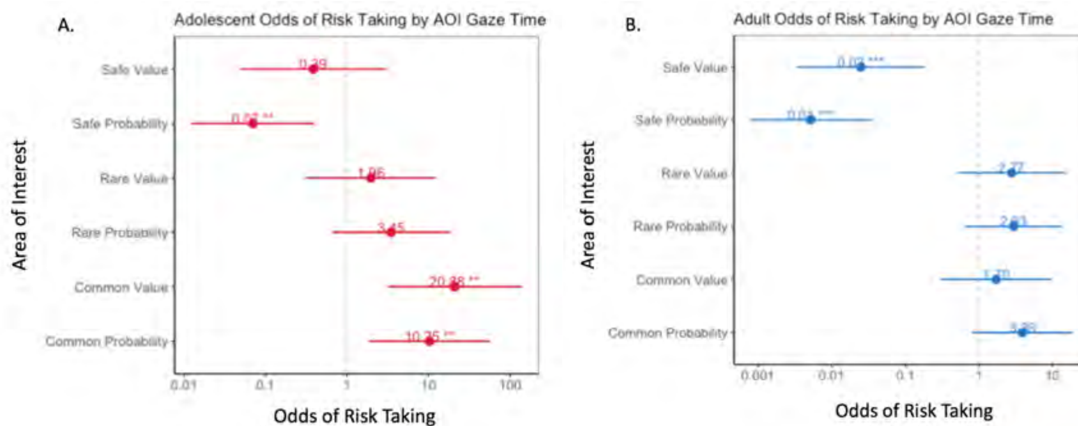


Figure 20. Odds ratios predicting risk taking by gaze time, from the regressions in Table 11 in A. adolescents and b. Adults.

Heart Rate Variability Analyses Across Tasks

During Task Completion

In order to assess affective arousal level, I focused on low-frequency heart rate variability (LFHRV), indexed as a change from the most recently collected baseline. I ran a repeated measures ANOVA examining task (DFD, DFE) and domain (single, mixed) by age group (adolescent, adult). The three-way task by domain by age group interaction did not reach significance, $F(1,50) = .38$, $p = .54$, $\eta_p^2 = .01$, and counter to my hypotheses, there were no significant effects (main effect or interactions) of domain. However, there was a significant age group by task interaction ($F(1,50) = 9.05$, $p = .004$, $\eta_p^2 = .15$; *Figure 21*). Interestingly, while the increase in LFHRV from baseline was higher in DFE than DFD within adolescents ($p = .04$), the difference from baseline was significantly *lower* in adults in DFE than in DFD ($p = .04$). Additionally, the difference between adolescent and adult LFHRV change from baseline within DFE was highly significant ($p < .001$).

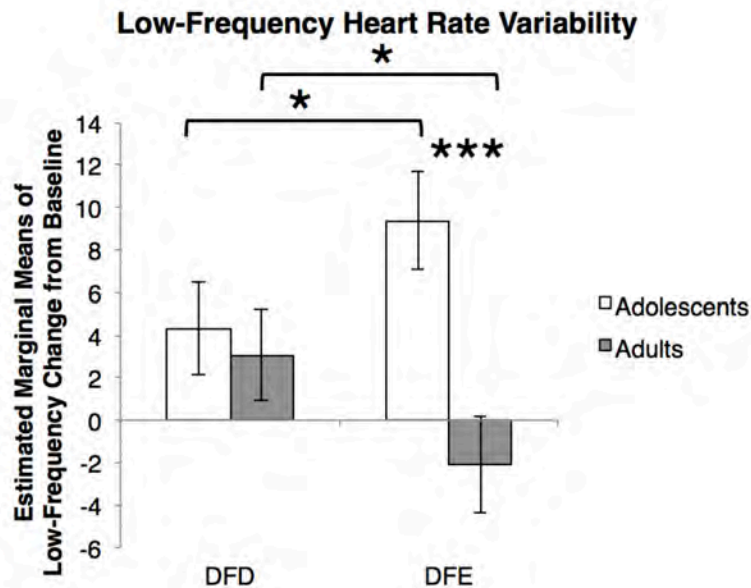


Figure 21. Low-frequency heart rate variability by task and age. *$p < .05$, ***$p < .001$.

Sustained Changes in Low-Frequency Heart Rate Variability

I also examined whether change in LFHRV was sustained after task completion into the baseline session directly following the first task. Because the first task completed (DFD or DFE) was counterbalanced, and because I ran a 3-minute baseline directly before and directly after the first task, I was able to examine whether the identity of the first task affected LFHRV during the baseline following the first task, relative to the baseline before the first task, as a function of age. A 2 (first and second baseline, within-subjects) by 2 (first task, between-subjects) by 2 (adolescents, adults, between-subjects) ANOVA did not show a significant 3-way interaction ($F(1,48) = 1.47$, $p = .23$, $\eta_p^2 = .03$). However, simple main effects showed that adolescents who completed DFE first showed a significant increase in LFHRV ($p = .03$; *Figure 22*), while none of the other groups showed a significant increase in LFHRV from baseline 1 to baseline 2 (*p*s > .14). This result provides further support for the notion that adolescents show a unique increase in LFHRV when they complete DFE relative to DFD, a pattern that is not seen in adults.

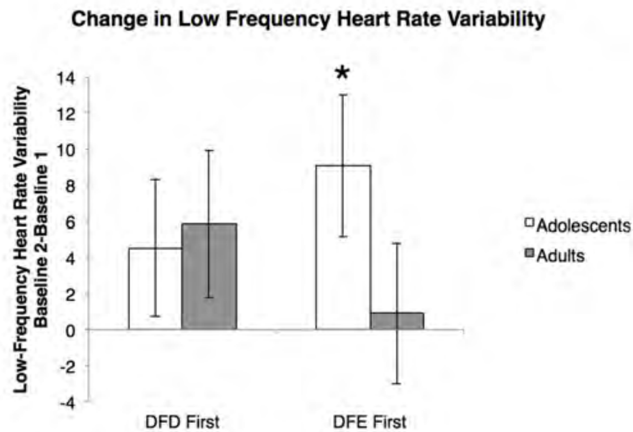


Figure 22. Change in low-frequency heart rate variability from baseline 1 to baseline 2 by first task and age group. *$p < .05$.

Low-Frequency Heart Rate Variability and Decision Making

I examined how arousal indexed by LFHRV during DFE related to behavior during DFE task performance. To this end, I ran correlations between LFHRV and the proportion of choices consistent with underweighting rare outcomes during DFD by age group. I did not find a significant correlation between task behavior and LFHRV change from baseline in either age group (*p*s > .19). I also assessed whether risk taking during DFE was related to LFHRV during DFE, and once again did not find a significant correlation between my HRV measure and behavior.

Individual Differences in the D-E Gap

Go/No-Go. I explored the possibility that go/no-go performance is related to the D-E gap as it was in the sample in Chapter 2. I ran a correlation between correct rejections (CRs) and the D-E gap measure separately in adolescents and adults. The correlation was not significant in either age group (*p*s >.90).

Delay Discounting. I additionally explored whether the mean indifference point in delay discounting was correlated with the D-E gap measure. The correlation was not significant in adolescents or adults (*p*s >.19).

Self-report measures. As an exploratory analysis, additional correlations were tested between self-report data and measures related to the D-E gap, for each age group separately. Three sets of correlations with behavioral variables were examined. The first involved general decision styles measures, including the 5 items from Generalized Decision Making Scale (Scott & Bruce, 1995), maximizing and regret (Schwartz et al., 2002). The second set of correlations involved measures thought to contribute to heightened adolescent risk taking, including the Barratt Impulsivity Scale (Patton et al., 1995), and the disinhibition and sensation seeking scales (Zuckerman et al., 1978, 1993). Last, I ran correlations between task behavior and self-reported real world risk taking and risk perception (Benthin et al., 1993).

The correlations are presented in *Table 12*. No correlations reached significance after controlling for family-wise error rate within each age group (.05/16 tests per age group = .003). However, the Generalized Decision Making Scale Avoidant measure correlated positively with the D-E gap measure using uncorrected p-values in both adolescents (r = .42, p = .03) and adults (r = .37, p = .05). None of the other correlations reached uncorrected significance, ps > .05.

Table 12

Correlations between self-report measures and the D-E gap by age

Test Set	Scale	Adolescent D-E Gap	Adult D-E Gap
Set 1	GDMS Intuitive	.36	.28
	GDMS Rational	-.10	-.03
	GDMS Avoidant	.42 †	.37 †
	GDMS Spontaneous	.25	.13
	GDMS Dependent	.13	.29
	Maximization	-.33	.28
	Regret	-.12	.29
	Need for Cognition	.03	-.18
	Faith in Intuition	.07	.20
Set 2	Future Orientation	-.27	-.15
	Zuckerman Disinhibition	-.14	.08
	Barratt Impulsivity	-.13	.16
	Sensation Seeking	-.04	-.19
Set 3	Health Risk Taking	.18	-.21
	Antisocial Risk Taking	-.17	.04
	Risk Perception	.03	-.06

Note. †p < .05 uncorrected.

Discussion

Table 13

Summary of hypotheses and data

Aim	Hypothesis	Supported by data?
1 D-E Gap by Age	a. Similar risk taking between adolescents and adults in DFD	Yes
	b. Higher risk taking in adolescents in DFE	No
	c. Adolescents do not conform to PT weighting in DFD	Yes
	d. Different processing in DFD between adolescents and adults	Yes
	e. Heightened underweighting of choices in DFE in adolescents	Maybe
	f. Similar or smaller D-E gap in adolescents relative to adults	Yes
	g. Fewer samples in adolescents relative to adults	Yes
	h. No differences in estimation of sampling distributions	Yes (if anything, adolescents are better)
2 Physiological Arousal	a. Heightened physiological arousal in adolescents in DFE	Yes
	b. No differences between adolescents' and adults' arousal levels in DFD	Yes
	c. More exaggerated underweighting of rare outcomes in mixed-domain DFE	No
	d. Higher physiological arousal in mixed- vs. single-domain DFD	No
	e. Highest arousal in adolescents in DFE mixed-domain block	No
3 Individual Differences	Correlation between behavioral inhibition measure and D-E gap	No

In this project, I tested the Spectrum of Affect model (*Figure 1*), which was proposed to expand upon existing theoretical accounts of adolescent risk taking. I additionally considered how adolescent risk taking relates to PT in DFD and DFE, and explored the possibility that adolescent risk taking is driven by heightened underweighting tendencies in DFE when the rare outcome is unfavorable. I addressed two primary aims and one secondary aim. Aims and related hypotheses, along with a summary of the data's support (or lack thereof) for those hypotheses, are outlined in *Table 13*.

Aim 1: D-E Gap by Age

For my first aim, I focused on age differences in the Description-Experience (D-E) gap paradigm. Adolescent and adult participants made the same 24 choices based on description

(Decisions from Description, DFD) and experience (Decisions from Experience, DFE). Consistent with the Spectrum of Affect model, I found that adolescents and adults exhibited similar risk taking in DFD (*Hypothesis 1.a.*). Although both adolescents and adults took more risks in DFE relative to DFD, I did not find overall heightened risk taking in adolescents compared to adults within DFE, contrary to *Hypothesis 1.b*. There are many possible explanations for this result. According to the Spectrum of Affect model, DFE paradigms without an additional affective component, may not elicit enough arousal to heighten overall risk taking. In my discussion of Aim 2 hypotheses and in the general discussion, I will return to this point. Another possibility is that the relative weighting of experience-based information in adolescents led to *less* risk taking when the rare outcome was favorable, leading to overall risk-neutral behavior in both adolescents and adults during DFE. Indeed, overall risk taking did not significantly differ from chance within DFE. I will return to this idea in discussion of *Hypothesis 1.e.*

Consistent with *Hypothesis 1.c.* and past studies investigating adolescents' probability weighting functions (Engelmann et al., 2012; Levin et al., 2014), adolescents decided quite differently than adults in DFD. My regression analysis revealed that adolescents were sensitive to the difference in expected value (EV) of the alternatives, although this effect was relatively weak and only emerged when all problem parameter moderators were included in the model. Adults, who were also sensitive to EV, systematically made decisions based on additional problem parameters in DFD. Adults exhibited choice behavior as predicted by PT: they took more risks when the rare outcome was favorable than unfavorable, and in loss problems relative to gain problems (consistent with the framing effect; Tversky & Kahneman, 1981, 1992). Contrasting the adolescent and adult data suggests that adolescents do not behave in accordance with PT predictions, and may not be sensitive to changes in problem parameters when they learn about decision alternatives through descriptions.

Eye tracking data also showed age differences in decision processes during DFD (in support of *Hypothesis 1.d.*). The most consistent age difference across problem types was that, whereas adolescents generally spent more time looking at probabilities (especially the common probability)

than values, adults tended to look more at values than probabilities. Adults' heightened attention to values may play a large role in their PT-consistent decision tendencies, and conversely, adolescents' lack of focus on values in DFD could reflect their immaturity in exhibiting the same patterns. Interestingly, in the analyses using gaze time to predict decision making, gaze time at two of the three probability AOIs (common and safe), as well as the common value, predicted adolescent decision making (such that looking more at the common probability or value predicts more risk taking, and looking more at the safe value predicts more safe choices). In adults, results were in the same direction as in adolescents, but gaze times within the risky choice display (the common or rare probabilities or values) did not predict risk taking. Only adults' gaze time on the safe probability and value predicted choices.

The fact that gaze time within the risky alternative AOIs did not predict adults' risk taking is puzzling because adults' decisions depended heavily on varying parameters in the risky alternative (e.g., whether the rare outcome is favorable or unfavorable). Conversely, while adolescents seemed less sensitive to problem parameter variation, their gaze patterns within the risky alternative in part predicted risk taking. One limitation of my gaze time analysis methods is that they are unable to detect the *sequence* of processing. In different but related decision making paradigms, processing sequence has been shown to strongly predict decision making (Venkatraman, Payne, & Huettel, 2014), and has been shown to vary significantly between adolescents and adults (Kwak, Payne, Cohen, & Huettel, 2015). Future analyses of these data will focus on indices of sequence of fixations, and will help to elucidate developmental differences in DFD processing. Putting the DFD behavior and eye tracking results together, it is clear that adults decide consistently with PT, while adolescents do not. These differences in choice behavior are likely due to processing differences (e.g., adolescents' inability to adequately integrate probabilities and values in DFD).

Adolescents and adults made similar choices in DFE, which was particularly clear in the analysis including all trials, even those where the rare outcome was not observed. In these trials, both adolescents and adults were considerably more likely to take a risk when the rare outcome was

unfavorable, consistent with underweighting of rare outcomes. Interestingly, while adults were significantly more likely to take a risk in loss vs. gain problems during DFD, this pattern was not observed in DFE. Like adolescents, adults' decisions were not significantly influenced by domain in DFE. This result is another indication that adult DFE choices deviate substantially from PT predictions.

Critically, the results provided some evidence supporting the hypothesis that adolescents show enhanced underweighting of rare outcomes relative to adults in DFE (*Hypothesis 1.e.*), even after accounting for sampling bias. First, in the mixed-domain problems, choices in trials where all outcomes were observed (all-outcome trials) were consistent with underweighting rare outcomes in DFE for participants with a high D-E gap, but only in the adolescent group. Moreover, in the regression analysis limited to all-outcome trials, adolescents were *still* more likely to take a risk when the rare outcome was unfavorable. This pattern was not significant in adults, suggesting adolescents may be more tolerant to observing rare unfavorable outcomes when they are experienced, and consistent with an overall rare-outcome underweighting bias.

Altogether, DFD and DFE choice results converged to show a slightly smaller D-E gap in adolescents relative to adults (compatible with *Hypothesis 1.f.*), driven by overall similar (or enhanced) DFE behavior, but decision making in DFD that did not correspond to PT predictions.

Further, although adolescents tended to take fewer samples per decision problem (consistent with *Hypothesis 1.g.*), the pattern of DFE decision behavior could not be explained by overall age differences in sampling bias. Neither sampling bias nor the number of all-outcome trials significantly differed between age groups. It is also unlikely that potential differences in weighting during DFE were due to differences in memory or estimation abilities during sampling (in accord with *Hypothesis 1.h.*). In the shapes estimation task, adolescents performed at least as well as adults (and for problems when the rare outcome comprised 10% of the distribution, there was evidence that adolescents were better at estimation than adults). In sum, adolescents seem to be sensitive to manipulation of problem parameters when they learn about alternatives through direct experience.

Aim 2: Physiological Arousal

A key component of the Spectrum of Affect model is that age differences in decision behavior (especially during DFE) are due to heightened affective arousal. Consistent with this idea (and *Hypothesis 2.a.*), the change in adolescents' low-frequency heart rate variability (LFHRV) from baseline was higher during DFE than during DFD. In agreement with *Hypothesis 2.b.*, there was no age difference in DFD LFHRV change from baseline. Although the relative differences in LFHRV signal by age were consistent with my hypotheses, my expectation was that adults would still show increased LFHRV in DFE relative to DFD, while adolescents would show the same pattern, but enhanced. This prediction was based on JDM literature suggesting that experience-based tasks are more affectively arousing in adults (e.g., Bechara et al., 1997; Figner et al., 2009). Surprisingly, I found the opposite pattern: adults showed *lower* LFHRV signal relative to baseline during DFE compared to DFD. There is reason to question the LFHRV measure as an index of affective arousal, as the physiological difference was not significantly correlated with behavioral indices within DFE in adolescents. While there may have been a general heightened arousal within adolescents during DFE, regardless of risk taking, these patterns are perplexing, and should be tested further in other studies using multiple measures of affective arousal.

Yet another hypothesis, stemming from two JDM studies (Hochman & Yechiam, 2010; Yechiam & Telpaz, 2011), was that problems involving both gain and loss outcomes (mixed-domain problems) would evoke more affective arousal than single-domain problems in DFD for both age groups (*Hypothesis 2.d.*), and in DFE especially in adolescents (*Hypothesis 2.e.*). Correspondingly, I hypothesized that mixed-domain problems would lead to heightened behavioral biases towards underweighting rare outcomes in DFE and overweighting them in DFD (*Hypothesis 2.c.*). None of my hypotheses regarding mixed- vs. single-domain problems were supported by the data, although eye-gaze patterns were slightly different between the single vs. mixed problem types. The most apparent difference in gaze times between the two domains was that in mixed-domain problems, gaze time was longer on the rare value than on the common or safe value in both adolescents and adults, while

the opposite was true in single-domain problems. This difference may stem from the fact that in mixed-domain problems, the common and rare values are of opposite valence (e.g., one outcome is a possible gain, while the other is a possible loss). The contrast between these two values likely attracts more attention and may represent heightened conflict between the two outcomes relative to single-domain problems, where one outcome matches the valence of the safe outcome, and the other is 0.

Although there is an interpretable difference in gaze patterns by domain, using mixed- relative to single-domain problems did not produce meaningful behavioral changes in DFD or DFE risk taking, contrary to my hypotheses. However, there were notable differences between my study and the studies motivating the problem type manipulation. First, these studies were run only in adults, did not contain rare outcomes (only 50/50 gambles), and used pupil dilation along with basic heart rate measurement rather than LFHRV (heart rate alone is not thought to be a useful measure of sympathetic nervous system activity; Kiyono et al., 2017; Tarvainen et al., 2013). Therefore, it seems probable after examining the results that this manipulation was not an appropriate one to increase affective arousal in my sample.

Aim 3: Individual Differences

I additionally explored the idea that impulsivity might be an underlying mechanism of the D-E gap (*Hypothesis 3*), based on the findings from adults in Experiment 1. I did not find evidence of a relationship between the D-E gap measure and motor or choice impulsivity as indexed by Go/No-Go or Delay Discounting, respectively. I also ran purely exploratory correlations between various self-report measures and the D-E gap. The avoidant measure on the Generalized Decision Making Scale (GDMS; Scott & Bruce, 1995), which is related to postponing or avoiding decision making when possible, was positively correlated with the D-E gap measure in both adolescents and adults. The link between avoidant decision making and the D-E gap is not immediately clear. In the past, this measure has been negatively correlated with rational decision making measures (Galotti et al., 2006; Scott & Bruce, 1995). Rational decision making could be related to the D-E gap, in that its traditional

definitions decision making involves decision invariance regardless of context (Tversky, Sattath, & Slovic, 1988; Tversky & Simonson, 1993), while the D-E gap is a measure of the effect of context on choice. However, if this were the mechanism for the observed correlation, one would think that the GDMS rational subscale would negatively correlate with the D-E gap, and this is not the case. In any case, it is important not to overinterpret this relationship, as the correlations did not survive correction. More importantly, it seems that task behavior did not relate to self-reported real-world decision behaviors. This study is hardly the first to not find a correlation between self-report questionnaires and decision behavior in the laboratory, and some have acknowledged that self-report measures have validity issues, especially in adolescent samples (Cornell, Lovegrove, & Baly, 2014; Fan et al., 2006). Future studies should work to better understand the relationship between risk taking task behavior and adolescents' real-world risk taking.

CHAPTER 4

GENERAL DISCUSSION

Adolescent risk taking is an important area of study, particularly in understanding mechanisms of heightened risk taking in order to improve intervention efforts. In this dissertation, I proposed that distinguishing between descriptive and experienced information may be helpful in understanding adolescents' heightened risk taking. I first discussed the Judgment and Decision Making (JDM) Description-Experience (D-E) gap literature, which differentiates contexts in where information about risk is learned through description versus experience. Next, I showed that this distinction can also be helpful in framing the adolescent literature, such that teens are more likely to take risks when they learn about risk alternatives through direct experience than through descriptions. This finding supported the Spectrum of Affect model, which suggests that paradigms that evoke heightened risk taking in the laboratory in adolescents are more affectively arousing than those that fail to show such a difference, and that descriptions from experience (DFE) paradigms are more arousing than decisions from description (DFD) paradigms.

The goal of the present dissertation was to test the Spectrum of Affect model. In order to do this, it was important to test both DFD and DFE within the same participant (in contrast to prior D-E gap literature, which largely employs between-subjects paradigms). Experiment 1 presented a novel paradigm that asked participants to make the same choices from description and experience, and quantified a D-E gap within each participant. Next, Experiment 2 tested adolescents and adults in the D-E gap paradigm in order to test whether adolescents were biased towards heightened risk taking in DFE, particularly when a rare outcome is unfavorable. I found that adolescents and adults took similar risks in DFD and DFE, but adolescents showed immature decision making relative to adults in DFD, accompanied by more gaze time on probabilities relative to outcomes in a decision problem. Additionally, I found evidence of heightened rare-outcome underweighting in DFE, which coincided with higher Low-Frequency Heart Rate Variability (LFHRV) in DFE relative to DFD and relative to adults.

Relating Results from Experiments 1 and 2

Experiments 1 and 2 used similar methods, yet some results were inconsistent between the two experiments. One clear behavioral difference was that the mean number of samples taken in DFE was around 9 in Experiment 1, while in Experiment 2, adolescents sampled about 11 times per trial, and adults sampled about 13 times per trial. It is possible that the difference in the number of samples taken was due to a difference in the testing environment: Experiment 1 was run in a group environment (multiple participants run at once, sitting at a conference table) while Experiment 2 was run one participant at a time. Past research from my lab has shown that task behavior can be influenced by the presence (or perceived presence) of other individuals (Chein et al., 2011; Gardner & Steinberg, 2005; O'Brien et al., 2011; Smith et al., 2014; Weigard et al., 2014). Despite efforts to sit participants far from one another in Experiment 1, participants may have inadvertently influenced one another (e.g., one participant may have seen another participant moving faster through the experiment, causing them to limit sampling). Additionally in Experiment 1, participants read instructions and advanced through the experiment on their own, while in Experiment 2, the experimenter read all instructions aloud to ensure understanding. It is possible that limited sampling was due to a lack of understanding. Indeed, several participants in the Experiment 1 experiment were excluded for not sampling at least once from each alternative per trial in DFE; none of the present experiment's participants needed to be excluded for this reason. Additionally in Experiment 1, there was a significant correlation between motor impulsivity and the within-subject D-E gap. This pattern was not observed in either age group in Experiment 2. It is possible that Experiment 2 was underpowered to detect a relationship between these constructs. Future studies should more closely examine the relationship between impulsivity and decisions from description vs. experience.

Even with experimental differences and small inconsistencies in results, there were many similar findings between the adult samples in the two studies. Within the shared problems across the two experiments, the direction and size of the D-E gap was similar. Additionally, after

accounting for sampling bias, those with a higher D-E gap in both experiments still showed evidence of rare-outcome underweighting. This effect was larger in Experiment 1, perhaps because the experiment tested more participants (*N* = 90, vs. 30 per age group in Experiment 2). Therefore, the within-subjects D-E gap paradigm and metric can serve as a useful tool for indexing individual differences in the D-E gap in future studies. However, the paradigm may be most useful in large samples of adults, where both DFD and DFE biases are pronounced and can be related to other decision traits.

Results in Relation to Past Adolescent Risk Taking Literature and Prospect Theory

DFD Results

Adolescents' overall risk taking in DFD was similar to that in adults. This result is consistent with a review of age differences observed in DFD- and DFE-like paradigms (Rosenbaum et al., Resubmitted; paraphrased in Chapter 1), showing that DFD paradigms are less likely than DFE paradigms to uncover age differences in risk taking. Interestingly, though, when DFD results were re-examined from the perspective of PT, adolescents' choices were quite different from adults'. Consistent with the idea that adolescents and adults could show substantial differences in processing of DFD information, eye-tracking results showed significant differences by age group. While complex to interpret, the fact that adolescents looked more at probabilities than values in general, and that the opposite pattern was seen in adults, may suggest a general difficulty integrating probabilities and values during adolescence, thus affecting adolescents' probability weighting function.

DFE Results

In my previous review of the developmental literature, I found that most experience-based tasks evoke heightened risk taking in adolescents relative to adults. While the DFE task showed (marginally) higher risk taking across age groups relative to DFD, I did not find an age difference in risk taking. While inconsistent with my hypothesis regarding overall risk taking, results showed the characteristic underweighting of rare outcomes within the adolescent and adult groups, as

predicted by the D-E gap literature (Hertwig et al., 2004). In addressing the relative weighting of rare outcomes, and to draw parallels between the JDM D-E gap literature and the adolescent risk taking literature, the DFE paradigm was different from most of those that showed age differences in risk taking in my prior review (c.f. van den Bos & Hertwig, 2017; discussed in the next section). Specifically, the present studies involved free sampling without consequence before making a final consequential choice, and did not display final decision feedback. In contrast, DFE studies in the developmental literature tend to involve repeated consequential decisions with feedback after each choice. It is possible that these task components are essential to showing heightened risk taking in adolescent laboratory task behavior.

Results in Relation to van den Bos and Hertwig (2017)

Experiment 2 was similar to the recent study by van den Bos and Hertwig (2017; henceforth referred to as V&H). There were, however, several differences in the designs of the two studies. V&H asked participants to make 10 DFE choices, and only 8 of those included an alternative with a rare outcome (participants made those same choices in DFD, among many other choices without a rare event that did not correspond to DFE choices, and with some ambiguous decisions interspersed). Additionally, V&H's choice problems were limited to single-domain problems which, in past studies of adult choices, have shown a large D-E gap (e.g., Hertwig et al., 2004). Using few decision problems that have been shown to elicit strong biases in the past may limit the generalizability of presented findings (Camilleri & Newell, 2013b). In contrast, Experiment 2 asked participants to make the same 24 choices in DFD and DFE, many of which were novel, and half of which were mixed-domain problems. Beyond these differences in the decision paradigms themselves, my study included additional behavioral measures not included in V&H (shapes estimation, Go/No-Go, Delay Discounting) and physiological/eye tracking measures.

Even considering these paradigm differences, the two studies converged to find lower sampling in adolescents relative to adults during DFE. Crucially, both showed evidence for

heightened underweighting of rare outcomes in DFE vs. DFD in adolescents, a pattern that persisted in trials where the rare outcome was observed. There were several results that diverged in the two studies. Interestingly, despite similar sampling patterns across the two studies, participants in Experiment 2 experienced the rare outcome more often than in V&H (around 60% across all participants in the present Experiment 2, vs. 38% in V&H). This was likely due to the fact that adolescents sampled far less in van den Bos and Hertwig (around 5 times) versus in the present study, where they sampled around 11 times.

Also interestingly, although V&H (2017) did not present DFD results in terms of PT biases (i.e., whether adolescents behaved as if they overweighted rare events in DFD), adolescents showed a more pronounced D-E gap than adults, in contrast to adolescents' lower D-E gap in Experiment 2. This suggests that adolescents likely showed similar PT biases to adults in their study. The inconsistency in PT findings by age may be due to a number of factors (e.g., culture – German vs. U.S. sample, gender distribution). The developmental trajectory of weighting functions in DFD and DFE is an interesting topic for future research.

Reinforcement Learning, the Ventral Striatum, and DFE

Although adolescents did not take more risks than adults in the present DFE task, their adult-like (and even potentially enhanced) rare-outcome underweighting in DFE is intriguing, particularly with evidence that adolescents showed a higher change in LFHRV from baseline relative to adults in this task. However, perhaps the most interesting finding comes from the contrast between DFD and DFE: adolescents were able to pick up on small variations in decision problems similarly to adults in DFE, but were unable to learn similarly from descriptions. This result shows a natural parallel to a different literature: reinforcement learning. Reinforcement learning tasks involve repeated choices, and updating choices based on feedback. There is a growing literature on reinforcement learning in developmental cognitive neuroscience. Prediction error from reinforcement learning (i.e., the error between a predicted outcome based on past experiences and the actual observed decision outcome) has been tied to activity in the ventral

striatum (Schultz, Dayan, & Montague, 1997), a reward-related area thought to be particularly relevant to adolescent risk taking (Casey, 2015; Shulman et al., 2016). Cohen and colleagues (2010) demonstrated enhanced ventral striatal reward prediction error signal in adolescents relative to children and adults. Another recent study showed enhanced reinforcement learning in adolescents, and enhanced memory for instances when decision outcomes were rewarded (Davidow, Foerde, Galvan, & Shohamy, 2016).

In addition to its relevance to reward prediction error, ventral striatum signal has been linked to affective arousal (Knutson et al., 2014), and more specifically to reward-related approach in adolescence (Casey, 2015; Ernst, 2014). Consistent with the dual systems model (Shulman et al., 2016) and the Spectrum of Affect model advanced earlier, arousal is thought to be the mechanism through which heightened reward seeking occurs. Together, these findings suggest that adolescence is a period during which the ventral striatum is integrally involved in both reward learning and risk taking, two concepts that are likely inextricably linked (Davidow et al., 2016). Reward prediction error itself may enhance adolescents' ability to learn in DFE but not in DFD. Consistent with this logic, adolescents performed slightly better than adults at the shapes estimation task, and may have found the task easier (rather than harder) than adults.

Relating Developmental Results to Studies of Description+Experience

I found small age differences in DFE, in the presence of immature DFD behavior within adolescents. This result is consistent with the idea that adolescents are unable to integrate descriptive information. At the same time, they are able to learn information from experience just as well as, if not better than, adults. A relevant question is how people incorporate descriptions and experiences when both types of information are available, as they often are in the real world. Another important question regards the developmental trajectory of how competing information sources are integrated. In my earlier discussion of the D-E gap literature (Chapter 1), I found that paradigms giving adult participants both descriptions and experiences revealed an experience-bias over time, even when descriptive information was available (Jessup et al., 2008; Lejarraga &

Gonzalez, 2011; Yechiam et al., 2005). This pattern of behavior is consistent with the speeding example mentioned earlier, in that adults tend to speed, underweighting the likelihood that they will encounter a rare possibility of getting a ticket, perhaps due to repeated experiences with speeding. Importantly, this behavior occurs despite descriptive knowledge that speeding can lead to tickets. Based on the present study, adolescents are better at learning from experience than from description. It is plausible that when both descriptions and experiences are available, adolescents are even *less* able to integrate descriptions into their choices relative to adults, leading to underweighting of rare unfavorable outcomes with less experience than adults. Future studies would benefit from using paradigms that involve repeated consequential decisions along with available descriptions to address this question.

Study Limitations and Future Directions

This study is novel in that it tested the D-E gap using behavioral and physiological methods from the JDM literature in a developmental sample, in an attempt to bridge the disconnect between adolescent and adult studies of risk taking. It additionally tested the Spectrum of Affect model, which stemmed from an integration of JDM and adolescent risk taking literatures. Although the data did not provide clear evidence supporting some of my hypotheses, it is important to address study limitations and to expand on current findings in future studies. Further, the rich dataset collected in the present study will allow for novel analyses in the future.

A major limitation of this study was that skin conductance data were not usable. Skin conductance level (SCL) is often used as an index of affective arousal. While I collected these data, it is unclear why SCL values were lower during the tasks than during baseline periods. If skin conductance data had been usable, finer-grained analyses of problem-level skin conductance responses (SCRs) would have been possible. Future analyses will attempt to remedy the data to render them useful.

As discussed earlier, new analyses of eye tracking data are forthcoming, involving the order of decision processes. In my eye tracking data collection, I also collected pupil dilation data,

which is thought to be another index of arousal (Joshi, Li, Kalwani, & Gold, 2016). Upcoming analyses will examine pupil dilation as it relates to various problem dimensions in DFD.

Another exciting future direction will involve exploring reinforcement learning modeling of sampled outcomes in DFE. Looking further into the sequence of experienced values will help us to better understand what problem parameters predict choice, and whether such parameters differ between adolescents and adults.

Additionally, it seemed clear that the problem type manipulation meant to increase affective arousal did not work as planned. In the future, it may be fruitful to run the D-E gap paradigm using a social context manipulation, which has been shown to increase adolescent risk taking in both description based paradigms (Haddad et al., 2014; Smith et al., 2014) and experience-based paradigms (Chein et al., 2011; Gardner & Steinberg, 2005).

The present analysis also has implications for understanding adolescents' choices in the real world and in implementing interventions designed to discourage adolescent risk taking. Obviously, there is a downside to learning about many risks through experience; a rare but disadvantageous outcome can have serious consequences. However, because experience-based tasks are more likely to evince heightened risk taking in adolescents, they therefore may be more useful in understanding the mechanisms behind adolescents' risk taking through laboratory studies. Another implication of my review is that interventions based on information alone are unlikely to be effective in deterring the sorts of risk taking for which adolescents' prior experiences have not produced negative outcomes. It is important to note, therefore, emerging evidence that virtual reality-based interventions can translate into long-lasting behavioral change in the real world (Ahn, Bailenson, & Park, 2014; Bailey et al., 2015). Accordingly, one might consider attempts to simulate certain risk-taking environments in order to give teens virtual "experiences," especially negative experiences, with risk taking in a safe, controlled environment. For example, an effective intervention against drinking and driving might use a driving simulator to demonstrate to a teen the difficulty of stopping in time for a red light while intoxicated. Yet another

possibility is the use of information-based interventions mainly with younger adolescents, before many of them have had experiences that have not led to negative outcomes. That is, it may be easier to communicate facts about safe sex to a sexually inexperienced 12-year-old than to a 17-year-old who has been having unprotected intercourse for a year but has not experienced a pregnancy or STI.

Finally, although teens take more risks than adults in general, there is considerable individual variability in risk taking among same-aged adolescents (Bjork & Pardini, 2014). Risk-taking interventions may be most effective if they target those teenagers who are prone to heightened risk taking, rather than those who are unlikely to take risks in the first place. Indeed, Experiment 2 attempted to explain some variability in task behavior by using a within-subjects paradigm and by running a series of correlation analyses between task behavior and a variety of self-report measures. However, I did not find any correlations that survived correction, and realistically, the sample size in Experiment 2 was too small to uncover such effects. Future studies would benefit from implementing within-subjects designs with larger sample sizes to improve understanding of individual differences in risk taking.

REFERENCES CITED

Abdellaoui, M., L'Haridon, O., & Paraschiv, C. (2011). Experienced vs. Described Uncertainty: Do We Need Two Prospect Theory Specifications? *Management Science*, *57*(10), 1879–1895. http://doi.org/10.1287/mnsc.1110.1368

Ahn, S. J. (Grace), Bailenson, J. N., & Park, D. (2014). Short- and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. *Computers in Human Behavior*, *39*, 235–245. http://doi.org/10.1016/j.chb.2014.07.025

Aïte, A., Cassotti, M., Rossi, S., Poirel, N., Lubin, A., Houdé, O., & Moutier, S. (2012). Is human decision making under ambiguity guided by loss frequency regardless of the costs? A developmental study using the Soochow Gambling Task. *Journal of Experimental Child Psychology*, *113*(2), 286–94. http://doi.org/10.1016/j.jecp.2012.05.008

Albert, D., & Steinberg, L. (2011). Judgment and Decision Making in Adolescence. *Journal of Research on Adolescence*, *21*(1), 211–224. http://doi.org/10.1111/j.1532-7795.2010.00724.x

Andre, J., Picchioni, M., Zhang, R., & Toulopoulou, T. (2015). Working memory circuit as a function of increasing age in healthy adolescence: A systematic review and meta-analyses. *NeuroImage: Clinical*. http://doi.org/10.1016/j.nicl.2015.12.002

Bach, D. R., & Friston, K. J. (2013). Model-based analysis of skin conductance responses: Towards causal models in psychophysiology. *Psychophysiology*, *50*, 15–22. http://doi.org/10.1111/j.1469-8986.2012.01483.x

Bailey, J. O., Bailenson, J. N., Flora, J., Armel, K. C., Voelker, D., & Reeves, B. (2015). The Impact of Vivid Messages on Reducing Energy Consumption Related to Hot Water Use. *Environment and Behavior*, *47*(5), 570–592. http://doi.org/10.1177/0013916514551604

Barkley-Levenson, E. E., & Galván, A. (2014). Neural representation of expected value in the adolescent brain. *Proceedings of the National Academy of Sciences of the United States of*

America, *111*, 1646–51. http://doi.org/10.1073/pnas.1319762111

Barkley-Levenson, E. E., Van Leijenhorst, L., & Galván, A. (2013). Behavioral and neural correlates of loss aversion and risk avoidance in adolescents and adults. *Developmental Cognitive Neuroscience*, *3*, 72–83. http://doi.org/10.1016/j.dcn.2012.09.007

Barron, G., Leider, S., & Stack, J. (2008). The effect of safe experience on a warnings' impact: Sex, drugs, and rock-n-roll. *Organizational Behavior and Human Decision Processes*, *106*(2), 125–142. http://doi.org/10.1016/j.obhdp.2007.11.002

Barron, G., & Yechiam, E. (2009). The coexistence of overestimation and underweighting of rare events and the contingent recency effect. *Judgment and Decision Making*, *4*(6), 447–460.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 51. http://doi.org/10.18637/jss.v067.i01

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8039375

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science (New York, N.Y.)*, *275*(5304), 1293–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9036851

Beitz, K. M., Salthouse, T. A., & Davis, H. P. (2014). Performance on the Iowa Gambling Task: From 5 to 89 Year. *Journal of Experimental Psychology: General*, *143*(3), 1677–1689. http://doi.org/10.1037/a0035823

Benthin, A., Slovic, P., & Severson, H. (1993). A Psychometric study of adolescent risk perception. *Journal of Adolescence*, *16*, 153–168.

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., … van der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1997.tb02140.x

Bjork, J. M., & Pardini, D. A. (2014). Who are those "risk-taking adolescents"? Individual

differences in developmental neuroimaging research. *Developmental Cognitive Neuroscience*. http://doi.org/10.1016/j.dcn.2014.07.008

Bjork, J. M., Smith, A. R., Danube, C. L., & Hommer, D. W. (2007). Developmental differences in posterior mesofrontal cortex recruitment by risky rewards. *The Journal of Neuroscience*, *27*(18), 4839–4849. http://doi.org/10.1523/JNEUROSCI.5469-06.2007

Blankenstein, N. E., Crone, E. A., van den Bos, W., & van Duijvenvoorde, A. C. K. (2016). Dealing With Uncertainty: Testing Risk- and Ambiguity-Attitude Across Adolescence. *Developmental Neuropsychology*, *5641*(March), 1–16. http://doi.org/10.1080/87565641.2016.1158265

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. http://doi.org/10.1111/j.1469-8986.2012.01384.x

Braams, B. R., van Duijvenvoorde, A. C. K., Peper, J. S., & Crone, E. A. (2015). Longitudinal Changes in Adolescent Risk-Taking: A Comprehensive Study of Neural Responses to Rewards, Pubertal Development, and Risk-Taking Behavior. *Journal of Neuroscience*, *35*(18), 7226–7238. http://doi.org/10.1523/JNEUROSCI.4764-14.2015

Brooks, P., & Zank, H. (2005). Loss averse behavior. *Journal of Risk and Uncertainty*, *31*(3), 301–325. http://doi.org/10.1007/s11166-005-5105-7

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology*, *17*(2), 243–250. http://doi.org/10.1016/j.conb.2007.02.005

Burnett, S., Bault, N., Coricelli, G., & Blakemore, S. J. (2010). Adolescents' heightened risk-seeking in a probabilistic gambling task. *Cognitive Development*, *25*(2010), 183–196. http://doi.org/10.1016/j.cogdev.2009.11.003

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2008). The Psychophysiology of Emotion. In *Handbook of Emotions, 3rd Edition* (pp. 173–191).

Camilleri, A. R., & Newell, B. R. (2009). Within-subject Preference Reversals in Description- and Experience-based Choice. *Cognitive Science Society*, 449–454.

Camilleri, A. R., & Newell, B. R. (2011a). Description- and experience-based choice: does equivalent information equal equivalent choice? *Acta Psychologica*, *136*(3), 276–84. http://doi.org/10.1016/j.actpsy.2010.11.007

Camilleri, A. R., & Newell, B. R. (2011b). When and why rare events are underweighted: a direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review*, *18*(2), 377–84. http://doi.org/10.3758/s13423-010-0040-2

Camilleri, A. R., & Newell, B. R. (2013a). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In *Progress in brain research* (1st ed., Vol. 202, pp. 55–71). Elsevier B.V. http://doi.org/10.1016/B978-0-444-62604-2.00004-6

Camilleri, A. R., & Newell, B. R. (2013b). The long and short of it: Closing the description-experience "'gap'" by taking the long-run view. *Cognition*, *126*, 54–71.

Casey, B. J. (2015). Beyond Simple Models of Self-Control to Circuit-Based Accounts of Adolescent Behavior. *Annual Review of Psychology*, *66*(1), 295–319. http://doi.org/10.1146/annurev-psych-010814-015156

Cassotti, M., Houdé, O., & Moutier, S. (2011). Developmental changes of win-stay and loss-shift strategies in decision making. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *17*(4), 400–411. http://doi.org/10.1080/09297049.2010.547463

Caswell, A. J., Bond, R., Duka, T., & Morgan, M. J. (2015). Further evidence of the heterogeneous nature of impulsivity. *Personality and Individual Differences*, *76*, 68–74. http://doi.org/10.1016/j.paid.2014.11.059

Cauffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., & Woolard, J. (2010). Age differences in affective decision making as indexed by performance on the Iowa

Gambling Task. *Developmental Psychology*, *46*(1), 193–207. http://doi.org/10.1037/a0016128

Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science*, *14*(2), F1-10. http://doi.org/10.1111/j.1467-7687.2010.01035.x

Clayton, R., Cattarello, A., & Johnstone, B. (1996). The Effectiveness of Drug Abuse Resistance Education (Project. *Preventative Medicine*, *31825*, 307–318. http://doi.org/10.1006/pmed.1996.0061

Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience*, *13*(6), 669–71. http://doi.org/10.1038/nn.2558

Cornell, D. G., Lovegrove, P. J., & Baly, M. W. (2014). Invalid survey response patterns among middle school students. *Psychological Assessment*, *26*(1), 277–87. http://doi.org/10.1037/a0034808

Crone, E. A., & van der Molen, M. W. (2004). Developmental changes in real life decision making: performance on a gambling task previously shown to depend on the ventromedial prefrontal cortex. *Developmental Neuropsychology*, *25*(3), 251–79. http://doi.org/10.1207/s15326942dn2503_2

Crone, E. A., Vendel, I., & van der Molen, M. W. (2003). Decision-making in disinhibited adolescents and adults: Insensitivity to future consequences or driven by immediate reward? *Personality and Individual Differences*, *35*, 1625–1641. http://doi.org/10.1016/S0191-8869(02)00386-0

Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(24), 9315–9320. http://doi.org/10.1073/pnas.0510088103

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1346), 1413–20. http://doi.org/10.1098/rstb.1996.0125

Davidow, J. Y., Foerde, K., Galvan, A., & Shohamy, D. (2016). An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence. *Neuron*, *92*(1), 93–99. http://doi.org/10.1016/j.neuron.2016.08.031

de Palma, A., Abdellaoui, M., Attanasi, G., Ben-Akiva, M., Erev, I., Fehr-Duda, H., … Weber, M. (2014). Beware of black swans: Taking stock of the description–experience gap in decision under uncertainty. *Marketing Letters*, *25*(3), 269–280. http://doi.org/10.1007/s11002-014-9316-z

Defoe, I. N., Dubas, J. S., Figner, B., & Van Aken, M. A. G. (2015). A Meta-analysis on Age Differences in Risky Decision Making : Adolescents versus Children and Adults. *Psychological Bulletin*, *141*(1), 29.

Downs, J. S., Murray, P. J., Bruine de Bruin, W., Penrose, J., Palmgren, C., & Fischhoff, B. (2004). Interactive video behavioral intervention to reduce adolescent females' STD risk: a randomized controlled trial. *Social Science & Medicine (1982)*, *59*(8), 1561–1572. http://doi.org/10.1016/j.socscimed.2004.01.032

Duncan, S., & Barrett, L. F. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition & Emotion*, *21*(6), 1184–1211. http://doi.org/10.1080/02699930701437931

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: a critical evaluation. *Neuroscience and Biobehavioral Reviews*, *30*(2), 239–71. http://doi.org/10.1016/j.neubiorev.2005.07.001

Elkins, I. J., McGue, M., & Iacono, W. G. (2007). Prospective effects of attention-deficit/hyperactivity disorder, conduct disorder, and sex on adolescent substance use and abuse. *Archives of General Psychiatry*, *64*(10), 1145–1152.

Engelmann, J. B., Moore, S., Capra, C. M., & Berns, G. S. (2012). Differential neurobiological

effects of expert advice on risky choice in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(5), 557–67. http://doi.org/10.1093/scan/nss050

Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of project DARE outcome evaluations. *American Journal of Public Health*, *84*(9), 1394–1401. http://doi.org/10.2105/AJPH.84.9.1394

Epstein, S., Pacini, R., Denes-raj, V., & Heier, H. (1996). Individual Differences in Intuitive-Experiential and Analytical-Rational Thinking Styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405.

Ernst, M. (2014). The triadic model perspective for the study of adolescent motivated behavior. *Brain and Cognition*, *89*, 104–111. http://doi.org/10.1016/j.bandc.2014.01.006

Ernst, M., Grant, S. J., London, E. D., Contoreggi, C. S., Kimes, A. S., & Spurgeon, L. (2003). Decision Making in Adolescents With Behavior Disorders and Adults With Substance Abuse. *American Journal of Psychiatry*, *160*, 33–40. http://doi.org/10.1176/appi.ajp.160.1.33

Eshel, N., Nelson, E. E., Blair, R. J., Pine, D. S., & Ernst, M. (2007). Neural substrates of choice selection in adults and adolescents: Development of the ventrolateral prefrontal and anterior cingulate cortices. *Neuropsychologia*, *45*, 1270–1279. http://doi.org/10.1016/j.neuropsychologia.2006.10.004

Fan, X., Miller, B. C., Park, K.-E., Winward, B. W., Christiensen, M., Grotevant, H. D., & Tai, R. H. (2006). An Exploratory Study about Inaccuracy and Invalidity in Adolescent Self-Report Surveys. *Field Methods*, *18*(3), 223–244. http://doi.org/10.1177/152822X06289161

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 709–30. http://doi.org/10.1037/a0014983

Fox, C. R., & Hadar, L. (2006). “Decisions from experience” = sampling error + prospect theory.

Judgment and Decision Making, *1*(2), 159–161.

Galotti, K. M., Ciner, E., Altenbaumer, H. E., Geerts, H. J., Rupp, A., & Woulfe, J. (2006). Decision-making styles in a real-life decision: Choosing a college major. *Personality and Individual Differences*, *41*(4), 629–639. http://doi.org/10.1016/j.paid.2006.03.003

Galván, A., & McGlennen, K. M. (2012). Daily stress increases risky decision-making in adolescents: A preliminary study. *Developmental Psychobiology*, *54*, 433–440. http://doi.org/10.1002/dev.20602

Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology*, *41*(4), 625–35. http://doi.org/10.1037/0012-1649.41.4.625

Glöckner, A., Fiedler, S., Hochman, G., Ayal, S., & Hilbig, B. E. (2012). Processing differences between descriptions and experience: A comparative analysis using eye-tracking and physiological measures. *Frontiers in Psychology*, *3*(JUN), 1–15. http://doi.org/10.3389/fpsyg.2012.00173

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523–51. http://doi.org/10.1037/a0024558

Gottlieb, D. A., Weiss, T., & Chapman, G. B. (2007). The Format in Which Is Information Uncertainty Decision Biases Presented Affects. *Psychological Science*, *18*(3), 240–246.

Griffin, K. W., Botvin, G. J., & Nichols, T. R. (2006). Effects of a school-based drug abuse prevention program for adolescents on HIV risk behavior in young adulthood. *Prevention Science*, *7*(1), 103–112. http://doi.org/10.1007/s11121-006-0025-6

Habib, M., Borst, G., Poirel, N., Houde, O., Moutier, S., & Cassotti, M. (2015). Socio-emotional context and adolescents' decision making: The experience of regret and relief after social comparison. *Journal of Research on Adolescence*, *25*(1), 81–91. http://doi.org/10.1111/jora.12092

Habib, M., Cassotti, M., Borst, G., Simon, G., Pineau, A., Houdé, O., & Moutier, S. (2012). Counterfactually mediated emotions: A developmental study of regret and relief in a probabilistic gambling task. *Journal of Experimental Child Psychology*, *112*(2), 265–274. http://doi.org/10.1016/j.jecp.2012.01.007

Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*(4), 317–325.

Haddad, A. D. M., Harrison, F., Norman, T., & Lau, J. Y. F. (2014). Adolescent and adult risk-taking in virtual social contexts. *Frontiers in Psychology*, *5*(December), 1–7. http://doi.org/10.3389/fpsyg.2014.01476

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk Attitudes of Children and Adults: Choices Over Small and Large Probability Gains and Losses. *Experimental Economics*, *5*, 53–84. http://doi.org/10.1023/A:1016316725855

Hartley, C. A., & Somerville, L. H. (2015). The neuroscience of adolescent decision-making. *Current Opinion in Behavioral Sciences*, *5*, 108–115. http://doi.org/10.1016/j.cobeha.2015.09.004

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. http://doi.org/10.1002/bdm.665

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*, *518*(May), 493–518. http://doi.org/10.1002/bdm.598

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–9. http://doi.org/10.1111/j.0956-7976.2004.00715.x

Hibino, G., Moritani, T., Kawada, T., & Fushiki, T. (1997). Caffeine enhances modulation of parasympathetic nerve activity in humans: quantification using power spectral analysis. *The*

Journal of Nutrition, *127*(February 1996), 1422–1427. http://doi.org/?

Hilbig, B. E., & Glöckner, A. (2011). Yes, they can! Appropriate weighting of small probabilities as a function of information acquisition. *Acta Psychologica*, *138*(3), 390–396. http://doi.org/10.1016/j.actpsy.2011.09.005

Hochman, G., & Yechiam, E. (2010). Loss Aversion in the Eye and in the Heart: The Autonomic Nervous System's Responses to Losses. *Journal of Behavioral Decision Making*, *24*, 140–156. http://doi.org/10.1002/bdm

Humphrey, G., & Dumontheil, I. (2016). Development of Risk-Taking, Perspective-Taking, and Inhibitory Control During Adolescence. *Developmental Neuropsychology*, *5641*(May), 1–18. http://doi.org/10.1080/87565641.2016.1161764

Jemmott III, J. B., Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a Theory-Based Abstinence-Only Intervention Over 24 Months. *Archives of Pediatric and Adolescent Medicine*, *164*(2), 152–159.

Jenkins, L., Myerson, J., Hale, S., & Fry, A. F. (1999). Individual and developmental differences in working memory across the life span. *Psychonomic Bulletin & Review*, *6*(1), 28–40. http://doi.org/10.3758/BF03210810

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science*, *19*, 1015–1022. http://doi.org/10.1111/j.1467-9280.2008.02193.x

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234. http://doi.org/10.1016/j.neuron.2015.11.028

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, *47*(2), 263–292.

Kann, L., Kinchen, S., Shanklin, S. L., Flint, K. H., Kawkins, J., Harris, W. A., … Zaza, S. (2012). Youth risk behavior surveillance--United States, 2011. *Morbidity and Mortality Weekly*

Report. Surveillance Summaries (Washington, D.C. : 2002), *61*(4), 1–168. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24918634

Kareev, Y. (1995). Through a narrow window: working memory capacity and the detection of covariation. *Cognition*, *56*, 263–269. http://doi.org/10.1016/0010-0277(95)92814-G

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, *107*(2), 397–402. http://doi.org/10.1037/0033-295X.107.2.397

Keulers, E. H. H., Stiers, P., & Jolles, J. (2011). Developmental changes between ages 13 and 21 years in the extent and magnitude of the BOLD response during decision making. *NeuroImage*, *54*(2), 1442–1454. http://doi.org/10.1016/j.neuroimage.2010.08.059

Kim-Spoon, J., Kahn, R., Deater-Deckard, K., Chiu, P. H., Steinberg, L., & King-Casas, B. (2015). Risky decision making in a laboratory driving task is associated with health risk behaviors during late adolescence but not adulthood. *International Journal of Behavioural Development*, (Mc 0436), 1–25.

Kim, S., & Lee, D. (2011). Prefrontal Cortex and Impulsive Decision Making. *Biological Psychiatry*, *69*(12), 1140–1146. http://doi.org/10.1016/j.biopsych.2010.07.005.Prefrontal

Kirby, D., & Laris, B. A. (2006). Effective Curriculum-Based Sex and STD/HIV Education Programs for Adolescents. *Child Development Perspectives*, *3*(1), 21–29.

Kiyono, K., Hayano, J., & Watanabe, E. (2017). Heart Rate Variability (HRV) and Sympathetic Nerve Activity. In *Clinical Assessment of the Autonomic Nervous System* (pp. 147–161). http://doi.org/10.1007/978-4-431-56012-8

Knutson, B., Katovich, K., & Suri, G. (2014). Inferring affect from fMRI data. *Trends in Cognitive Sciences*, *18*(8), 422–428. http://doi.org/10.1016/j.tics.2014.04.006

Kühberger, A., & Tanner, C. (2009). A test of prospect theory and fuzzy-trace theory in risky choice framing. *Journal of Behavioral Decision Making*.

Kwak, Y., Payne, J. W., Cohen, A. L., & Huettel, S. A. (2015). The Rational Adolescent: Strategic Information: Processing during Decision Making Revealed by Eye Tracking. *Cognitive*

Development, *36*, 20–30. http://doi.org/10.1016/j.cogdev.2015.08.001

Lejarraga, T. (2010). When Experience Is Better Than Description: Time Delays and Complexity. *Journal of Behavioral Decision Making*, *23*, 100–116. http://doi.org/10.1002/bdm

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, *116*(2), 286–295. http://doi.org/10.1016/j.obhdp.2011.05.001

Levin, I. P., Bossard, E. A., Gaeth, G. J., & Yan, H. (2014). The combined role of task, child's age and individual differences in understanding decision processes. *Judgment and Decision Making*, *9*(3), 274–286.

Lichtenstein, S. (1965). Bases for Preferences Among Three-Outcome Bets. *Journal of Experimental Psychology*, *69*(2), 162–169.

Loomes, G. (2010). Modeling choice and valuation in decision experiments. *Psychological Review*, *117*(3), 902–924. http://doi.org/10.1037/a0019807

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The Development of Nonverbal Working Memory and Executive Control Processes in Adolescents. *Child Development*, *76*(3), 697–712.

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? *PLoS ONE*, *6*(6). http://doi.org/10.1371/journal.pone.0020262

Luna, B., Garver, K. E., Urban, T. a, Lazar, N. a, & Sweeney, J. a. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357–1372. http://doi.org/10.1111/j.1467-8624.2004.00745.x

Mailhes, C., & Castanié, F. (2013). Spectral Analysis by Parametric Modeling. In F. Castanié (Ed.), *Digital Spectral Analysis: Parametric, Non-Parametric and Advanced Methods* (pp. 145–168). Hoboken, US.

Malliani, A., Pagani, M., Lombardi, F., & Cerutti, S. (1991). Cardiovascular neural regulation

explored in the frequency domain. *Circulation*, *84*(2), 482–492. http://doi.org/10.1161/01.CIR.84.2.482

Mayo Clinic. (2014). Caffeine content for coffee, tea, soda and more. Retrieved from http://www.mayoclinic.org/healthy-lifestyle/nutrition-and-healthy-eating/in-depth/caffeine/art-20049372

Mitchell, S. H., Schoel, C., & Stevens, A. A. (2008). Mechanisms underlying heightened risk taking in adolescents as compared with adults. *Psychonomic Bulletin & Review*, *15*(2), 272–277. http://doi.org/10.3758/PBR.15.2.272

Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S.-J., & Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16396–401. http://doi.org/10.1073/pnas.1305631110

O'Brien, L., Albert, D., Chein, J., & Steinberg, L. (2011). Adolescents Prefer More Immediate Rewards When in the Presence of their Peers. *Journal of Research on Adolescence*, *21*(4), 747–753. http://doi.org/10.1111/j.1532-7795.2011.00738.x

Olson, E. A., Hooper, C. J., Collins, P., & Luciana, M. (2007). Adolescents' performance on delay and probability discounting tasks: Contributions of age, intelligence, executive functioning, and self-reported externalizing behavior. *Personality and Individual Differences*, *43*(7), 1886–1897. http://doi.org/10.1016/j.paid.2007.06.016

Overman, W. H., Frassrand, K., Ansel, S., Trawalter, S., Bies, B., & Redmond, A. (2004). Performance on the IOWA card task by adolescents and adults. *Neuropsychologia*, *42*, 1838–1851.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774.

Paulsen, D. J., Carter, R. M., Platt, M. L., Huettel, S. A., & Brannon, E. M. (2012). Neurocognitive Development of Risk Aversion from Early Childhood to Adulthood. *Frontiers in Human*

Neuroscience, *5*(January), 1–17. http://doi.org/10.3389/fnhum.2011.00178

Paulsen, D. J., Platt, M. L., Huettel, S. A., & Brannon, E. M. (2011). Decision-making under risk in children, adolescents, and young adults. *Frontiers in Psychology*, *2*(April), 72. http://doi.org/10.3389/fpsyg.2011.00072

Payne, J. W., & Venkatraman, V. (2011). Opening the Black Box: Conclusions to A Handbook of Process Tracing Methods for Decision Research. In M. Schulte-Mecklenbeck, A. Kuhberger, & R. Ranyard (Eds.), *A Handbook of Process Tracing Methods for Decision Research* (pp. 223–249). New York: Psychology Press.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

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*(2), 168–179. http://doi.org/10.1016/j.obhdp.2008.02.001

Rakow, T., & Rahim, S. B. (2010). Developmental Insights into Experience-based Decision Making. *Journal of Behavioral Decision Making*, *23*, 69–82. http://doi.org/10.1002/bdm.672

Rauh, R., Burkert, M., Seipmann, M., & Mueck-Weymann, M. (2006). Acute effects of caffeine on heart rate variability in habitual caffeine consumers. *Clinical Physiology and Functional Imaging*, *26*(2), 163–166. http://doi.org/10.1016/S0002-9149(02)02725-X

Reyes del Paso, G. A., Langewitz, W., Mulder, L. J. M., van Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. *Psychophysiology*, *50*(5), 477–487. http://doi.org/10.1111/psyp.12027

Reyna, V. F., Chick, C. F., Corbin, J. C., & Hsia, A. N. (2014). Developmental Reversals in Risky Decision Making: Intelligence Agents Show Larger Decision Biases Than College Students. *Psychological Science*, *25*(1), 76–84. http://doi.org/10.1177/0956797613497022

Reyna, V. F., Estrada, S. M., Demarinis, J. A., Myers, R. M., Stanisz, J. M., & Mills, B. A. (2011). Neurobiological and memory models of risky decision making in adolescents versus young adults. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(5), 1125–1142. http://doi.org/10.1037/a0023943

Reyna, V. F., & Farley, F. (2006). Risk and Rationality in Adolescent Decision Making: Implications for Theory, Practice and Public Policy. *Cognition*, *7*(1), 1–44. Retrieved from http://psi.sagepub.com/content/7/1/1.abstract

Reyna, V. F., & Mills, B. A. (2014). Theoretically motivated interventions for reducing sexual risk taking in adolescence: a randomized controlled experiment applying fuzzy-trace theory. *Journal of Experimental Psychology. General*, *143*(4), 1627–48. http://doi.org/10.1037/a0036717

Rosenbaum, G. M., Chein, J. M., & Venkatraman, V. (In Revision). A Within-Subjects Measure of Individual Differences in the Description-Experience Gap.

Rosenbaum, G.M., Venkatraman, V., Steinberg, L., & Chein, J.M. (Resubmitted). The Influences of Described and Experienced Information on Adolescent Risk Taking.

Scherer, K. R. (2005). What are emotion? And how can they be measured? *Social Science Information Sur Les Sciences Sociales*, *44*(4), 695–729. http://doi.org/10.1177/0539018405058216

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

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. http://doi.org/10.1126/science.275.5306.1593

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing Versus Satisficing: Happiness Is a Matter of Choice. *Journal of Personality and Social Psychology*, *83*(5), 1178–1197. http://doi.org/10.1037//0022-3514.83.5.1178

Scott, S. G., & Bruce, R. A. (1995). Decision-making style: The development and assessment of a new measure. *Educational and Psychological MEasurement*, *55*(5), 818–831.

Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*(2015), 103–117. http://doi.org/10.1016/j.dcn.2015.12.010

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. http://doi.org/10.1016/j.ejor.2005.04.006

Smith, A. R., Chein, J., & Steinberg, L. (2014). Peers increase adolescent risk taking even when the probabilities of negative outcomes are known. *Developmental Psychology*, *50*, 1564–8. http://doi.org/10.1037/a0035696

Sondermeijer, H. P., van Marle, A. G. J., Kamen, P., & Krum, H. (2002). Acute effects of caffeine on heart rate variability. *American Journal of Cardiology*, *90*(8), 906–907. http://doi.org/10.1016/S0002-9149(02)02725-X

Steinberg, L. (2008). A Social Neuroscience Perspective on Adolescent Risk-Taking. *Developmental Review*, *28*(1), 78–106. http://doi.org/10.1016/j.dr.2007.08.002

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Developmental Psychology*, *44*(6), 1764–78. http://doi.org/10.1037/a0012955

Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development*, *80*(1), 28–44. http://doi.org/10.1111/j.1467-8624.2008.01244.x

Tarvainen, M. P. (2014). *Kubios HRV User's Guide, Version 2.2*. http://doi.org/10.1016/j.enconman.2005.10.016

Tarvainen, M. P., Niskanen, J., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2013).

Kubios HRV – Heart rate variability. *Computer Methods and Programs in Biomedicine*, *113*(1), 210–220. http://doi.org/10.1016/j.cmpb.2013.07.024

Tarvainen, M. P., Ranta-aho, P. O., & Karjalainen, P. A. (2002). An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering*, *49*(2), 172–175. http://doi.org/10.1109/10.979357

Task Force. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of The European Society of Cardiology and the North American Society for Pacing and Electrophysiology. *Eur Heart J.*, *17*(3), 354–381. http://doi.org/10.1161/01.CIR.93.5.1043

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology*, *2*(MAR), 1–9. http://doi.org/10.3389/fpsyg.2011.00039

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, *211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. http://doi.org/10.1007/BF00122574

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, *95*(3), 371–384. http://doi.org/10.1037/0033-295X.95.3.371

Tversky, A., & Simonson, I. (1993). Context-dependent Preferences. *Management Science*, *39*(10), 1179–1189.

Tymula, A., Rosenberg Belmaker, L. a, Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences*, *109*(42), 17135–40. http://doi.org/10.1073/pnas.1207144109

Ungemach, C., Chater, N., & Stewart, N. (2009). Are Probabilities Overweighted or

Underweighted When Rare Outcomes Are Experienced (Rarely)? *Psychological Science*, *20*(4), 473–479. http://doi.org/10.1111/j.1467-9280.2009.02319.x

van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Nature Scientific Reports*, *7*(December 2016), 40962. http://doi.org/10.1038/srep40962

van Duijvenvoorde, A. C. K., Huizenga, H. M., Somerville, L. H., Delgado, M. R., Powers, A., Weeda, W. D., … Figner, B. (2015). Neural Correlates of Expected Risks and Returns in Risky Choice across Development. *Journal of Neuroscience*, *35*(4), 1549–1560. http://doi.org/10.1523/JNEUROSCI.1924-14.2015

van Duijvenvoorde, A. C. K., Jansen, B. R. J., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192–203. http://doi.org/10.1037/a0025601

Van Leijenhorst, L., Gunther Moor, B., Op de Macks, Z. A., Rombouts, S. A. R. B., Westenberg, P. M., & Crone, E. A. (2010). Adolescent risky decision-making: neurocognitive development of reward and control regions. *NeuroImage*, *51*(1), 345–55. http://doi.org/10.1016/j.neuroimage.2010.02.038

Van Leijenhorst, L., Westenberg, P. M., & Crone, E. A. (2008). A developmental study of risky decisions on the cake gambling task: age and gender analyses of probability estimation and reward evaluation. *Developmental Neuropsychology*, *33*(2), 179–196. http://doi.org/10.1080/87565640701884287

Venkatraman, V., Huettel, S. A., Chuah, L. Y. M., Payne, J. W., & Chee, M. W. L. (2011). Sleep deprivation biases the neural mechanisms underlying economic preferences. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(10), 3712–3718. http://doi.org/10.1523/JNEUROSCI.4407-10.2011

Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., & Huettel, S. A. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making.

Neuron, *62*(4), 593–602. http://doi.org/10.1016/j.neuron.2009.04.007

Venkatraman, V., Payne, J. W., & Huettel, S. A. (2014). An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes*. http://doi.org/10.1016/j.obhdp.2014.06.003

Wagar, B. M., & Dixon, M. (2006). Affective guidance in the Iowa gambling task. *Cognitive, Affective & Behavioral Neuroscience*, *6*(4), 277–290. http://doi.org/10.3758/CABN.6.4.277

Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annural Review of Psychology*, *60*(October 2008), 53–85. http://doi.org/10.1146/annurev.psych.60.110707.163633

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological Review*, *111*(2), 430–445. http://doi.org/10.1037/0033-295X.111.2.430

Weigard, A., Chein, J., Albert, D., Smith, A., & Steinberg, L. (2014). Effects of anonymous peer observation on adolescents' preference for immediate rewards. *Developmental Science*, *17*(1), 71–8. http://doi.org/10.1111/desc.12099

West, S. L., & O'Neal, K. K. (2004). Project D.A.R.E. outcome effectiveness revisited. *American Journal of Public Health*, *94*(6), 1027–1029. http://doi.org/10.2105/AJPH.94.6.1027

Wysong, E., & Wright, D. W. (1995). A Decade of DARE: Efficacy, Politics and Drug Education. *Sociological Focus*, *28*(3), 283–311.

Yechiam, E., Barron, G., & Erev, I. (2005). The Role of Personal Experience in Contributing to Different Patterns of Response to Rare Terrorist Attacks. *Journal of Conflict Resolution*, *49*(3), 430–439. http://doi.org/10.1177/0022002704270847

Yechiam, E., & Busemeyer, J. R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making*, *19*(1), 1–16. http://doi.org/10.1002/bdm.509

Yechiam, E., & Telpaz, A. (2011). To take risk is to face loss: A tonic pupillometry study. *Frontiers*

in Psychology, *2*(NOV), 1–9. http://doi.org/10.3389/fpsyg.2011.00344

Zuckerman, M., Eysenck, S., & Eysenck, H. J. (1978). Sensation seeking in England and America: cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology*, *46*(1), 139–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/627648

Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology*, *65*(4), 757–768. http://doi.org/10.1037/0022-3514.65.4.757

APPENDIX

COMPREHENSION QUESTIONS FOR D-E GAP PARADIGM

Participants were presented with the following three comprehension questions at the end of the experiment to exclude out participants who may not have understood or been paying attention to instructions:

1. True or false: You had to sample each sampling machine at the beginning of the experiment 5 times

2. How many points did you start the experiment with?

   a) 25

   b) 50

   c) 100

   d) 200

3. How will your monetary reward be determined?

   a) From a running total of all the points you won/lost during all of the games

   b) From a randomly drawn game

   c) From the combined total of 5 randomly drawn sampling machine