DELAYING DECISIONS IN ORDER TO LEARN THE DISTRIBUTION OF OPTIONS

by

ADAM D. I. KRAMER

A DISSERTATION

Presented to the Department of Psychology
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2010

University of Oregon Graduate School

Confirmation of Approval and Acceptance of Dissertation prepared by:

Adam Kramer

Title:

"Delaying Decisions in Order to Learn the Distribution of Options"

This dissertation has been accepted and approved in partial fulfillment of the requirements for the degree in the Department of Psychology by:

Sara Hodges, Chairperson, Psychology
Sanjay Srivastava, Member, Psychology
Paul Slovic, Member, Psychology
Lynn Kahle, Outside Member, Marketing

and Richard Linton, Vice President for Research and Graduate Studies/Dean of the Graduate School for the University of Oregon.

September 4, 2010

Original approval signatures are on file with the Graduate School and the University of Oregon Libraries.

An Abstract of the Dissertation of

Adam D. I. Kramer for the degree of Doctor of Philosophy

in the Department of Psychology to be taken September 2010

Title: DELAYING DECISIONS IN ORDER TO LEARN THE DISTRIBUTION OF OPTIONS

Approved: ______________________________

Sara D. Hodges

This dissertation explores two basic hypotheses about how humans make decisions when presented with a sequential series of options: 1) people have a desire to learn about or experience qualities of the set of options available to them and will delay choice to gain such knowledge; and 2) delaying decision-making in order to better understand the set of options available will lead to better knowledge of the distribution of potential options and better decision outcomes. Three studies, conducted on a total of 302 college student participants, used an "optimal stopping" paradigm, in which participants viewed a series of options (in Studies 1 and 2, the options determined how their time would be spent in the latter part of the study; in Study 3, the options represented qualities of a hypothetical potential housemate). Participants had to choose or reject each in turn. I show consistent support for two hypothesis: Decision-makers continue to view and review options in order to gain a better understanding of the distribution of potential options, and decision-makers who have a better understanding of the option space end up with higher-quality decisions, using objective, subjective, and revealed-preference measures of quality. These results were consistent for multi-attribute decisions, single-attribute decisions, hypothetical decisions, and non-hypothetical decisions using both within-subjects and between-subjects designs. Individual differences among decision-makers did not show any consistent individual difference results, though decision-makers higher in numeracy appear to make better use of the cues available to them. In sum, decision-making appears to be aided by understanding the distribution of options, suggesting that it is occasionally wise to delay or "procrastinate" choice in order to gain an understanding of potential options when choosing.

CURRICULUM VITAE

NAME OF AUTHOR: Adam D. I. Kramer

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR

Carnegie-Mellon University, Pittsburgh, PA

Greenfield Community College, Greenfield, MA

DEGREES AWARDED:

Doctor of Philosophy in Psychology, 2010, University of Oregon

Master of Science in Psychology, 2005, University of Oregon

Master of Science in Logic and Computation, 2004, Carnegie-Mellon University

Bachelor of Science in Logic and Computation, 2001,
Carnegie-Mellon University

AREAS OF SPECIAL INTEREST:

Decision-Making

Human-Computer Interaction

Quantitative Methods

PROFESSIONAL EXPERIENCE:

Data Scientist, Facebook Inc., July 19, 2009-Present

Intern, Facebook Inc., June 20, 2008 – October 11, 2008; July 10, 2007 – October 11, 2007

Intern, Google Inc., July 5, 2006 – September 29, 2006; June 10, 2005 – September 21, 2005

GRANTS, AWARDS AND HONORS:

Student Travel Award, Society for Personality and Social Psychology, 2009

Graduate Student Research Award, University of Oregon, 2007

Graduate Student Research Award, University of Oregon, 2005

PUBLICATIONS:

Kramer, A. D. I. (2010). An unobtrusive behavioral model of "gross national happiness." *Proceedings of the 2010 Conference on Human Factors and Computing Systems (CHI 2010),* 287–290.

Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villanueva, B. R. (2010). Giving birth to empathy: The effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality and Social Psychology Bulletin, 36,* 398–409.

Newton, A. T., Kramer, A. D. I., & McIntosh, D. (2009, April). Autism online: A comparison of word usage in bloggers with and without Autism Spectrum Disorders. *Proceedings of the 2009 Conference on Human Factors and Computing Systems (CHI 2009),* 463–466.

Kramer, A. D. I., & Rodden, K. (2008, April). Word usage and posting behaviors: Modeling blogs with unobtrusive data collection methods. *Proceedings of the 2008 Conference on Human Factors and Computing Systems (CHI 2008),* 1125–1128.

Kramer, A. D. I. (2007). Placebo effects. In R. F. Baumeister & K. D. Vohs (Eds.), *The Encyclopedia of Social Psychology, 2*, (pp. 670–671). Thousand Oaks, CA: Sage.

Kiesler, S, Lee, S-L., & Kramer, A. (2006). Relationship effects in psychological explanations of nonhuman behaviors. *Anthrozoös, 19,* 335–352.

Kramer, A. D. I., Oh, L. M., & Fussell, S. R. (2006, April). Using linguistic features to measure presence in computer-mediated communication *Extended Abstracts of the 2006 Conference on Human Factors and Computing Systems (CHI 2006),* 913–916.

Powers A., Kramer, A. D. I., Lim, S., Kuo, J., Lee, S., & Kiesler, S. (2005, August). Eliciting information from people with a gendered humanoid robot. *Proceedings of Robot and Human Interactive Communication (ROMAN 2005)*, 158–163.

Fussell, S. R., Setlock, L. D., Yang, J., Ou, J., Mauer, E. M., & Kramer, A. (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction*, 19, 273–309.

Kramer, A. D. I., Fussell, S. R., & Setlock, L. D. (2004). Text analysis as a tool for analyzing conversation in online support groups. *Extended Abstracts of the 2004 Conference on Human Factors and Computing Systems (CHI 2004),* 1485–1488.

ACKNOWLEDGMENTS

I thank, in chronological order, Robyn Dawes and Margaret Clark for instilling in me an interest in social psychology and decision making and helping me get a foothold, Susan Fussell for playing the role of my first advisor though she was technically my boss, and Sara Hodges for six years of cheerful, creative encouragement, support, and advisorial training in topics ranging from psychology to politics to food and perhaps unknowingly preventing untold amounts of stress. I also thank Sreevardhan and Vydehi Mekala for teaching me how to make the hundreds of meals of curry that fed me throughout graduate school and the writing of this dissertation; the graduate students of the UO Psychology Department for supporting me through the thick of it; DPB, the RainSong winery, and the Jameson distillery for distracting me whenever necessary; and my family for repeatedly asking what I study and smiling at whatever I said.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Overview 1

Prior Work 3

Optimal Stopping 3

Procrastination 11

Decisional Delay: Qualities of the Decision 15

Individual Differences 28

Summary of Hypotheses and Prior Work 36

Overview of Studies 36

II. STUDY 1 39

Hypotheses 40

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable 40

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful 43

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information 44

Summary of Hypotheses 45

Method 47

Chapter Page

Participants 47

Procedure 47

Results 51

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable 53

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful 56

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information 59

Discussion 60

III. STUDY 2 67

Hypotheses 68

Summary of Hypotheses 70

Method 72

Participants 72

Procedure 72

Results 74

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable 75

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful 78

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information 84

Discussion 85

Chapter Page

IV. STUDY 3 92

Hypotheses..98

Method..101

Participants 101

Procedure 101

Results 103

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable 104

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful...................105

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional

Information 108

Discussion...109

V. GENERAL DISCUSSION 112

Overview 112

Implications 114

Limitations..134

Future Work..142

Summary...146

REFERENCES 147

LIST OF FIGURES

Figure Page

1. Graph of option quality as a function of order 40
2. Participants' view of the decision they were making in Studies 1 and 2 49
3. Illustration of the X-shape found between subjective and objective value 53
4. Participants' view of the decision they were making in Study 3........................... 102

CHAPTER I

INTRODUCTION

Overview

The purpose of this dissertation is to examine one reason why people delay making decisions. Specifically, I hypothesize that one reason why people put off choosing is to gain an understanding of the distribution of options potentially available to them (gaining distributional information is *desirable*), and that gaining and using this information will lead to better decision-making processes (having distribution information, and delaying choice in order to gain such information, is *useful*).

The three studies described here are designed to explore the way in which people make decisions in contexts in which options are presented in series and an accept/reject decision must be made for each item. The studies will examine how decision-makers differ in both the strategies they use and the outcomes they receive for these kinds of decisions. These kinds of decisions are found frequently in real-world scenarios. For example, imagine choosing a roommate: Usually each potential roommate "interviews," one at a time, and the search ends when one roommate is chosen. When the decision-maker says "no" or "we'll call you later" to a potential roommate, there is a very real chance that the option will disappear—the applicant needs somewhere to live, and is certainly looking at other places.

The roommate scenario is a good illustration of the phenomenon this dissertation will explore for other reasons as well. There are, of course, many qualities of a roommate that bear evaluation—does the roommate seem personable? Does he have the means to pay rent reliably? Is she sufficiently clean? When a roommate search begins, most people are able to list desirable qualities of such a roommate, but they may not be in a position to "hold out" for the perfect candidate: If the person with the room to rent deliberates or interviews for too long, it could result in having to pay rent on an empty room, though taking the first applicant to walk through the door could result in many other problems. However, if one has not searched for a roommate before, or perhaps has not before searched for a roommate in a particular town or for a particular dwelling, it is not always clear exactly what one should or could reasonably hold out for. If the applicant is unlikable and unemployed, the "reject" decision may be easy, but if a perfectly average fellow applied, without knowledge of the pool of potential roommates, the decision-maker might well wonder whether one could do better. When a second applicant arrives, however, the two can be judged relative to each other—who is more likeable, who more financially solvent? After interviewing a few more individuals, the person with the room to rent may start to form an idea of the distribution of available candidates, and thus be able to better evaluate individuals within that context—and also able to estimate how likely it is that searching for longer will get somebody better than the candidate at hand, or a candidate of a certain quality. This decision process, broadly speaking, is referred to as an "optimal stopping" or "optional stopping" decision (Shaklee, 1958)—decision-makers may stop viewing options whenever they like, by making a choice. As each option is reviewed, the choice (thus stopping) is "optional," and the decision-maker's task is to discover when it is "optimal" to stop.

Delaying a decision in order to learn the distribution is unique in that it presents a case where delaying choice may be wise. Learning what counts as a "good" choice may prevent the decision-maker from making a sub-optimal choice. Learning the distribution places decisions in context: In many real-world decisions made in the course of everyday life, some background knowledge exists regarding not only the options among which a decision-maker chooses, but also regarding the entire population of options. The purpose of this dissertation is to show that gaining this knowledge is both desirable and useful, and constitutes a wise reason for delaying choice.

Prior Work

This project ties together research from social, cognitive, quantitative, and personality psychology. This section will first review the literature on "optional stopping" paradigms, as that is the primary methodology used in the studies to follow. I will then cover research on delaying action in general (which is mainly research on procrastination), research on decisional procrastination specifically, and then individual differences that may be related to delaying action and choice.

Optimal Stopping

An optimal stopping problem is a problem where it is unclear when it is "optimal" to "stop." In the most commonly studied optimal stopping problem, called the "secretary problem," decision-makers are told that they need to hire a secretary, but that they are very busy and very picky: Only the very best secretary will do (anyone but the best is useless), and they will spent one afternoon interviewing secretaries, making a choice to hire or reject each secretary interviewed in turn. In other words, the goal is to interview and reject secretaries until the one thought to be "the best" has been found. At this point, the hire is made, and all secretaries hypothetically waiting in the lobby for their interview are sent home (Ferguson, 1989). Optimal stopping problems are also sometimes called "optional" stopping problems, because though the task is to discover when it is *optimal* to stop, "…the subject has the *option* of waiting" (Corbin, 1980, p. 54, emphasis added).

While optimal stopping problems were not originally designed to study decision delay *per se,* the paradigm is ideal for doing so: At each point in the problem, participants can either choose an option or delay longer. Using hierarchical modeling techniques not available when most research into optimal stopping was conducted, I will look at qualities such as participants' subjective ratings of the option and relative qualities of the option (for example, whether a given option is part of a positive or negative slope, and whether it is "good" relative to other options seen so far), as variates affecting the likelihood of choice.

The original "secretary problem" serves as the paradigm most frequently used to study optimal stopping. A problem is considered to be a "secretary problem" if it follows six basic assumptions (Ferguson, 1989); optimal stopping problems in general tend to follow the same set. The studies in this dissertation more or less resemble a "secretary problem" and thus I review these assumptions here (using "secretary problem" notation), and note how and when the current studies deviate from them. Assumptions are quoted from Ferguson (1989, p. 282):

1) "There is only one secretarial position available." My thesis retains this assumption, as there are many real-world situations in which there is only one choice that can be made.

2) "The number n of applicants is known." My studies retain this assumption, though I note that in many cases the *precise* number n is not known: For example, in a roommate search, the decision-makers may not have knowledge of how many people will be looking for rooms, but they likely know when rent is due and about how frequently new roommate options appear.

3) "The applicants are interviewed sequentially in random order, each order being equally likely." In the real world, options may get better or worse over time, and the literature shows that people are sensitive to improving or worsening option sets (Corbin, Olson, & Abbondanza, 1975; Hsee & Abelson, 1991; Hsee, Salovey, & Abelson, 1994), so I retain this assumption.

4) "It is assumed that you can rank all of the applicants from best to worst without ties. The decision to accept or reject an applicant must be based only on the relative ranks of those applicants interviewed so far." In the real world, "ties" are indeed possible. When choosing which plane ticket to buy online, one might see the same "best ticket" for two days in a row. Further, most hard decisions are ones in which there are multiple dimensions of quality, such as weighing the personability of a potential roommate against

her financial stability (Yates, Venoitt, & Patalano, 2003), leading to ties when one dimension for one option perfectly (or stochastically) balances out another dimension in another option. Many financial decisions lead to ties whenever the payoffs (or expected utilities, or expected values) are equivalent. Further, making decisions based on solely relative ranks is often not a reasonable assumption for the real world, in which we may have objective points of comparison—we may evaluate potential housemates relative to a prior housemate, or to our notion of what's "best" or "unacceptable:" If the best applicant out of 100 interviewed is still someone with whom we do not get along, a decision to reject this individual would indeed be due to objective quality rather than relative rank. As I am not trying to derive a normative strategy, this assumption is not directly relevant to my studies, though my paradigms may at times assign the same value to more than one option, or allow participants to do so (see Methods for Study 1, below).

5) "An applicant once rejected cannot later be recalled." In some real-world cases, it may be possible to "recall" an option (e.g., to tell an applicant "we'll call you back later," understanding that the applicant might find another job before a final decision about whom to hire is made). I retain this assumption for these studies, but note that future work should include an exploration of decisions for which this assumption is violated.

6) "You are very particular and will be satisfied with nothing but the very best. (That is, your payoff is 1 if you choose the best of the *n* applicants and 0 otherwise)." This assumption, couched at the end of the problem statement and with a plea to imagine "very particular" bosses, would be a sign of psychopathological obsession in the real

world. Further, it creates an internal inconsistency with assumption 4, above: If there are at least three secretaries applying for the job, then assumption 4 requires that the second- and third-best secretaries be of different, determinable rank (i.e., not be equivalent) while assumption 6 ultimately defines their values as equivalent. Further, even the most particular boss who considers his decision an abject failure if he does not select the best secretary will still derive differential utility from the services of different secretaries (once he gets through fuming). This assumption is dropped for my studies. However, I will explore individual differences in how important it is to deciders to get the "maximally" best option.

This sort of problem has attracted a notable amount of research from quantitative psychologists, ranging from theoretical proofs of when it is "optimal to stop" (e.g., Edwards, 1965; Mehle & Gettys, 1979; see Ferguson, 1989 for a review), to computer simulations designed to "experimentally" explore the efficacy of different decision stopping rules in different implementations of the problem (e.g., Seale & Rapoport, 1997; 2000). Empirical research on human decision-making (rather than research using mathematical models of decision-making) in optimal stopping problems has a history in the cognitive psychology literature prior to around 1981. Shaklee (1958) suggested that researchers should study what leads people to "cease observation and act," coining the term "optional stopping" as it applies to any planned action.

Various researchers of the secretary problem have already relaxed various assumptions of the secretary to various uses. For example, Rapoport and Tversky (1970) provided complete distributional information to subjects and assumed that the subjects would understand and use it, but they did not test or whether or how this information was used. Seale and Rapoport (1997; 2000) systematically tested each assumption in turn against a number of decision rules, but did not use human decision-makers. Other researchers used very similar paradigms to examine how long it takes participants to correctly guess some quality of a sequence of numbers (e.g., the slope or the frequency of some response; Corbin, Olson, & Abbondanza, 1975; Edwards, 1965; Fried & Peterson, 1969; Howell, 1966; Rapoport & Tversky, 1966, 1970), largely to discover stopping rules in certain contexts. It is striking, however, that few of these researchers studied choice contexts in which participants actually got the choice they made (something that would, of course, be quite impractical in the secretary problem). These researchers also did not have participants make decisions *based on* the distributions they viewed. Instead, they had participants make direct judgments *about* the distributions themselves (such as which deck of cards has which distribution; Fried & Peterson, 1969).

Two papers did at least examine human decision-making processes in which participants were rewarded for optimal choosing, however: Rapoport and Tversky (1970) gave participants a very straightforward optimal stopping problem: pay to see the next card, or decide to "stop." When the decision-makers stopped, they were paid a sum determined by the cards they had seen so far. However, Rapoport and Tversky's question of interest in this study was whether participants' choices were close to a previously derived optimal model of stopping, rather than addressing the question of how the choice is made (Howell's 1966 investigation was similar).

Corbin and her colleagues (1975) noted this lack of investigation into how participants utilized information about viewed options in order to infer qualities of future options, and attempted to examine "…the processes by which subjects make their selections," (p. 209), and did so by having participants draw cards from a deck and to stop when they believed that they had drawn the maximal card in the entire deck. Corbin's analysis, like mine, examines whether the values of the first few cards drawn (i.e., the distribution of prior options) determines whether participants believe that a chosen card is maximal for the deck. Corbin and colleagues, however, used decks with only five cards in them, and examined only two factors in determining whether the second or third card was chosen, making their conclusions very narrow. First, they examined whether the second card is chosen when the first card is lower (it is, sometimes), and whether this likelihood changed if the first card is far lower or slightly lower (bigger jumps increase choice likelihood). Second, they examined whether the third card is more likely to be chosen if the second card is far below the first card (an "initial decrease") or mildly higher than the first card (the initial decrease did increase choice likelihood).

The following studies will suggest that observed stopping choices are part of a much broader effect: Decision-makers are learning the distribution of options, and that these qualities can be represented in broader terms than describing the jump in value from the first option to the second option and to the third. For example, in the two-card case that Corbin and colleagues (1975) studied, they simultaneously manipulated the mean and variance of the training set; in the three-card case, they also manipulated the skew and slope. In my study I will examine effects of mean, variance, and slope explicitly, treating them as qualities of the underlying distribution of options rather than qualities of the set of options under consideration. Corbin and colleagues also failed to take into account any training effects across decks: Each participant evaluated 120 decks, but though the "history" of cards viewed was their primary topic of interest, they did not examine "history" beyond the current deck. Corbin and colleagues also rewarded participants only for choosing the optimal choice (secretary problem assumption #6, above). There has also been no research to date on how individual differences affect delay length (these previous studies were conducted during the heyday of behaviorism), nor has there been research exploring the processes by which delay is engendered.

I will extend these studies by going in a more applied direction: The current studies will model when people choose to stop in problems such as this, rather than trying to determine a single optimal stopping point. They also seek to evaluate whether people gather distributional information at all, and if so, whether people then make use of this information to improve their choices. The goal of the present studies is less oriented towards mathematically determining how to find the *best* secretary, so much as whether people faced with situations such as the secretary problem are able to choose a *good* secretary (objectively defined via quality metrics, but also in terms of the decider's subjective perceptions and revealed preferences). In essence, I hope to use the secretary problem paradigm in order to understand how real people make decisions—and how good those decisions are—in real-world analogs of the secretary problem.

Procrastination

Learning the distribution of options, of course, does not come for free. In most real-world decisions, one must view a set of real options (without choosing any) in order to have seen enough options to have a reasonable idea of what counts as a "good" option. This necessary evaluation period constitutes a delay in decision-making—if one does not choose the first option, one has delayed choice, and if one chooses the first option offered, no distribution learning has taken place. As such, the delay of choice is effectively a delay of action. There is an extensive literature on delaying actions more broadly: the study of procrastination. However, while definitions of procrastination vary across researchers (Steel's extensive review and meta-analysis spends half a page and cites 17 references in the interest of defining it; Steel, 2007), procrastination in its essence is negatively valenced: It involves putting off something that one intends to do and could do now, without "good reason" to delay or even with an expectation to be harmed by the delay. In this paper, I will use the word "delay" in a neutral manner, unless otherwise noted, in order to distinguish it from more negatively valenced "procrastination." I will argue that delay may be useful in many cases, and that there is much to study in the realm of delayed decisions beyond that which is called "procrastination."

Procrastination's negative reputation, however, is not unmerited. Prior work shows that individual differences in procrastination do exist, and that those who procrastinate also tend to be poor at scheduling their lives, either actively refusing or passively ignoring scheduled tasks (Ariely & Wertenbroch, 2002; Tice & Baumeister, 1997). People procrastinate on everyday activities because they predict that the activities will be unpleasant, because they view the tasks as impositions or unfair demands, or because they fear they will not complete the task adequately (Effert & Ferrari, 1989). Procrastination in school contexts has been shown to lead to poorer grades on work for which people chose to procrastinate, as well as greater stress and worse health (ostensibly due to the stress; Tice & Baumeister, 1997). Procrastinators are poorer at self-regulation, and the extra "free time" they gain from not doing tasks when they should be optimally done does not protect them from lowered overall satisfaction with life (Milgram, Sroloff, & Rosenbaum, 1988).

In their defense, procrastinators also feel less stress while their tasks are put off than those who do not procrastinate, leading in some contexts to an experience of less stress overall (Tice & Baumeister, 1997). Procrastinators also are not easily classified as simply "haphazard actors," and do indeed behave in a manner that indicates careful or systematic thought, often caring about completing their task and working towards the task (albeit at an insufficient rate; Ferrari & Dovidio, 2000). Some researchers have found a striking absence of negative personality correlates of procrastination (e.g., they are not more neurotic overall; Ferrari & Dovidio, 2000).

While the above research has defined "procrastinators" as those who self-report procrastinating, other research has identified behavioral measures that distinguish between procrastinators and non-procrastinators. Ariely and Wertenbroch (2002) show that in fact most people do indeed procrastinate (to their detriment) to some degree, indicating that procrastination is a problem for many or even most individuals (and that the median split methodology of self-reported procrastinators used by Steel, Ferrari, and colleagues may lump many people who procrastinate into the "non-procrastinator" category; Effert & Ferrari, 1989; Harriott & Ferrari, 1996; Steel, Brothen, & Wambach, 2001). However, Ariely and Wertenbroch also suggest that people show adaptive resistance to procrastination. For example, when not provided a deadline for a school project, many people created their own (earlier) deadlines in an effort to not miss a deadline that actually matters, and these individuals fare better than those who do not create their own deadlines (though not as well as individuals whose deadlines were externally imposed).

Whereas Steel and colleagues (2001) suggest that the lack of correlation with personality variables may be due to the difference between self-report and behavioral scales (the "procrastinator" label may threaten one's self-concept), another explanation is that procrastination itself may in some cases be justifiable—or even admirable. For example, if further reflection leads one to realize a flaw in one's plans, procrastination may avert crisis. Indeed, in the case of delaying decisions, the parallel holds—if one puts off making a choice and then better options become available, decision makers would likely consider themselves wise for waiting, even if they didn't know that the better option would appear. Similarly, the research literature on procrastination has not (to my knowledge) experimentally manipulated participants into situations where delay may provide a benefit other then the oft-maligned benefit of not having to think about an undesirable action.

Some evidence for positive effects of delaying action in general (rather than decisions specifically) does exist. Recent research on "unconscious information processing," in this context, suggests that those who delay (either actions or decisions) may in fact be better prepared to act when the time comes (Dijksterhuis & Aarts, 2010). Tice and Baumeister (1997) find that procrastinators experience less stress while procrastinating although they experience much more stress when deadlines come due (leading even to illness when many deadlines come due at once, e.g., during "finals week" for a college student sample). Vohs, Baumeister and their colleagues have long argued that any self-control task (such as completing an assigned task, doing effortful work, or even simply making a decision) depletes one's ability to for self-control on subsequent (irrelevant) tasks (Baumeister, Bratslavsky, Muraven, & Tice, 1998; Muraven & Baumeister, 2000; Vohs et al., 2008). This suggests that procrastinators may be conserving personal resources by waiting until the last minute or putting off action. In other words, the utility of procrastination (and, I will later argue, the delay of decisions) may depend on the activity (or decision) being procrastinated. When all deadlines come due at once (as in the case of "finals week" for oft-studied student samples), procrastinators may suffer quite negative effects (Tice & Baumeister, 1997), though the reduced overall stress that Tice and Baumeister show may in fact lead to better outcomes overall when deadlines are more spread out (Tykocinski & Ruffle, 2003).

This also brings into question another item sometimes noted in definitions of procrastination: Whereas Steel (2007) argues that in order for a delay to be considered "procrastination," individuals must *expect* to be worse off for delaying, a conservation of energy model may suggest that people often either expect things to go fine (or at least have no expectation of things going awry). Furthermore, research on the "Planning

Fallacy" shows that people frequently delay action until the last minute because waiting until the last minute is fine so long as nothing goes wrong (see Buehler, Griffin, & Ross, 1994 for a review of the planning fallacy). Whether or not this sort of delay should be called "procrastination" is not widely agreed-upon. This suggests that there may be two basic processes underlying an instance of procrastination: a tendency to avoid acting, and a (separate, but perhaps correlated) misregulation in estimating required effort and time to task completion. In other words, it is possible that people can delay without expecting the delay to hurt (even though on average it does, and they should arguably eventually know better). A two-process theory of procrastination indeed suggests that there would be situations, such as the delay of decisions in order to gain knowledge of the option space, in which delay is non-harmful or even helpful.

Decisional Delay: Qualities of the Decision

Decisional delay differs from procrastination in a few important respects. First and most obviously, decisional delay is the delay of decisions and not actions in general, making it a more specific construct. This suggests both that lessons learned of procrastination can likely be applied to decisional delay, but also that there may be more to learn about the delay of decisions specifically. Second, decisions do not always require any actual activity—we frequently act immediately upon deciding, leading these to be conflated, but the action is not the decision: As Yates and colleagues (2003) put it, making a decision is making a "...commitment to a course of action..." (p. 15). This draws another distinction between decision delay and procrastination, as procrastination is usually defined as putting off an action (Steel, 2007). This is an important distinction, as procrastination has usually been studied in contexts where there is a non-negligible amount of work necessary to complete a task: Writing even a lousy paper takes some amount of time, while committing to a lousy course of action can be trivial. Procrastination of acting and of deciding are related, of course: One can commit to a course of action and then put off the action indefinitely, and the term "action" is sufficiently broad as to include many trivial acts. One process that sits in the middle is the Planning Fallacy, in which those who have decided upon an action misestimate how long it will take to complete it (Buehler et al., 1994).

Decision delay also differs from procrastination in that it is hard to argue that "procrastination" is good, while delaying choice may well be. When deciding, further deliberation (either conscious or unconscious) prior to committing to a course of action may lead to a more reasonable choice, or at least to gaining additional information (Bastardi & Shafir, 1998; Dijksterhuis, Bos, Nordgren, & van Baaren, 2006). This in turn may lead to better outcomes in some contexts, while procrastination (by definition) is consistently negatively valenced.

Similarly, there is some evidence that people simply enjoy waiting to make certain decisions, ostensibly to allow for time to reflect or deliberate on the information at hand or to allow new information to come to light. Tykocinski and Ruffle (2003) make this argument well, in their response to Bastardi and Shafir (1998): They show that even when decision-makers are given all the information potentially provided to them, they will still delay if not forced to choose immediately. They do this over several studies by setting up three basic conditions in which participants must choose to enroll in an interesting elective class, or to not enroll: 1) They learn that the good professor is on vacation; 2) They learn that the good professor might be on vacation (and are given the option of waiting to find out; conditions 1 and 2 were both in Bastardi and Shafir's paper), and 3) They learn that the good professor is on vacation (but are given the option to wait until tomorrow for no reason). As many students opted to "wait for no reason" as opted to "wait to learn about the good professor's vacation," indicating a preference for waiting in general.

Beyond whether people wish to delay choice or not, I also separate qualities of the decision that engender delay into two broad categories: qualities of the decision context, and qualities of the options available. If a decision is a "commitment to a course of action that is intended to produce a satisfying state of affairs" (Yates et al., 2003, p. 15), then many decision-making researchers have argued (at least implicitly) that when the same courses of action are available (the choice set is the same) and the same states of affairs are expected from those courses of action (the outcomes are the same), then choosing a particular course of action more than once constitutes making the same decision. However, research has shown that the "same" decisions may be decided upon differently due to qualities of the decision-making context. Such context effects in decision-making are argued to indicate some amount of irrational (or at least haphazard) choosing behavior on the part of the decision-maker, as changing one's mind for reasons unrelated to the options or their expected outcomes is hard to defend.

Context effects broadly relate to the delaying of decisions because any delay in choosing may bring about a change in context. For example, putting off making a decision because one is in a bad mood may be "wise," as affect and emotions are known to alter the perception of courses of action or expected states of affairs (indicating that the decisions made under the influence of some emotions are not made via the same processes as those made under other affective states; c.f. Clore, Gasper, & Garvin, 2001; Finucane, Peters, & Slovic, 2003; Lerner & Keltner, 2001). Affective states also "carry over" into the decisions for which the affect is irrelevant, altering the choices people make (c.f. Forgas & Bower, 1987; Lerner, Small, & Loewenstein, 2004; Loewenstein & Lerner, 2003; Schwarz & Clore, 1983).

Many qualities of a decision-making context, defined colloquially, may be relevant to the decision at hand. For example, in a poker game where the cards in front of one define the decision-making context, the context is certainly related to the courses of action and expected outcomes. The case in which context effects are unrelated to the components of the decision, however, are more interesting for the purpose of this paper—specifically when delaying a choice induces a change in context (gaining a better knowledge of what future options might be), or when a certain context leads a decision to be delayed.

Some researchers, notably Patalano and Wengrovitz (2007), have created experimental contexts in which delaying decisions may have positive or negative consequences. They use a methodology of simulating a course selection process in which students choose a course to take in the following academic term, over several simulated days, including a "risk" condition in which new courses may be offered and currently available courses may fill up. Categorizing people into "decisive" and "indecisive" groups using a median split on the Frost and Shows (1993) scale, Patalano and Wengrovitz showed that the most "indecisive" (i.e., obsessive or choice-averse) half of individuals waited longer and viewed more options than the less indecisive half, regardless of whether the courses in question had a high chance of filling up. When the course had no chance of filling up, the best (normative) strategy for deciding would be to wait until all courses are available and then choose the best; when the courses were more likely to fill, waiting is no longer normative. Patalano and Wengrovitz also showed that the less decision-averse individuals were indeed more sensitive to risk (i.e., when in the risk condition, they chose faster), while the more decision-averse individuals took the same amount of time in both conditions. However, indecisive individuals waited longer overall—in other words, even when there was little risk for waiting, more decisive individuals appeared to ascribe a cost to waiting, leading them to choose faster and obtain less optimal outcomes (defined nomothetically, without reference to the participants' stated preferences about what they look for in a course) when the better choices appeared later in the choice set.

This constitutes evidence that delay can lead to better outcomes so long as the delay is for "good reason." Whether the more decision-averse individuals were basing their choices of whether to delay on information they were shown or not was not addressed by Patalano and Wengrovitz (2007); I also note that their two measures of decision aversion (Frost & Shows' 1993 Indecisiveness Scale and the Melbourne Decision Making Questionnaire's decision procrastination scale by Mann, Burnett, Radford, & Ford, 1997, discussed at greater length in Individual differences: Decisional delay, below), which were reasonably correlated, do not produce consistently similar results when used interchangeably in Patalano and Wengrovitz's studies.

Beyond decision context, the options that are available to decision-makers appear to influence the relative valuation of other options in the choice set. When delay is an available option, the choice of whether to delay should similarly be affected by the other options available. A better understanding of the distribution of options available may lead decision-makers to delay when they expect better options to appear. For example, if one has never booked a flight to San Francisco, a $200 ticket might seem reasonable (or not), but many people (anecdotally) report waiting to "see what the prices do"—even if they might go up, or if paying $200 would not be a burden. Indeed, research on the relevance of irrelevant alternatives suggests that in many cases, irrelevant or wholly dominated options can lead decision-makers to avoid a choice, seek other options, or reverse their preferences (Dhar, 1997; Shafir, Simonson, & Tversky, 1993; Tversky & Shafir, 1992). This is not necessarily a pure "context effect," however: From a statistical standpoint, the presence of additional options communicates to decision-makers information about the distribution of the whole option set (including options that are unavailable, and thus options that might become available if the decision were delayed). Thus, seemingly "irrelevant" data points may actually be relevant: Delay is sought for the express purpose of "gaining information," even if relying upon that information may not be wise (c.f. Bastardi & Shafir, 1998). Though Bastardi and Shafir argue that this process (specifically, the use of irrelevant information) is irrational, this work does show that "gaining information" is one reason people delay, a tactic that may prove wise in other contexts.

Tversky (1972) also discusses the theoretical problem of "independence of irrelevant alternatives" in his elimination by aspects model. This model suggests that difficult decisions may be made, in part, by rejecting options outright (i.e., in a non-compensatory manner) if they have or lack some important "aspect:" If someone is indifferent between a vacation in Europe or the Far East (probability of 0.5 of choosing either), they should be indifferent between two similar trips to Europe and two similar trips to the Far East (probability of 0.25 of choosing any). This leads to a conclusion that such a decision-maker would also be indifferent among one trip to Europe and two trips to the Far East (with a probability of 0.33 of choosing any of the three). Rather, we might more intuitively expect something closer to a probability of 0.5 for the trip to Europe and a probability of 0.25 for either trip to the Far East, effectively evaluating the "aspect" of locale first and eliminating those options that do not pass muster for that aspect. Conversely, if some triviality were provided (such as a free bottle of wine accompanying only one of the trips to Europe), then the European trip that lacked the "wine" aspect would serve to eliminate one European option (if Europe was chosen during the evaluation of the "locale" aspect, earlier). This finding, and Tversky's elimination by aspects model, also provide evidence for "choice set" effects, or that the overall choice context is necessary to predict choice; knowing the qualities or desirabilities of the options is not enough. Tversky's model suggests that in this case, an aspect of the choice (such as the trip being to the "Far East" or "Europe") would be eliminated from consideration. My options-as-context model suggests another route to the same result: The value of an option depends intrinsically on the other options available to the decision-maker. Effectively, a trip to Europe without free wine means something different when it is evaluated relative to a wineless trip to the Far East than when it is evaluated relative to a similar trip to Europe with free wine. Additions to (or alteration of) any option in a pre-existing choice set may lead a decision-maker to revalue the other options—for example, the option of "waiting to choose"—as their knowledge about the population of options has changed.

Studies of the so-called "disjunction effect" have also shown effects of shifting option sets on decisional delay. For example, Bastardi and Shafir (1998; see also Tversky & Shafir, 1992) show that participants offered the opportunity to choose whether or not to take a vacation prefer to wait for contextual certainty (knowing whether or not they passed a difficult exam), even though they would have chosen to go on vacation in either case (participants who were asked to suppose that they passed or failed the exam before choosing both chose to vacation). While this may be due to experimenter demand effects, as the "wait" option was only offered in the "unknown outcome" condition may have suggested to subjects that waiting was a wise option (Grice, 1975; but see also Tykocinski & Ruffle, 2003, who suggest that people simply prefer to wait when possible), it is reasonable to conclude from these studies that people prefer to wait 1) if there is no cost to waiting, and 2) if they can identify information that they would gain by waiting (Bastardi & Shafir, 1998; Duncan, Wengrovitz, Sedlovskaya, & Patalano, 2007; Tykocinski & Ruffle, 2003). However, humans have difficulty taking the perspective of someone who has knowledge they don't or who does not have knowledge they do—likely including themselves at a future time (see Camerer, Loewenstein, & Weber, 1989; Fischhoff, 1975; Royzman, Cassidy, & Baron, 2003 for cognitive failures to take perspective, and Ariely & Loewenstein, 2006; Loewenstein, 1996; Van Boven & Loewenstein, 2003 for visceral and emotional failures to imagine one's own perspective at a different time). Given this difficulty in imagining hypotheticals, it was arguably wise for Bastardi and Shafir's decision-makers not to judge the information as "non-instrumental" before they had the information available (especially as the experimental paradigm called attention to this lack of information immediately before asking decision-makers to choose).

Dhar (1997) also suggests that choice deferral, or preference for a "no choice" or "seek more options" option, can result from changes in the option set—when a new option is included in the option set, there may be a higher or lower chance that the decision maker will choose deferral. Consider a choice with three options: A, B, and delaying the decision. Dhar shows that people are more likely to defer when option B is similar to A and less likely to defer when B is far inferior to A, compared to a two-option choice (including only item A and delay). While these studies make an important point about how adding or deleting options from a set can alter the relative desirability of those options, the explanations previously provided still remain at the level of the options themselves (in this case, interactions among the options such as "one is inferior to the other"), but stop short of addressing a broader issue of the choice set composition or the decision-maker's understanding of the population of options from which the proffered options are drawn. I suggest that the process at work in Dhar's studies is an effect whereby decision-makers are seeing a more variable option set as deserving of more consideration than a less-variable option set, and are thus delaying to gain more knowledge of the distribution.

These studies together consistently suggest that individual options take on different meaning depending on what the other available options are; that making changes in the overall choice set may alter delays in choosing when delay is an option; and that the qualities of the individual options apparent to participants at any one point may be insufficient to explain these delays. The explanations of extreme options, tradeoff contrasts, and option conflict may be specific cases of a more general phenomenon of understanding the distribution of options available to the decision maker, and determining whether, given that decision maker's estimate of the distribution, she is likely to find options in the future that are superior to the ones currently available. For example, consider a variance-based description of Dhar's (1997) study: Dhar shows that when there is only one option in the option set (and thus no variance or information about the distribution), deferral happens at a certain rate depending solely on the quality of the option. When another option is added, deferral goes up when the variance is low (the options are similar and thus it is hard to tell what counts as "good") and down when the variance is high (one option is clearly superior). However, it is less clear whether this pattern would continue with the addition of subsequent options: I predict that at some point, adding enough similar options (i.e., choosing among a low variance set with enough observations to determine the distribution, such as the vast array of nearly-equivalent jams on a supermarket shelf) is likely to lead subjects to feel confident about making a choice and to not further waste their time by examining additional options. In a high variance set, however, the next option viewed could be far superior to prior options (or far worse), leading decision-makers to keep searching for longer, hoping for a good, great, or fantastic option. For example, contrast wine-tasting tours taken to find the "next great wine" with the process of buying an acceptable wine from the supermarket, or scavenging tag sales for a "diamond in the rough" furnishing item rather than going to Ikea for something that's "good enough."

Similarly, adding a third option to Dhar's study (defined above, in which B is inferior to A) could generate a skewed choice set, with one option far above or below the others. If decision-makers are expecting a certain distribution of option quality (in Dhar's example, this would mean quality within product categories such as bookshelf speakers, answering machines, or laptop computers), then the mismatch between the expected distribution and the observed distribution should predict choice delay: Whatever options are "missing" from the observed set of options are the ones most likely to appear in the future, and so if these options are good, participants should delay.

This process is formalized in the literature on applying bounded-rational satisficing procedures to everyday life (e.g., Dawes, 1979; Hastie & Dawes, 2001). The rationale of "evaluate a subset of choices, then choose the best" is justified in this literature via an argument that a sample of a distribution represents the population distribution, and thus gaining information about the options within the sample set is useful. Some research, however, has suggested that there is such a thing as "too much choice" (Botti & Iyengar, 2004). This would, in turn, suggest that there is a limit to the payoff from knowing more about the option space. In situations where many similar choices are available (as in the cavalcade of jams used in Botti and Iyengar's studies), the information could be considered largely redundant—the distribution is learned quickly, and nothing really stands out. As people are largely focused on the relative (rather than absolute) value of options (c.f. Kahneman & Tversky, 1979, and their discussion of "reference points"), a distribution-focused account of Botti and Iyengar's result might be that too much *choice* is not aversive, but that unnecessary or redundant evaluation is. In terms of distribution learning, a great number of very similar options constitute a distribution with low variability; I suggest that this may promote less evaluation and less delay in choosing. As such, when yet more options (that scarcely differ) demand evaluation, the overall choice process should be more aversive. In terms of the distribution of quality, a "too much choice" study in which options are highly similar essentially produces a histogram with one lone bar—any option will do, and no option is great, potentially explaining the lack of interest in (and subsequent enjoyment of) the purchased jam in their "too much choice" condition. Findings using this methodology—in which items are close to indistinguishable—do not contradict the hypothesis that an individual seeks to know (and perhaps benefits from knowing) the relative standing of his or her options. Knowledge of the distribution of options is helpful, even if it means knowing not to hold out or continue to evaluate options hoping for a "better" or "great" option in a sample that shows little to no variance.

Beyond the sort of experimentally controlled context found in the "too much choice" studies, however, there are few studies that examine the relationship between the context in which a decision is made and the *extent* to which the decision is delayed, or which actually test whether participants are waiting to choose based on how well they understand the set of (potential) options available to them. One reason for this is likely the fact that varying experimental context often constitutes manipulating many variables simultaneously, making it difficult to argue that the contextual variable of interest is the one causing a noted difference.

For example, Patalano and Wengrovitz (2007) studied whether participants would "put off" choosing which hypothetical course to take in a "course sign-up" paradigm. They compared a "riskless" condition (courses were very unlikely to fill up) to a "risky" condition (courses were nearly full), and allowed students to delay choice for several "days" to see whether additional courses would come available; to make the point that choice was risky, during the delay period, some courses would in fact "fill up." As the same courses were used in both conditions, Patalano and Wengrovitz may be tempted to argue that the "contextual variable" of "riskiness" is at work when choice delay behavior is altered across conditions, when they have in fact manipulated more than just risk: For each "day" that participants evaluated, the set of options available was different in the "risky" choice condition, meaning that the relative (objective) values of each course had changed. They are still comparing apples to apples, but how good "the best apple in the bunch" is viewed as may depend on the quality of the rest of the apples as well.

Individual Differences

There is likely more to delay than simply qualities of the situation and the action involved. Qualities of the individual may determine or interact with these situational effects to help predict delay of both action and decision-making.

Individual differences: Procrastination. The earlier discussion of procrastination treats it both as an action and "procrastinators" as a class of individuals. Supporting this distinction, many researchers have treated procrastination as an individual difference (e.g. Effert & Ferrari, 1989; Harriott & Ferrari, 1996; Solomon & Rothblum, 1984; Steel, 2007), referring to those higher in procrastination (those who report procrastinating frequently) or "procrastinators." More recent meta-analytic work has addressed questions about what sort of person a procrastinator is. Steel (2007) discusses the trait-like status of procrastination: People tend to respond similarly over time to procrastination scales. However, research and personal experience also tell us that there is a task-based component: Certain activities, especially aversive ones, are more likely to be delayed than others or avoided entirely (Milgram et al., 1988; Solomon & Rothblum, 1984; Steel et al., 2001), whether the decisions are objectively aversive or just aversive to the decision-maker. Steel's (2007) meta analysis also revealed that participant impulsivity, irrational beliefs, lower self-esteem, lower self-efficacy, self-handicapping, depression, boredom-proneness, lower conscientiousness, a larger intention-action gap, and younger age also predict procrastination, whereas others have found evidence suggesting that being a procrastinator might not be so bad: Ferrari and Dovidio (2000) used these individual difference measures to show that those who are more prone to "procrastination" may be using the time gained by delaying action in an adaptive manner, for example deliberating more or gaining more knowledge about the options at hand. In some cases, this may lead to better outcomes in situations where there is sufficient time to complete these evaluative tasks without risking outcome quality. Tice and Baumeister (1994) show that those who delayed experienced less stress overall (because they were entirely unconcerned with deadlines for the first while). However, measurement of procrastination has been overly focused on the negative: One self report scale for procrastination (the Procrastination Assessment Scale - Students; Solomon & Rothblum, 1984) simply asks students how frequently they procrastinate, and validates this scale against procrastinatory behaviors known to be negative and dangerous to the student—namely, putting off class assignments.

In sum, the evidence for being a procrastinator is generally negative, but not wholly so. It is important to remember, however, that these previous studies are all approaching or measuring procrastination as an inherently negative trait: In the above-mentioned meta-analysis, Steel (2007) notes in his work that, "...the positive form of procrastination...is secondary in usage. The focus of [Steel's] article is on the primary negative form of procrastination" (p. 66). This is understandable, as the positive side of procrastination is delaying an action for a good reason! Things done "for good reason" are also less interesting to study from a psychological perspective, as most psychological research seeks to identify the causes of behavior, and a "good reason" may be viewed as a highly plausible cause, requiring little further investigation. Further, in the absence of a reason to delay acting, it is hard to argue for delay: The further into the future we look, the less certain things become, and if a task must be accomplished and can be accomplished now, acting now reduces uncertainty regarding our ability to complete the task later or our inability to predict what might stand in our way (Buehler et al., 1994). Given this, "trait procrastinators," or those who procrastinate compulsively or avoid acting if possible as a rule, are likely to suffer ill effects of this inaction. This does not make a case against delay; rather, it makes a case against thoughtless or compulsive delay.

Individual differences: Decisional delay. Much as the delay of decisions can be distinguished from the delay of action in general, individual differences in propensity to delay making decisions differ from individual differences propensity to procrastinate overall. Research on "decisional procrastination" and individual-difference measures of indecisiveness has indeed been conducted, though to date without addressing the distinction between wise and foolish delay. First, there have been several measures used to determine how indecisive or prone to procrastinating on decisions individuals are. Among these, Frost and Shows (1993) developed an indecisiveness scale as a diagnostic measure of obsessive-compulsive disorder. However, their scale used items that were either negatively valenced (e.g., "I try to put off making decisions," "I do not get assignments done on time because I cannot decide what to do first"), positively valenced and reverse-coded (e.g., "I find it easy to make decisions"), or related to anxiety or obsession rather than a propensity to delay (e.g., "I become anxious when making a decision," "after I have chosen or decided something, I often believe I've made the wrong choice or decision"), suggesting that the scale is unlikely to measure any variance in decision delay which would be considered adaptive or positive. Frost and Shows show that their scale has consistent predictive validity for non-clinical levels of obsessive-compulsive disorder, and also show that individuals higher on their scale are also higher on general measures of (maladaptive, negatively valenced) procrastination (the Procrastination Assessment Scale - Students; Solomon & Rothblum, 1984). However, they operationalize "indecisiveness" as the amount of time subjects take to make an immediate decision (e.g., they were shown two articles of clothing and asked verbally which they would prefer without any chance to delay). While this appears to do a good job of measuring sub-clinical decision obsession or choice aversion, these constructs are decidedly different from decisional delay as I have described it. For this reason, the results using this scale are more correctly described as results regarding choice-averse or decision-obsessed individuals; I henceforth refer to the construct measured with this scale as "decision aversion."

Mann and colleagues (1997) describe the Melbourne Decision Making Questionnaire (MDMQ), a scale designed to measure patterns for coping with decisional conflict (described in Janis & Mann, 1977). While this does not explicitly attempt to measure decision delay (except insofar as we would expect decisions higher in conflict to be more aversive, and thus counting as aversive tasks that are more likely to be delayed; Anderson, 2003; Janis & Mann, 1977; Solomon & Rothblum, 1984), the questionnaire has four subscales that hint at parts of decisional delay: "buck-passing" (decision avoidance) and "procrastination," both explicitly negatively valenced; "hypervigilance," referring to the same mixture of obsession and anxiety as Frost and Shows' (1993) indecisiveness scale; and a "vigilance" component, which measures positively valenced items regarding care and deliberation in making decisions, but does not have any explicit time component. For example, one "vigilance" item asks decision-makers whether they take a lot of care or try to collect a lot of information, but does not address whether or not doing so lengthens their decision-making process, while the "procrastination" items ask about delay from an explicitly negative standpoint. Though their factor analysis provides evidence that these four structures are independent, they show rather strong between-subjects correlations (*r*'s between .72 and .78 for the three negatively valenced subscales, but much lower *r*'s between -.17 and -.30 between the vigilance subscale and the other three). Though this suggests that there may well be two components of delaying choice, one positive and one negative, the literature on decisional procrastination using these scales has used only the "procrastination" subscale (see Harriott & Ferrari, 1996; Ferrari & Dovidio, 2000; Patalano & Wengrovitz, 2008), while the "vigilance" subscale has remained basically untouched.

Finally, Kuhl's (1981) action vs. state orientation scale has also been used to predict the delaying of decisions: Those who are oriented towards action are likely to do things (such as make decisions), while those who are oriented towards the state of the world as it is are more likely to work to (or prefer to) maintain the status quo (Diefendorff, Hall, Lord, & Strean, 2000). The dichotomy drawn here is interesting, as maintaining the status quo may occasionally require action (for example, politicians who would seek to maintain the status quo often must argue and lobby vociferously), though in most real-life contexts it does not. Manipulating individuals to be more action-oriented has been shown to reduce “inaction inertia” (Anderson, 2003), a tendency for people not currently acting to avoid making decisions (van Putten, Zeelenberg, & van Dijk, 2009), though it has not been used in other decisional contexts.

Using the aforementioned scales, however, most researchers have studied what they call “decisional procrastination,” or maladaptive indecisiveness. Taking the procrastination literature and applying it to a “maladaptive pattern of postponing a decision” (p. 127), Ferrari and Dovidio (2000) showed that those who detrimentally put off making decisions may not simply be disorganized or choice-fearing, but rather that they may engage in longer, more specific information-seeking strategies, suggesting that they may hold themselves to a higher standard of certainty (or be more uncertainty-averse) than others. However, their discussion was speculative and they did not provide any empirical evidence addressing the idea that decision delay may be non-negative or potentially positive. Indeed, the literature still lacks a judgmentally positive (or even neutral) approach to decision delay.

Other individual differences. Other scales pertaining to differences among individuals may relate to propensity to delay choice, to good or bad effect. For example, propensity to seek to achieve the best possible goal (to “maximize” outcomes) rather than to accept a sufficient “good enough” goal (to “satisfice”) has obtained a lot of attention recently. Although it dates back to Simon (1955), Simon never discussed this tendency as an “individual difference,” or scale on which anybody may fall anywhere. Rather, Simon conceptualized satisficing as a feature of human cognition, generating the study of bounded rationality, referring to a decision strategy that may or may not be used or not in a given situation and illustrating some theoretical situations in which satisficing would be the only rational strategy. The newer literature conceptualizes maximizing/satisficing as a bipolar scale representing a continuous trait characteristic, and suggests that individual differences on this trait may affect life satisfaction and quality of decisions (Schwartz, Ward, Monterosso, Lyubomirsky, White, & Lehman, 2002). Schwartz and colleagues developed this scale to measure the extent to which people maximize (as opposed to satisfice), though there is also recent research suggesting that many of these overarching effects may be overstated or artifactual due to flaws in the scale: Diab, Gillespie, & Highhouse (2008) address these psychometric issues (including poor item choice and unreliable alphas) cogently, and propose an alternative scale that more purely measures the tendency to maximize.

The maximizing/satificing dimension (measured via whatever scale) is potentially related to the incidence of delay in optimal stopping tasks: A tendency to seek the single *best* option suggests that these individuals would be utilizing distributional information in a different way, looking for an option above all other options rather than above some distributionally defined cutoff. In other words, maximizers[1] may delay much longer, achieving minimal (but potentially reliable) gains relative to satisficers—but also suggesting that the potentially extreme added effort may not be worth it. This is addressed briefly by Patalano and Wengrovitz (2007), who used Schwartz and

---

[1] I frequently will refer to "maximizers" or "satisficers" throughout this paper for ease of expression; in all such cases, I am referring to those who are higher or lower, respectively, on a continuous maximizing-to-satisficing scale.

colleagues' maximizing scale (Diab and colleagues' critique and scale was unpublished as of that time), and suggest that decision-averse individuals are best conceptualized as satisficers with a very high "satisfactory" cutoff, while maximizers are more decisive, effectively determining that they have found the "best" option even when many options remain unexamined, leading them to choose sooner and thus occasionally end up with paradoxically worse outcomes than the satisficers. Patalano and Wengrovitz did not, however, attempt to test this rather convoluted explanation empirically.

Finally, because the notion of distribution learning is inherently numerical (being based on statistics and probability), there is reason to believe that those who are more enamored of numbers may show faster learning and more nuanced use of the information provided. Peters and colleagues' (2006) numeracy research suggests that some individuals are not only better at mathematical tasks, but also take pleasure in the processing and manipulation of numbers. In tasks which lack a clear "right answer" or even a clear indication of what makes a choice "good" (for example, multiattribute decisions for which it is not clear how attributes should be weighted against each other; c.f. Timmermans, 1993), the mathematical complexity of determining which options are better than the others may lead more numerate individuals to make choices based on more of the available options or to utilize more information in general when making choices.

Summary of Hypotheses and Prior Work

The question of how long people will delay decision-making in order to learn about a distribution is hypothesized above to depend on how variable the options available are. Support for this hypothesis is drawn from numerous studies (e.g., Botti & Iyengar, 2004; Corbin et al., 1975; Dhar, 1997; Patalano & Wengrovitz, 2007; Tversky & Shafir, 1992) that show effects that could be explained as extrapolation from the distribution of available or already-viewed options, to a population of potential options. Other research on procrastination has lacked a judgmentally neutral approach to delaying actions (and, by extension, decisions): I plan to extend this work by providing evidence that delay of decisions is, in some cases, useful or adaptive. Individual difference research into propensity to delay and decision-making style suggests that the tendency to delay choice and the efficacy of doing so should vary across people.

Overview of Studies

The three studies described here take a similar format: Decision-makers participate in an optimal stopping problem, and I examine how good their choices are, how soon they make their choices, and how individual differences moderate these effects. In Study 1, I ask participants to choose, in an optimal stopping context, the number of minutes they will spend completing math problems (as opposed to having "free time" online) during the second half of the study. I systematically vary the mean and variance of the distribution of options in a 2x2 design: high versus low mean across the options of how many minutes of free time participants will get, crossed with high versus low variance in minutes of free time. In addition, I collect subjective ratings from participants about the quality of each option before the option is chosen or rejected. Behavior is evaluated in terms of the quality of outcome achieved (both via ratings of the option, but also using the objective number of minutes "received" to spend freely online), how much choice was delayed (how many options were viewed), and sensitivity to the distribution (the extent to which subjective ratings predict objective ratings and the slope at which this sensitivity increases with more data), as well as individual difference variables.

In Study 2, I replicate and extend Study 1 by varying the distribution itself (presenting each subject with a particular random distribution rather than using one distribution for all participants), examining four decisions per subject (asking each subject to make a choice in each of the four mean and variance conditions, in order to replicate Study 1's results within subjects), and adding a cost to delay: Participants are either rewarded for choosing efficiently ($100 will be given to the single participant who made the "best" choice, weighting time and quality equally) or punished for choosing inefficiently (1% of the stated free time was discounted for each option they choose to view).

In Study 3, I utilize a more "real-world" vignette in order to extend these results into a multi-attribute framework: I have participants consider a potential roommate, who varies in terms of personability (how well they will get along with the potential roommate) and financial stability (how easily the potential roommate can afford rent). I examine sensitivity in this case by analyzing at the extent to which participants use the two cues (personability and financial stability) and examine individual differences in the extent to which participants use one or both cues.

CHAPTER II

STUDY 1

The purpose of the first study is to evaluate, as cleanly as possible, whether participants are sensitive to qualities of the option set, whether they delay decision-making to gain knowledge of these qualities, and whether sensitivity to these qualities leads to better decision outcomes. For this reason, the study uses an easy-to-understand unidimensional quality metric: in this case, time.

The basic design of this study is to have decision-makers complete an optimal stopping problem in which they determine how much of the time spent participating in the study will be spent having access to a computer and free time online, as opposed to more structured, less enjoyable activities in the post-decision part of the study. Thus, in the study, the "options" among which participants were asked to choose, were each a number of minutes of free online time, with the knowledge that the remainder of their time in the study would spent on the other activities, adding up to a total of 30 minutes. Two variables were manipulated: The mean of the amount of free time across the options (either a low or high number, corresponding to a mean of either 21 or 11 minutes of free time across the options) and the variance of the amount of free time across the options (also either low or high, corresponding to 2.65 or 10.59 squared minutes). The distribution of options was a single random sequence of "study time" quantities, shown in

Figure 1; this set was drawn from a random-normal distribution and then shifted, in order to guarantee that the relative worth of subsequent options was held constant. As such, the shape of **Figure 1** is the same for all four conditions; conditions vary only in terms of the scale and range of the y-axis.

Figure 1. Graph of option quality as a function of order.

Hypotheses

In this study, I address two basic questions regarding the manipulated variables: Do people delay decisions in order to learn what counts as a "good" choice, and does delaying in this manner lead to better choices? I also address several hypotheses regarding individual differences in decisional delay for these options. My hypotheses are outlined below, and summarized at the end of this section.

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable

To test whether knowledge of the choice set is desirable, I predict that the quality of an option (including "objective quality," or the amount of free time offered to participants, as well as "subjective quality," measured as a Likert-scale rating of how "good" or "bad" participants believed the option to be) alone will be insufficient to predict choice. In other words, the quality of the options should not fully mediate the relationship between the number of options seen and the likelihood of choice: If participants were uninterested in seeking distributional information, they should choose based only on the quality of options, and the raw number of options viewed would not be significant once controlling for quality. If, on the other hand, number of options seen is a significant predictor of choice even after quality metrics are included in the model, participants are waiting for something other than simply the quality of the option. This would support the hypothesis that people seek to learn about the options available to them as part of the decision-making process. In sum, I predict that quality metrics will not fully mediate the relationship between number of options seen and likelihood of choosing an option.

Similarly, I predict that there will be no difference between the "high" and "low" mean condition when predicting the amount of time participants spend deciding. In other words, whether the average amount of free time is high or low should not affect how many options are seen. This is a more direct test of the meditational hypothesis above: If the quality of options is systematically varied while the relative quality of options is held stable, a quality-driven decision process would predict quicker choice in the "high mean" condition. If, on the other hand, choice is not based on raw quality (for example, if it is based on the quality of each option relative to the set of options available), then the two mean conditions may not differ.

I conversely predict that the variance manipulation *will* predict choice delay: Participants in the high variance condition will take longer to choose than those in the low variance condition, regardless of the mean quality. If people are sensitive to their knowledge of the distribution of options and prefer to gain this information before deciding, then a more variable set of options should lead participants to feel as if they did not understand the distribution as well. Because the options shown in the two variance conditions will have the same relative distribution (the same rank order of options as well as the same standardized difference between options), a preference for viewing more options in the high variance condition should indicate that peoples' desire to learn about the distribution is based on a known truth of statistical inference: Inference (such as inferring that the offered option is optimal) is less precise or trustworthy when there is high variance. Thus, I predict that there will be no difference in length of delay between the high and low mean conditions; however, there will be an effect of variance condition.

Finally, I predict that the number of options seen will interact with the quality (both subjective and objective) of the option, such that seeing more options causes quality to become a better predictor of an option being chosen. I predict this because, as distributional knowledge is gained, participants should gain a better understanding of what "option quality" really means—they will learn what counts as a "good" choice for the set of options available to them. As they learn what counts as a good choice, the quality of the options should become a stronger (i.e., more positive) predictor of choice.

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful

To test whether knowledge of the choice set is useful, I predict that the subjective ratings decision-makers provide will, with the viewing of more options, become a more accurate measure of the viewed options' objective qualities (number of minutes of free time). In other words, I predict that as participants gain more knowledge of the choice set (by simply viewing more options), their subjective ratings of option quality will show increased correspondence to the options' objective quality. I also predict that the amount of delay before choosing will show a "diminishing returns" relation with decision-quality such that delay will predict higher quality at first, but as more and more options are viewed, additional delay will not continue to lead to yet-better decisions being made. In other words, I predict a positive linear effect of delay on the quality of the eventually chosen option, but only in the context of a negative quadratic effect.

In addition, I predict that subjective ratings (the measure of how good options "seem" to participants) will show a relationship to *relative* objective ratings, over and above the effect of raw objective ratings. In other words, while I expect that subjective ratings of options will be related to the objective values (equal to the number of minutes of free time), I expect a further effect to be present when looking at *relative* objective value, or how objectively good an option is *relative to the other options seen so far.* Such an effect would reinforce the argument that distributional information is useful, showing that, within the context of what is known about a distribution, our affective responses to options depend on their contextual meaning (is this "good" because it is better than most of the options seen?) over and above the absolute objective meaning (is this "good" because 17 is a good number of minutes of free time?).

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information

To test individual differences in delay, I predict that those who are averse to making decisions will be more "extreme" in their decision-delaying behavior (i.e., will be further from the mean). I will count how many options are viewed by each individual, and then examine both the raw number (the total amount of delay) and the mean absolute deviation (how far an individual's amount of delay is from the average amount of delay), in order to simultaneously capture those who make their decisions hastily or perseverate unnecessarily. I expect the decision-averse to be higher on the latter variable (absolute extremity of delay), but nonsignificant on the former variable (raw quantity of delay). This pattern of results would show that extremity of delay is the factor at work rather than simply picking up on a higher or lower amount of delay.

I also predict that more vigilant decision makers will take longer to choose than the less vigilant, but that this additional delay will not reach the point of diminishing returns. Thus, there should be a positive relationship between vigilance and delay, a positive relationship between vigilance and decision quality, and a positive interaction between vigilance and delay in predicting quality, showing that more vigilant decision-makers make better use of the information they gain. For "maximizers," however, I predict the same positive relationship between maximizing and delay (and the corresponding positive relationship between maximizing and quality), but either no interaction or a negative interaction between maximizing and delay in predicting quality, showing that maximizing does not lead to better outcomes outside of the fact that maximizers delay longer (or, if negative, that maximizing blunts the effects of delay on decision quality).

To test differences in sensitivity to the distribution, I further predict that vigilance and satisficing scores will positively predict correspondence to the true distribution—in other words, subjective ratings of the quality of decisions will be a stronger predictor of relative objective quality (defined above) for those who are higher on vigilance or satisficing. I also predict that decision aversion will be a negative predictor of relative objective quality (i.e., that those more averse to decision-making will show a weaker relationship between objective and subjective quality). Finally, I predict that, overall, people whose subjective ratings correspond more closely to the objective ratings (and thus have a more accurate internalized representation of the objective distribution), conceptualized as an individual difference, will choose better outcomes, both objectively and subjectively.

Summary of Hypotheses

In sum, I hypothesize the following:

1. Gaining knowledge of the option set is desirable.

1a. Quality metrics will not fully mediate the relationship between number of options seen and likelihood of choosing an option.

1b. Delay will not vary between the high and low mean conditions.

1c. Delay will be higher in the high-variance condition than in the low-variance condition.

1d. Decision quality and number of options seen will interact positively to predict how long choice is delayed: Option quality will become a better predictor of choice as participants view more options.

2. Gaining knowledge of the option set is useful.

2a. Subjective and objective quality metrics will become more similar over time.

2b. Delay will lead to better choices, but this effect will taper off eventually.

2c. Subjective quality will be related to the relative objective value of proffered options, over and above the extent to which ratings track the absolute objective values.

3. Individuals will differ in how the utilize distributional information.

3a. People who are averse to decisions will have more extreme delaying behavior than those less averse to decisions.

3b. Vigilant decision-makers will delay longer than less vigilant decision-makers, and delay will be a better predictor of outcome quality for those who are more vigilant.

3c. Maximizers will delay longer than satisficers, but delay will not result in especially better decisions for maximizers than satisficers (and in fact, delay may be a better predictor of outcome quality for satisficers).

3d. Decision-makers who are more vigilant, less maximizing, and less decision-averse will produce subjective ratings that are closer to the relative objective values of the option at hand than the less vigilant, more maximizing, and more decision averse, respectively.

3e. People whose subjective ratings more closely track the relative objective values of the option at hand will obtain better outcomes.

Method

Participants

Participants were 138 college students enrolled in a Psychology or Linguistics class at the University of Oregon, who participated as part of their coursework. Participants signed up for the study via an online study-management system that was designed to reduce selection bias by keeping participants blind to the subject of study. The Human Subjects pool is comprised of students participating for credit or extra credit in psychology courses; most students in the pool are in an introductory course. The pool is predominantly female (58-64% across the terms analyzed in these studies), white (77-78% Caucasian), and college-aged (M=20.07 years old, sd=3.32).

Procedure

Participants were told that they would be participating in an experiment on decision-making. The basic study layout was described to them as part of the consent process. Participants were then told that they would decide how to spend the second 30-

minute half of the study: They would see a series of options, presented on a computer one at a time, indicating how many minutes of "free time" they would have to use the computer for their own purposes (for example, checking Facebook, playing computer games, or surfing the web), and were told that the other time (30 minutes, minus the number of minutes they chose) would be spent piloting math problems for a study that might run in a later term (which was just a cover story). This decision process was an optimal stopping task: Participants were told (truthfully) that at any point they could select an option offered to them, or they could reject it and move on to the next option. Participants were also told that they had up to 100 options to view, and that if they did not choose one of the first 99, they would be assigned the 100^{th}. When they chose an option, the decision would be over, and they would later complete the decided-upon amount of free time, followed by the remainder of the 30 minutes working on math problems, and then they would be free to go. Participants were told that they would also be required to provide a subjective rating of how good they thought the option they chose was on a 7-point Likert scale with anchors labeled "very", "moderately", and "slightly" good or bad (depending on which side of the scale they were on), and "neither good nor bad" as the central option's anchor. This process was displayed graphically and rated on a computer, as in the example shown in **Figure 2**.

When participants were shown the example page, the computer also assigned them to one of the conditions (low/high mean, and low/high variance) in a quasi-random manner designed to ensure balance (i.e., they were assigned randomly to a condition selected from the set of conditions with the fewest participants thus far). Participants were not aware that they had been assigned to a condition, nor were they told at all that the options they saw were generated in any specific way. The research assistant running the survey was also blind to the participant's assigned condition.

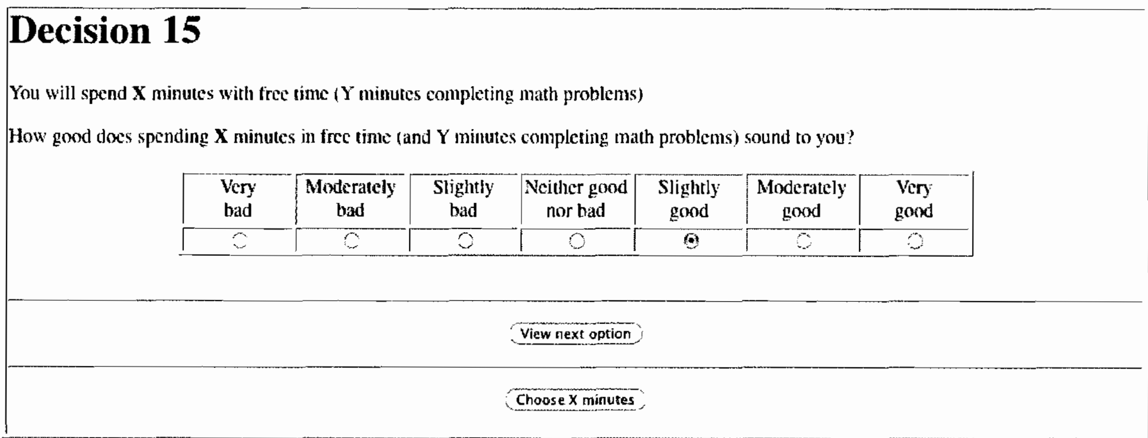

Figure 2. Participants' view of the decision they were making in Studies 1 and 2.

Participants then began the decision-making task as described. After completing the decision-making task, the computer told participants to alert the researcher. The researcher then noted how much free time they had chosen (it was displayed on the screen at this point), but told them that before it began, they would first complete a series of questionnaires, also administered via computer. These questionnaires[2] were:

1. Schwartz and colleagues' (2002) maximizing/satisficing scale, a 13-question scale answered via 7-point Likert response items anchored at "strongly disagree" and "strongly agree." Schwartz and colleagues

---

[2] Alphas reported here are the ones reported in the previously published and cited papers; alphas observed during my studies are reported in the Results sections, below.

report α = .71 for this scale, though Diab and colleagues (2008) report α = .68.

2. Diab and colleagues' (2008) maximizing/satisficing scale, a 9-item scale using the same response format as Schwartz and colleagues (2002), designed to more precisely measure maximizing/satisficing tendency. As this scale and Schwartz and colleagues' scale use the same response items and share some actual items, these scales were presented as one scale to participants to save time and eliminate redundancy; α = .80.

3. Frost and Shows' (1993) decision aversion scale, a 15-item measure using 5-point Likert scales anchored at "strongly disagree" and "strongly agree," designed to measure the extent to which people are "indecisive" or choice-averse; α = .87.

4. Mann and colleagues' (1997) Melbourne Decision-Making Questionnaire, a 22-item scale with four subscales that measure decision-making styles: Vigilance (6 items, α = 0.80), Buck-Passing (6 items, α = .87), Procrastination (5 items, α = .81), and Hypervigilance (6 items, α = .74). The original scale was designed with three-option scores of "true for me," "sometimes true," and "not true for me," though in the current study, scores were computed as the sum of items that the respondent checked as being "true for me."

After completing these measures, participants then had their "free-time," which was timed by a research assistant who then brought them to a separate webpage for the "math section." The first page of the math section was in fact Lipkus, Samsa, and Rimer's (2001) numeracy scale (a 10-item scale with both multiple-choice and free-response sections designed to measure mathematical skill; $\alpha = .78$; see also Peters et al., 2006), which made sense to include in the "math section" as the scale is mostly numerical questions such as, "If Person A's risk of getting a disease is 1% in ten years, and Person B's risk is double that of A's, what is B's risk?" The scale was scored in terms of how many responses were correct. This scale was kept separate from the other individual difference measures described above, so that participants would not be confused or believe that they had started the math portion of the study during the administration of the other individual difference measures. They then completed a series of math problems taken from online example SAT tests (which were not scored). Upon completing the full 30-minute free time and math section, participants were debriefed.

Results

The *lme4* package (Bates & Maechler, 2010) for the R package for statistical computing (R Development Core Team, 2010) was used for hierarchical analyses. This package allows analysis of mixed effects models, fit to both linear (commonly called a "hierarchical linear model" or HLM) and logistic distributions. The data produced in Study 1 included several observations for each decision-maker. For example, each decision-maker made a choice to choose or reject each option that he or she saw, constituting a within-subjects design with a varying number of observations per subject. For this reason, participant's subject ID is treated as a random variable.

An unexpected effect was detected during exploratory data analysis: The relationship between a participant's ratings of an option's relative objective value and its relative subjective rating was in fact "X-shaped" (see Figure 3, a density-enhanced scatterplot of each decision made, not accounting for dependencies within subjects), indicating that for some participants the relationship between objective and subjective value was negative and for others it was positive. This was due to an individual difference in this relationship found across subjects: Among the correlations between objective value and subjective rating for subjects, 33% of these correlations were in fact negative, showing an effect indicating that about a third of the participants preferred options involving *more* math rather than less; this behavior was confirmed by the research assistants that who conducted the study, who noted that some people appeared to have chosen very low quantities of free time. In a subsequent study with a similar design (Study 2, described in the next chapter), participants were queried about this preference; some participants mentioned wanting to "help out" presumably by completing more of the pilot math problems. To account for this effect in further analyses, the objective values for those participants who showed this negative correlation were reverse-scored by reflecting the value around the mean for the given distribution (high vs. low mean).

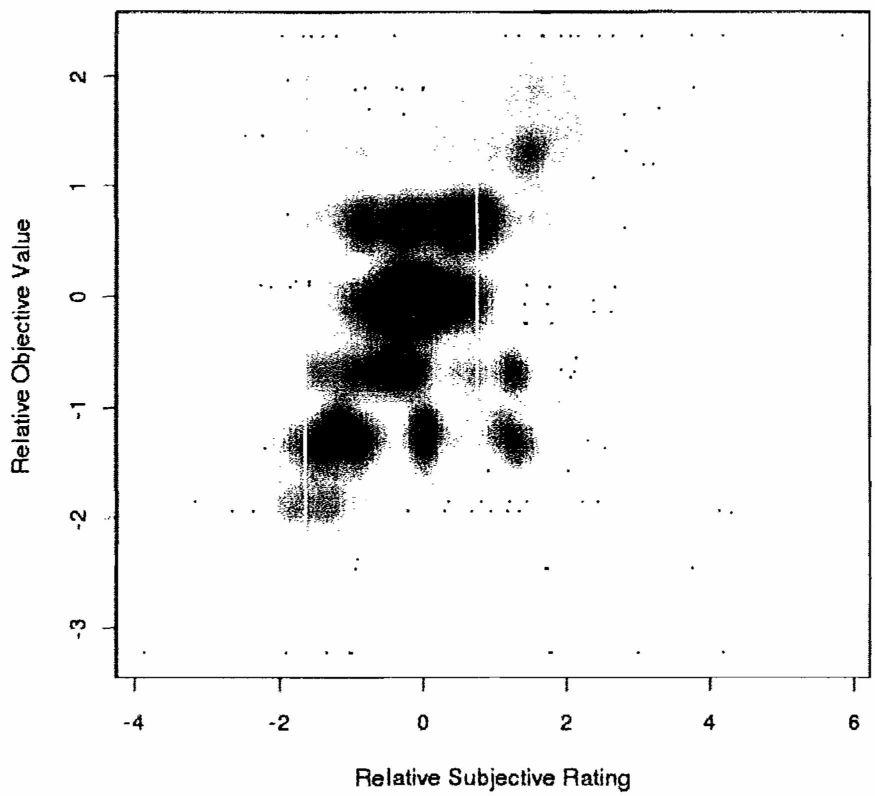

Figure 3. Illustration of the X-shape found between subjective and objective value.

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable

These hypotheses regard the extent to which holding and gaining knowledge of the choice set is desirable, thus making the obtainment of this knowledge a reason for delaying choice. Hypothesis 1a was that option quality alone would be insufficient to predict choice. A model in which option quality alone is sufficient to predict choice would be consistent with finding that the number of options already seen (hereafter, "delay") would not significantly predict whether or not an option is chosen in a model that already contained quality as a predictor. This hypothesis was tested using a model that included both objective quality (number of minutes of free time provided) and subjective quality (Likert-scale rating of the option). A hierarchical logistic model predicted whether an option was chosen or not and included fixed effects for subjective quality, objective quality, and the number of options already seen, with a random intercept for each decision-maker.[3] This model showed a significant effect for subjective rating, $B = 1.46$, $z = 9.81$, $p < .001$, as well as number of options already seen, $B = 0.04$, $z = 5.16$, $p < .001$, but no effect for the objective value of the option, $B = -0.04$, $z = -0.72$, $p = .47$. In this context, it appears that the worth of the objective value of the object was (unsurprisingly) not predictive of choice once the subjective rating was accounted for. However, the fact that delay was also predictive of choice (over and above the effects of the option's subjective and objective quality) supports Hypothesis 1a.

Hypotheses 1b predicted that participants in the low (versus high) mean condition would not differ in amount of delay, whereas Hypothesis 1c predicted that participants in the low (versus high) variance condition would show less delay. To this end, I used the same modeling strategy as above, but included a main effect for the "mean" condition and a simple effect for the "variance" condition within both mean cells. The simple-effects approach was taken rather than using a 2x2 interaction strategy because these effects are not expected to interact in a synergistic fashion: If there is an effect of mean, it

[3] Because there is only one option that is eventually chosen in these tasks, it is inappropriate to estimate other random effects within subjects. In this case, estimation of the intercept controls for how likely a given option was to be chosen, based on the number of options that were eventually viewed.

is of interest to examine the effect of variance separately. Objective and subjective quality and the number of options seen were also included as controls. As predicted, high versus low mean condition was not a significant predictor of choice, $B = 0.40$, $z = 0.84$, $p = .40$. Variance, conversely, was a significant predictor of choice in the "low mean" condition, $B = 1.95$, $z = 2.72$, $p = .007$, indicating that participants in the low variance condition were about 14% more likely to choose. In the "high mean" condition, there was no effect of variance ($B = 0.08$, $z = 0.12$, $p = .90$). When the nonsignificant effect of mean was dropped from the model, variance remained a significant predictor overall, $B = 1.00$, $z = 2.02$, $p = .04$. In sum, Hypotheses 1b and 1c were largely supported, showing that the amount of time participants delayed choice was affected by the variance of the options available to them, but not by the mean level of objective quality.

Hypothesis 1d predicted that quality and delay would interact to predict whether an option was chosen. Because I expected subjective and objective quality to be collinear, I tested Hypothesis 1d in two separate models, one for subjective quality and one for objective quality, each of which contained a random intercept for each participant. There was no interaction between delay and objective quality in predicting choice, $B = -0.001$, $z = -1.16$, $p = .25$, but there was a significant interaction between delay and subjective quality in predicting choice, $B = -0.01$, $z = -3.93$, $p < .001$. However, contrary to prediction, the interaction was negative, not positive. This indicates that as people learn more about the option set, their subjective values become worse predictors of choice. In other words, as decision-makers delay more, higher-valued options are less likely to be chosen, rather than more likely; conversely, this could indicate that earlier in the decision-making process, higher-valued options are more likely to be chosen.

As an exploratory analysis, based on the theory underlying Hypothesis 2b regarding how efficacy of delay is expected to "taper off" over time, I added a term for the square of delay to the model, which as well as the interaction between quality and the square of delay. The resulting model showed the hypothesized positive interaction between delay and objective quality, $B = 0.009$, $z = 2.43$, $p = .01$, and a negative interaction between the square of delay and objective quality, $B = -0.0001$, $z = -2.82$, $p = .005$, indicating that the objective quality of an option becomes a stronger determinant of choice as delay increases, but that this effect tapers off in the long run. The same pattern of results was found when predicting subjective ratings, with a positive interaction between the linear effect of delay and subjective quality, $B = 0.05$, $z = 3.21$, $p = .001$, and a negative interaction between the quadratic effect of delay and subjective quality, $B = -0.0006$, $z = -3.99$, $p < .001$. While this pattern of results was not explicitly hypothesized, it does provide one explanation for not finding the initially predicted relationship, and supports the conceptual (non-operational) hypothesis.

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful

The second broad hypothesis was that there is utility in gaining information about the option set. The first hypothesis in this vein, 2a, sets the groundwork for this: As people see more options, their subjective ratings of the options they see should become a better predictor of the objective quality of the options. As such, I hypothesize a positive interaction between subjective quality and delay in predicting objective quality. This was tested using an HLM predicting objective quality from fixed effects of subjective quality, number of options seen, and the interaction between subjective quality and number of options seen, with a random intercept for each subject as well as a random slope for both subjective quality and delay. This interaction was significant, $B = 0.003$, $t = 2.17$, $\chi^2(1) = 4.71$, $p = .03$.[4]

Hypothesis 2b predicted not only that delaying would improve choices, but that eventually this effect would taper off. Operationally, this hypothesis is that delay will show a positive linear effect, predicting higher-quality choices, but a negative quadratic effect, indicating diminishing returns. This hypothesis was tested in two separate regression models, one predicting the chosen option's objective quality and one predicting subjective quality. Because the high and low mean condition had different expected values, this effect was modeled separately within the high and low mean condition (i.e., as a simple effects analysis). For objective quality, the predicted effects were present, with a marginal positive linear effect of delay in the "high mean" condition, $B = 1.32$, $t(85) = 1.85$, $p = 0.07$, a significant positive linear effect of delay in the "low

---

[4] Because there is some debate on the accuracy of estimating degrees of freedom in hierarchical models such as this (Baayen, Davidson, & Bates, 2008), I report the *t*-value and a *p*-value inequality such as $p < .05$, rather than a precise *p*-value. When the *p*-value is close to .05, *p* is reported as marginal, $p < .10$. In some cases, when precise *p*-values are necessary, I will report either the significance of the χ^2 log-likelihood ratio test of whether the term in question significantly reduces the deviance when added to the model or the bootstrapped highest posterior-density implied *p*-value using the *languageR* package (Baayen, 2010; Baayen et al., 2008).

mean" condition, $B = 1.51$, $t(85) = 2.60$, $p = .01$, a marginal negative quadratic effect of delay in the "high mean" condition, $B = -0.001$, $t(85) = -1.63$, $p = .10$, and a significant negative quadratic effect of delay in the "low mean" condition, $B = -0.002$, $t(85) = -2.79$, $p = .01$. For subjective quality, the same basic structure replicated: For the low mean condition there was a significant positive linear effect, $B = 0.06$, $t(85) = 2.37$, $p = 0.02$, and a significant negative quadratic effect, $B = 0.0008$, $t(85) = -3.13$, $p = .002$. For the high mean condition the linear effect was nonsignificant but positive, $B = .05$, $t(85) = 1.33$, $p = .19$, and the quadratic effect was negative and marginally significant, $B =-.0005$, $t(85) = -1.67$, $p = .10$.

Hypothesis 2c predicted that subjective quality of each option would be related to the *relative* objective quality of an option over and above the objective quality of the option. Objective quality in this analysis as in earlier analyses is the raw number of minutes of free time the participant would experience. Relative objective quality, however, requires a transformation of objective quality. This variable was computed as a z-score representing the objective quality of each option shown to participants, scaled relative only to options the participant had already seen. So, if m_i is the mean of objective quality for items seen, s_i the standard deviation of quality for items seen, and x the absolute objective quality of the ith option a participant sees, then x has relative objective quality given as $(x - m_i)/s_i$.

This hypothesis was tested using an HLM model predicting subjective quality from relative objective quality and "absolute" objective quality (plus the mean and variance condition as covariates), allowing each person to have a random intercept (their own average subjective quality, or scale effect) as well as a random slope for relative and absolute objective quality (their own relationship between these variables and their subjective rating). I only examined ratings for the third and subsequent options seen by participants, as the first two options had the same absolute quality (meaning there was zero variance for the first two options, so relative quality could not be computed). This model produced a significant effect for both relative objective quality, $B = 0.34$, $t = 5.00$, $p < .001$, and absolute objective quality, $B = 0.19$, $t = 7.59$, $p < .001$, supporting Hypothesis 2c.

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information

Tests of hypotheses of individual differences took one of three forms: predictions of decision delay, predictions of decision quality, and predictions of distribution sensitivity (defined as the correlation that each individual shows between subjective quality and relative objective quality, transformed into a *z*-score using Fisher's method). Initial analyses suggested that the scales were functioning as expected: Alphas for all scales exceeded .60. Lower alphas were shown by the MDMQ scales (vigilance alpha=0.69 and decisional procrastination alpha=0.75; these were implemented as dichotomous checklist scales rather than the three-valued scales used by Mann and colleagues, 1997, so lower alphas are to be expected in my study). The Schwartz and colleagues (2002) maximizing scale also had a low alpha of .61, which is more consistent with Diab and colleagues' (2008) report of the Schwartz scale than Schwartz and colleagues' report. Similarly, these scales correlated with each other in the expected manner: The Diab and Schwartz scales, which both measure maximizing, correlated at $r = 0.53, p < .001$. Diab and colleagues report a correlation of 0.48 between these scales; this is not significantly different from observed correlation in Study 1, $z = 0.59, p = .55$. The Frost and Shows (1993) indecisiveness questionnaire correlated with the MDMQ decision procrastination scale at $r = 0.58, p < .001$. Patalano and Wengrovitz (2007) reported $r = .66$ which does not differ from the current data, $z = -0.83, p = .40$.

However, these individual difference measures were almost wholly unrelated to the behavioral measures collected in this study (amount of delay, subjective quality, objective quality, and sensitivity to the distribution). While a few significant or marginal results were found, these results became insignificant when accounting for family-wise type-I error. Furthermore, when they were found, they were found in only one of two measures of a construct (e.g., were significant for the Schwartz maximizing scale but not the Diab scale); these results are thus not considered reliable and thus not reported.

Discussion

Study 1 provided broad support for Hypothesis 1 and Hypothesis 2, supporting the idea that people delay choice in order to gain information about the distribution of options (Hypothesis 1), and that doing so helps them achieve better outcomes (Hypothesis 2). These findings were demonstrated in several ways: The objective quality of a choice was insufficient to predict choice—those who had viewed more options were also more likely to choose (Hypothesis 1a). This suggests that people are in fact "holding out" for more information or better options: It appears that as people gain knowledge of what the option space is like, the quality (either objective or subjective) of the option becomes more relevant (Hypothesis 1d). When option sets are more variable, people also prefer to view more options before choosing (Hypothesis 1c). This result was significant only in the "low-mean" condition, however, which is puzzling. This could have occurred for several reasons: The first is that, for the high-mean condition, several of the options were effectively "at ceiling": Because there were only 30 minutes of free time to be had, options such as 28 and 29 might have been considered "ideal" in the high-mean condition, rendering the variance of the option set much less relevant. People often took longer to make a choice when they saw more variable options (Hypothesis 1c), even when the means and relative differences among the objective values of the options remained the same, further suggesting that people "hold out" for more information when they are more unsure about the range of options, again supporting the idea that people seek information about distributions of options.

Hypothesis 2, that people are in fact learning the distribution better as they view more options, also found broad support. For example, as people viewed more options, their subjective ratings also grew closer to the objective value of the items they saw (Hypothesis 2a): People aren't just "waiting around" to view more choices just because people prefer to wait or enjoy procrastinating, as was suggested by Tykocinski and Ruffle (2003) and Ariely and Wertenbroch (2002). Instead, decision-makers' subjective ratings were becoming more "attuned" to the distribution shown to them as they saw more options (Hypothesis 2c), as evidenced by their subjective and objective ratings growing closer together as participants viewed more options, indicating that the objective values were effectively being "internalized" into decision-makers' subjective ratings.

Gathering distribution information and becoming better-attuned to the distribution that one has viewed also led people to make better choices: People who delayed longer were shown to make better choices (Hypothesis 2b), both in terms of their subjective ratings (people rated the option they chose higher) and the relative objective value of the options they have seen (people chose options that provided relatively higher numbers of minutes of free time). This is not a linear effect, however: Viewing more options eventually shows an effect of diminishing returns, with increased viewing of options leading to better choices at first, but tapering off as more and more options have been seen. Of course, my ability to detect this point of diminishing returns is based on the fact that some people reached this point: Why some people reach this point of diminishing returns and still continue to view more options is a question for future research. One possibility is that it is an artifact of the single distribution used for this study: The distribution in question may have a "point of diminishing returns" that occurs prior to when participants feel that they have gathered enough contextual information to make a choice. Another possibility is that this is necessarily the case: In order to induce that the distribution in question is normal, participants would need to see data consistent with a normal distribution: most options being similar, with a few being much higher or lower than the mean. To discover that a few options are higher or lower, of course, one must see some "higher" options and not take them; when they come up again, they are unlikely to be very superior to those already seen, thus indicating that the point of diminishing returns has been reached.

As predicted by Hypothesis 2c, peoples' subjective judgments of whether options are good or bad depend on relative value metrics as well as absolute metrics: When participants viewed many options, the relative value of those options (i.e., the standardized value relative to the options seen) predicted subjective ratings over and above the absolute or "raw" value of the options (that is, the number of minutes of free time). This is consistent with a wide body of literature: For example, the classic "anchoring and adjustment" effect (Tversky & Kahneman, 1974, and many subsequent papers) shows that when one number is provided to decision-makers (even a completely irrelevant number), that number partially determines estimates of quantities provided by participants later. The results supporting Hypothesis 2c also support the literature regarding how judgments of value and quality depend on a reference point (e.g., are "relative;" Kahneman & Tversky, 1979).

However, this behavior should not be considered an "error." It is entirely appropriate, as each additional option provides new information about the relative worth of all prior options (c.f. Dawes, 1990 for a similar argument regarding a different judgment process). Given support for Hypothesis 2c, future research should be wary of paradigms that show "preference shifts" when seemingly irrelevant options are added to or removed from a choice set: If each option colors the worth of the others, it is not appropriate to compare choice behavior across decision contexts wherein unrelated options vary as if some options are “the same,” even if their objective values are equivalent.

Surprisingly, there was essentially no support for the Hypothesis 3, the set of predictions involving individual differences. Some of the sub-hypotheses of Hypothesis 3, regarding individual differences, had no past empirical support or extensive theoretical basis beyond that which I covered in my Introduction, above (for example, those regarding length of decision delay for those who prefer maximizing to satisficing). However, the failure to find that more decision-averse people make decisions more extreme on the delay attribute than those who are less averse to decision-making is perplexing. There is, however, some previous support for this lack of results: Patalano and Wengrovitz (2007) found only tenuous links between measures of decisional procrastination and decisional delay. They found results using only one decision aversion scale; not all of the metrics that could support their hypotheses did. Similarly, more generally (and beyond the scope of this dissertation), psychologists have noted the difficulty in predicting behavior from personality trait variables (starting with Walter Mischel’s 1968 book); this difficulty is increased when few behaviors or types of behaviors are examined. As this study examined only one decision, my lack of results may not be so surprising. To address this point, Study 2 examines more than one decision. An in-depth discussion of the general lack of individual difference effects can be found in the General Discussion.

One final point deserves additional discussion: The (unexpected) oddity that several subjects appeared to be weighting options higher if they had *less* free time than more. A majority, as expected, did prefer free time to math problems, but some participants preferred the opposite. There are several possible reasons for this, alluded to above. One possible simple reason could be that these participants were confused about the options they were choosing from and what a certain option meant. However, this seems somewhat unlikely, as the graphical depiction of each option repeated both the amount of free time and of time completing math problems (see **Figure 2**). A more likely possibility is that participants had some intrinsic reason for actually preferring options with more math problems to free time. Numerous possible intrinsic reasons could exist: participants may have been eager to help pre-test math problems, disinterested in the "free time" (perhaps due to mistrust of the privacy of the lab computer, the incongruity of engaging in "fun" during an experiment, or perhaps having no specific desire to use the computer at that time), or—as difficult as it may be to believe—simply enamored of math. This latter possibility is unlikely, however, as an exploratory analysis controlling for participants' numeracy scores did obviate this effect. In all of these cases, however, their desire can be considered "real," and so recoding their objective value scores such that more math (rather than free time) is considered "better" is appropriate.

In addition to the "math-loving" or helpful participants, Study 1 also has several potential drawbacks. The first and perhaps most obvious issue is that all participants effectively saw the same distribution of options—the results discovered here may be artifactually related to the specific order seen in **Figure 1**. For example, the first few options showed to participants trend vaguely downward, but then jump back up. Though random numbers do not follow a pattern, people are very poor at identifying random sequences (Wagenaar, 1972; Wolford, Newman, Miller, & Wig, 2004; but see also Gilovich, Vallone, & Tversky, 1985 for a nice illustration), and may have believed that a "downward trend" was present in the data, leading to hastier choice. Similar effects could have occurred at any large or small drop due to participants interpreting the random sequence as a pattern, or as showing a meaningful trend in one direction or the other. There was also no clear reason to *stop* choosing at any point: As there was no penalty for viewing many choices (besides the potential tedium of the choice process), these results may not generalize well to many decisions people face in their everyday lives. These concerns are addressed in Study 2.

CHAPTER III

STUDY 2

Study 2 was designed to both replicate and extend the findings of Study 1. Specifically, it sought to make the optimal stopping task "more real" by providing participants with a motivation to choose sooner rather than later. Most optimal stopping problems in the real world contain such a motivation, be it to shorten the decision process, to lower the risk of a bad outcome, or simply to reduce uncertainty. Study 2 accomplishes this goal in two ways: First, one condition (the "cost" condition) added a cost for delaying decisions: Participants were told (truthfully) that the amount of free time they choose would be "discounted" or decreased by a percentage equal to the number of options they had seen. Second, in the "reward" condition, quicker decisions were motivated in a positive manner: Participants were told (also truthfully) that the participant who chose the "best option," weighing time to make the decision and obtained amount of free time on the computer equally, would receive a $100 reward. A third condition, the "control" condition, was also included in which no motivation was provided (thus replicating the methods used in Study 1).

An additional issue with the methodology in Study 1 was the fact that there was no variation in terms of the distribution of options that participants viewed. In other words, all participants saw the same sequence of options in the same order. While this allows stronger conclusions to be drawn regarding the different measured variables (they were all equal in terms of the quality of each subsequent option), it reduces the ability to generalize the results to more variable optimal stopping problems. To address this, in Study 2, each participant's sequence was randomly generated by the computer, using the same mean and variance as in Study 1. In addition, each decision-maker participated in four optimal stopping problems—one for each of the low and high mean and variance conditions, in random order. (In Study 1, this was a between-subjects variable.)

Hypotheses

Study 2 was expected to replicate all effects detected in Study 1, including effects which were discovered via exploratory analysis. Many of the hypotheses in Study 1, however, failed to find support. In Study 2, these hypotheses are guardedly retained, though replications of null results will receive special attention in the results and discussion, below. Study 2 also introduces several additional predictions. First and most broadly, the between-subjects effects shown in Study 1 are expected to replicate in Study 2 as within-subjects effects. In sum, the effects found in Study 1 will be present in Study 2 when they are measured within subjects rather than between subjects and also with a separate random distribution for each subject, rather than a single random distribution for all subjects.

The two added efficiency manipulations (reward and cost) are predicted to motivate faster decisions for those in the reward and cost conditions, as compared to those in the control condition. However, faster choices need not mean that decisions will turn out worse. Instead, it could be that if there is motivation to choose quickly, decision-making (and distribution learning, and distribution use) will all speed up in sync. Thus, in Study 2, I predict that not only will subjective and objective quality metrics become more similar over time (Hypothesis 2a), but they will also become similar *more quickly* with motivation to choose quickly. In addition, delay will lead to better choices, but this effect will taper off eventually (Hypothesis 2b), and it will taper off *faster* with motivation to choose quickly. Finally, relative objective quality of an option seen will predict subjective quality, over and above the extent to which absolute objective values predict subjective quality (Hypothesis 2c), and this effect will be *stronger* with motivation to choose quickly.

My individual differences hypotheses for Study 2 parallel those stated in Study 1. While few of these hypotheses were supported in Study 1, Study 2 provides a potential for reprisal, with the potential that individual difference effects will be more apparent when there is a motivation to choose quickly or when there is variance in the actual distribution shown to participants. Thus, even if main effects are not found (as in Study 1), interactions with efficiency condition are still possible.

Study 2 also adds variation in the values of options shown to participants: In Study 1, all participants saw the same values in the same order (e.g., the first option offered was always the same after accounting for the mean and variance manipulations). By generating random data, it is possible to test additional hypotheses that require the value of an option to vary independently of that option's serial position in the list of options, as each participant experiences his or her own sequence of options. For decision sets with a positive slope, I predict that choosing an option will be less likely, while for decision sets with a negative slope, choosing an option will be more likely.

Summary of Hypotheses

Study 2 will test the following hypotheses, with new hypotheses added to those from Study 1 in **bold**:

1. Gaining knowledge of the choice set is desirable.

1a. Quality metrics will not fully mediate the relationship between number of options seen and likelihood of choosing an option.

1b. Delay will not vary between the high and low mean conditions.

1c. Delay will be longer in the high-variance condition than in the low-variance condition.

1d. Decision quality and number of options seen will interact positively to predict how long choice is delayed: There will be a positive linear interaction (between quality and delay), and a negative quadratic interaction (between quality and the square of delay), indicating that quality becomes a better predictor that an option will be chosen as additional distributional information is gained.

1e. People will choose faster when they are motivated to choose quickly.

2. Gaining knowledge of the option set is useful.

2a. Subjective and objective quality metrics will become more similar over time

2b. Delay will lead to better choices, but this effect will taper off eventually.

2c. Subjective quality will relate to the relative objective value of proffered options, over and above the extent to which ratings track the absolute objective values.

2d. Hypotheses 2a, 2b, and 2c will have greater support when people are motivated to choose quickly.

2e. The observed slope of the quality of options will predict whether people choose: Positive slopes will predict delayed choice, negative slopes will predict quicker choice.

3. Individuals will differ in how they utilize distributional information.

3a. People who are more averse to decisions will have more extreme delaying behavior than those less averse to decisions.

3b. Vigilant decision-makers will delay longer than less vigilant decision-makers, and delay will be a better predictor of outcome quality for those who are more vigilant.

3c. Maximizers will delay longer than satisficers, but delay will not result in especially good decisions for maximizers than satisficers (and in fact, delay may be a better predictor of outcome quality for satisficers).

3d. Decision-makers who are more vigilant, less maximizing, and less decision averse will produce subjective ratings that are closer to the relative objective values of the option at hand than the less vigilant, more maximizing, and more decision averse, respectively.

3e. People whose subjective ratings more closely track the relative objective values of the option at hand will obtain better outcomes.

3f. Hypotheses 3a through 3e will show greater support when efficient choice is motivated than when efficiency is not motivated.

Method

Participants

Participants were 109 college students from the University of Oregon recruited in the same manner (from the Human Subjects Pool, sharing the same demographic qualities) as in Study 1.

Procedure

The study procedure was equivalent in almost all respects to the procedure from Study 1, with a few exceptions. First, the description page explained that participants would be completing the optimal stopping problem four times. Although participants were not told this, the four problems corresponded to each of the between subjects conditions in Study 1 (low versus high mean crossed with low versus high variance), which were presented in a random order. Participants were told that for each of the four problems, they would choose the number of minutes they wanted to spend in free time. After doing this four times, the computer would randomly select one of the four times they chose as the "actual" time they would spend in free time.

Second, the "efficiency" manipulation was presented to subjects in the form of the initial introduction screen they saw on a computer (this screen was, as in Study 1, also read to participants and explained to them by the experimenter), in the "cost" condition, participants were told:

> … there is a cost to waiting! Specifically, there is a 1% penalty for each option you reject. In other words, you will get 100% of the free time you choose if you choose the first option, 99% of the free time you choose if you choose option 2, etc. In other words, the faster you decide (assuming your choice is good), the greater the amount of free time you will receive!

Participants in the "reward" condition were told:

> … there will be a $100 prize awarded at the end of the term to whichever participant makes the ***best*** decision in the ***shortest*** amount of time, *weighting time and quality equally*. In other words, the faster you decide (assuming your choice is good), the better your chances are of winning!

Emphases were present in the description read by participants. Participants in the "control" condition were given the same instructions as in Study 1.

Finally, not all participants were given the individual difference measures after completing their decisions: Nineteen participants instead completed these surveys in an apparently unrelated study (a subject-pool-wide "General Survey," designed to collect survey data such as this for many researchers) in order to test for whether the study procedure biased responses on these tests. During the actual study, after participants had made their decisions, participants were asked to provide electronic consent for these earlier survey responses to be accessed. Participants were not told which survey responses were accessed so as not to bias their further behavior.

Results

The analytic strategy taken in Study 2 was basically the same as the strategy taken in Study 1, though in Study 2, there were four decisions made per person. As such, the random intercept estimated for each person is based on more information, but still represents that person's overall likelihood of choosing across all four mean and variance manipulations. As in Study 1, effects that can be measured within persons are treated in all reported models as random effects estimated separately for each person and then aggregated.

To account for the fact that some participants once again showed a negative correlation between subjective and objective quality (indicating a preference for less "free time"), the same strategy was used as in Study 1: For the four such participants in the control condition (11% of the control condition), their effective value scores were reflected around the mean for the distribution they were viewing. For the eight participants in the "reward" condition (21%), the same strategy was used. For the "cost" condition, however, the cost was implemented by decreasing the amount of free time as participants waited; for participants who preferred less free time, this fundamentally changes the discount function. As such, the three participants in the cost condition (9%) who showed a negative relationship were dropped from analyses that examined objective value.

Finally, it should be noted that participants in the "cost" condition have two "objective quality" measures—the discounted objective quality (i.e., the number of minutes of free time they actually experienced), and the raw objective quality (the number of minutes of free time they were shown, pre-discount). Analyses were conducted on both raw and discounted objective quality. When results for these two values differ below, both effects are noted, but only discounted objective quality is reported when the effects are of the same significance and in the same direction.

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable

These hypotheses regard the extent to which holding and gaining knowledge of the choice set is desirable, thus making the obtainment of this knowledge a reason for delaying choice. As in Study 1, support was found for Hypothesis 1a, showing that number of options viewed is a significant predictor of choice beyond the objective and subjective quality of the choice, $B = .02$, $z = 6.12$, $p < .001$. Also as in Study 1, objective quality was not a significant predictor, $B = .005$, $z = 0.31$, $p = .76$, and subjective quality was, $B = 1.28$, $z = 20.49$, $p < .001$.

Hypotheses 1b and 1c predicted a significant effect of variance (higher variance leading to more delay in choosing) and no significant effect of mean (low mean and high mean should not predict faster or slower choice), respectively. Whereas in Study 1 found an effect for variance and no effect of mean, Study 2 produced nearly the opposite results: There was a significant effect of mean, such that being in the high mean condition predicted a lower likelihood of choosing (and thus delayed choice), $B = 1.04$, $z = 2.42$, $p = .02$. As for the variance effect, there was a nonsignificant effect of variance for participants in the high-mean condition ($B = 0.22$, $z = 0.46$, $p = .65$), and only a marginal effect of variance in the low mean condition using discounted objective value,

$B = 0.46$, $z = 1.77$, $p = .08$. However, for analyses using "raw" objective value, the effect of variance was significant in the low mean condition, $B = 0.57$, $z = 2.19$, $p = .03$. This replicates the results of Study 1, which also showed no effect of variability in the high-mean condition but an effect of variability in the low-mean condition. Together, this provides support for the hypothesis that people take longer to decide when choice sets are more variable (Hypothesis 1c).

Additional follow-up analyses were conducted, focusing only on those in the control condition (to parallel Study 1) or only the first decision made (to control for any effects of problem order). These analyses showed similar patterns to those described above, with the people taking significantly longer to choose when presented with the high mean distribution (B's > 1.04, z's > 2.40), people taking consistently more time to choose when in the high variability condition and the low-mean condition in all analyses (B's > 0.33, z's > 0.87), and people showing no effect of the variability condition when in the high-mean condition (B's between 0.48 and -0.68, z's between 0.69, -1.69).

Hypothesis 1d predicted that quality would become a stronger (i.e., more useful, or better attended-to) predictor of choice once people had gained distributional information, operationalized as a positive interaction between quality (objective and subjective) and the linear component of delay, and a negative interaction between quality and the quadratic component of delay, when predicting whether a decision-maker chose a given option. The predicted pattern of results was present in Study 2 for subjective quality, which showed the expected positive linear interaction, $B = .002$, $z = 3.13$, $p = .002$, and the negative quadratic interaction, $B = -0.0003$, $z = -4.18$, $p < .001$. The effects for objective quality found in Study 1, however, did not replicate using discounted objective quality in Study 2 (though the coefficients were in the predicted direction, linear $B = 0.25$, quadratic $B = -1.23$, $|z|$'s > 0.57). When using raw objective quality, however, these effects were also present, in the correct direction, and marginally significant, $B = 0.0003$, $z = 1.71$, $p = 0.09$ for the linear effect, $B = -0.00004$, $z = -1.67$, $p = 0.09$ for the quadratic effect.

Finally, Hypothesis 1e (which was new to Study 2) predicted that those in the conditions that manipulated motivation to decide (via cost or reward) would choose more quickly than those in the control condition. To test this, I examined whether the total number of options viewed before choosing varied by condition, using hierarchical Poisson regression. This model was hierarchical due to the fact that each participant made four decisions, and Poisson regression was used as my the measure of delay was a positively-skewed "count" variable (the count of the number of options viewed prior to deciding; see Atkins & Gallup, 2007 for a discussion of Poisson regression). Condition was dummy-coded to produce two predictors, which measured the difference between the "cost" condition and the control, and the difference between the "reward" condition and the control. These contrasts are used to test condition effects henceforth.

When controlling for objective and subjective quality of ratings (as a random slope for each participant), this regression showed only a significant effect for the "cost" condition, $B = -0.85$, $z = -2.68$, $p = 0.007$, indicating that those in the cost condition chose faster than those in the control condition. No effect was shown for the reward condition, possibly because the instructions for the reward condition told participants to "weight time and choice quality equally," which may have increased the motivation to choose *a good option* relative to the control condition (thus increasing choice time) while simultaneously increasing the motivation to choose *quickly.*

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful

My second broad hypothesis, as in Study 1, was that there is some utility in gaining information about the option set. To replicate the results for Hypothesis 2a from Study 1, as people see more options, their subjective ratings should more closely resemble the objective quality metrics. In Study 2, this was tested in the same manner as Study 1, and the results were similar: Subjective rating and amount of delay showed a marginal positive interaction in predicting discounted objective quality, $B = 0.002$, $t = 1.86$, $\chi^2(1) = 3.43$, $p = 0.06$; this effect attained significance for raw objective quality, $B = 0.464$, $t = 3.70$, $p < 0.05$.

Hypothesis 2b predicted that the effect of delay on choice quality would be positive but taper off. This was tested using a hierarchical linear model that estimated the linear and quadratic components of delay as simple effects within levels of the "mean" condition, but allowed these effects to vary randomly within participants as each person made four decisions. A separate model was examined for objective versus subjective quality, and each model controlled for the other quality metric. Whereas the results from Study 1 were in the right direction but occasionally marginal, in Study 2 the predicted effect was found for objective quality with all coefficients being significant: There was a positive linear effect in both the high mean condition, $B = 0.10$, $t = 2.77$, $p < .05$, and the low mean condition, $B = 0.09$, $t = 3.36$, $p < .05$, and a negative quadratic effect of delay in both the high mean condition, $B = -0.001$, $t = -3.29$, $p < .05$, and the low mean condition, $B = -0.001$, $t = -3.89$, $p < .05$. In contrast, although the effects for subjective quality were significant or marginally significant in Study 1, these effects were not in Study 2.

Hypothesis 2c predicted that the relative objective value would be predictive of subjective quality, over and above the effect of absolute objective value. To address this in Study 1, I computed a *z*-score for each objective value based on the values seen up to that point. However, in Study 2, this value could be computed "relative" to all options viewed within a particular decision of the four that each participant made, or relative to *all* options already encountered across all decisions. To address this, analyses for Hypothesis 2c were conducted twice: once with relative objective value scores being computed relative to all options seen (including those from past decisions), and once with these scores being computed relative to only those options seen for the given decision. Both constructions of relative objective quality showed significant positive effects in predicting subjective value, t's > 3.80, p's $< .001$, showing that the value of an option relative to those seen contributes to the subjective judgment of that option over and above the absolute value, replicating Study 1 in support of Hypothesis 2c.

Hypothesis 2d, new to Study 2, predicts that the predicted results for Hypotheses 2a, 2b, and 2c will be stronger when people are motivated to choose quickly. Though the results for Hypothesis 1e above showed that only those in the "cost" condition actually chose more quickly than the control condition, both "cost" and "reward" conditions are examined here. Study 2 showed marginal support for Hypothesis 2a, with decision-makers' subjective quality ratings growing closer to the objective quality ratings over time. To support Hypothesis 2d, this effect must be stronger in the "cost" and/or "reward" conditions than in the "control" condition. Indeed, a two-way interaction was found between condition (cost/reward/control) and delay in predicting the objective quality of the option selected, indicating that subjects in the "cost" condition chose options with significantly lower objective quality when they delayed choice longer, $B = -0.14$, $t = -6.23$, $p < .05$. This may be due simply to the discounting procedure itself, however, as discounted quality necessarily decreased with time in the "cost" condition; using raw objective quality obviated this effect. No other 2-way interactions were found between condition and subjective quality or delay in predicting objective quality of the option chosen. The three-way interaction between motivation condition, amount of delay, and subjective quality in terms of predicting objective quality of the option chosen, which tests Hypothesis 2d, was significant for the "reward" condition relative to control, $B = 0.01$, $t = 5.22$, $p < .001$, and significant but in the counter-hypothesized direction for the "cost" condition relative to control, $B = -0.03$, $t = -6.09$, $p < .001$. Thus, being in the cost condition, relative to control, led to a *lower* relationship between objective quality and subjective quality over time when objective value was the "to-be-experienced" value (the number of minutes they would actually attain, after discounting). Repeating these analyses using the non-discounted values (i.e., "apparent" value; the number of minutes promised, pre-discount) produced the same positive interaction for the "reward" condition, $B = 0.01$, $t = 5.13$, $p < .05$, though the significant counter-hypothesized negative interaction for the "cost" condition did not replicate for raw objective quality.

As noted above, support for Hypothesis 2b was found in Study 2, in that the effect of delay on (objective) choice quality was positive, but tapered off. If motivation to choose enhances this effect, there should be a positive interaction between the "motivated choice" contrasts and the linear effect of delay on choice quality (indicating that when choice is motivated the linear effect is yet more positive), and/or a negative interaction between the "motivated choice" contrasts and the quadratic effect of delay on choice quality (indicating that when choice is motivated the quadratic effect is yet more negative). However, neither of these effects were found: Only one interaction (out of eight) between condition and amount of delay was significant, a positive quadratic interaction with the "cost" condition, relative to control, for subjective quality, $B = 0.001$, $t = 2.01$, $p < .05$.

The pattern of results representing Hypothesis 2c—that relative objective value would be a significant predictor of subjective ratings, over and above the effect of objective value—was also predicted to interact with motivation condition in Study 2. To test whether this effect is enhanced when participants are motivated to choose quickly, I examined the interaction between condition and relative objective quality; a positive interaction would support this hypothesis. However, the enhancement effect was significant for only one of 8 measures of relative objective quality: There was an effect for raw (rather than discounted) objective quality, measured relative to the current decision (rather than relative to all decisions), for those in the "cost" condition relative to control (rather than for "reward"), $B = 0.21$, $t = 2.18$, $p < .05$, indicating a very inconsistent enhancement effect, if any.

Hypothesis 2e predicted that the slope of the options participants had seen so far would predict delay: Positive slopes would predict further delay (in hopes that options would continue to improve), whereas negative slopes would predict less delay (due to a fear that options would get worse). This result could not be examined in Study 1, as Study 1 gave every participant the same sequence of options; as such, there was no variance in slope. To test this hypothesis, a variable was created that indicated the point-slope of a series of options. In other words, much like the *z*-scores computed for analysis of Hypothesis 2c, *z*-slopes were computed: I predicted the objective quality (corrected for the cost condition, as discussed above) for each option from the index of that option, using a linear model. A separate linear model was fit for every option viewed for every participant, relative to all options viewed previously for that specific decision. The standardized beta for the "option index" effect in this linear model was then taken as the point-slope, or the slope that the decision-maker had experienced thus far when making that specific decision. So, if the first three options were 15, 16, 18, and 17 respectively, these values were predicted from the numbers 1, 2, 3, and 4 (corresponding to the first, second, third, and fourth choice). This would yield a "point-slope" of 0.98 for the first three options viewed, and a point-slope of 0.80 for the first four options. These point-slope values were then used in a hierarchical logistic model as a predictor of whether the decision-maker chose the given option. As was the case for analyzing objective quality, point-slopes were computed both based on raw objective quality and separately for discounted objective quality.

The original model fit was a hierarchical logistic model, predicting whether an option was chosen or not from the point slope, controlling for amount of delay and the interaction between point-slope, delay, and objective quality (because the objective quality of an option has more influence on slope for options viewed earlier than options viewed later), as well as subjective quality. None of these models showed a significant effect of point-slope.

Two exploratory analyses showed, however, that point slope does play an important part in evaluating choices: Subjective quality was predicted by a two-way interaction between point slope and objective quality. In this model, point-slope did not predict subjective quality (as a main effect), but interacted with objective quality, $B = 0.03$, $t = 3.50$, $p < .05$, indicating that more-positive slopes make yet-better options seem (subjectively) even better (or that more-negative slopes make yet-worse options seem subjectively even worse). A marginal positive three-way interaction was also found when amount of delay was added to the model, $B = 0.72$, $t = 1.85$, $\chi^2(1) = 3.44$, $p = .06$, indicating that this effect becomes stronger as more options are viewed. The same models were not fit for objective quality, as objective quality scores were randomly generated and thus predicting them is of vacuous interest.

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information

Individual difference variables were collected in one of two ways: Some participants ($n = 19$) completed their individual difference surveys as part of an unrelated study, while the others ($n = 90$) completed their surveys as part of this study, after making their decisions. These groups were compared using Welch's *t*-test to account for the large difference in sample size. These groups differed significantly in terms of two individual difference measures: Diab and colleagues' (2008) measure of maximizing tendency, such that those who completed their surveys as part of the study reported themselves as being higher on maximizing, $t(48.15) = -7.20, p < .001$. This effect was nonsignificnt for Schwartz and colleagues' (2002) maximizing scale. Those who completed their surveys as part of the study also reported themselves as being lower in decision aversion according to Frost and Shows' (1993) indecisiveness scale, $t(52.05) = 4.33, p < .001$, but not the MDMQ (1997) procrastination scale.

Hypotheses regarding individual differences in Study 1 were largely unsupported, though in Study 2 these effects were given a second chance – both on their own, and in interaction with whether choice was motivated. Specifically, it was predicted that the set of Hypotheses 3 from Study 1 would show stronger effects in the efficient choice conditions than in the control condition. The individual difference scales showed generally the same reliabilities and correlation patterns as in Study 1. Once again, very few significant results were observed, and those that did occur showed the same inconsistent pattern: Significant but potentially spurious relationships observed in Study 1 (i.e., those that were marginal or significant but then rendered insignificant when correcting for family-wise type-I error) did not replicate, and new effects observed did were not consistent (e.g., across scales that measure the same construct, such as maximizing and decision aversion). In sum, the results for individual difference measures are inconsistent within each of Study 1 and Study 2, as well as being inconsistent across studies.

Discussion

Study 2 was a mixed bag in terms of replicating Study 1: Many of the patterns replicated, but unfortunately this included the numerous null effects observed when testing Hypothesis 3. In terms of replicating predicted results, across both studies, participants preferred to wait for more information (Hypothesis 1a), especially when that information appeared potentially useful: More variable option sets led to delayed choice (at least when the mean level of option quality was below ceiling; Hypothesis 1c), and the objective quality of options became a stronger predictor of an option being chosen as the participants viewed more options (Hypothesis 1d), indicating that delaying choice led participants to feel more comfortable basing their decisions on the apparent quality of the available options (though this effect tapers off eventually). Delaying in order to gain more information about the distribution was useful (in terms of achieving better outcomes; Hypothesis 2a), though this effect tapers off (for objective and subjective quality in Study 1, and for objective quality in this study; Hypothesis 2b). Additional information was also useful in terms of gaining more knowledge of the distribution

(Hypothesis 2c): People who viewed more options showed a closer correspondence between subjective ratings and objective quality.

These replications are notable because Study 2 tested these hypotheses using methods different from those in Study 1. First, the fact that each participant saw random data across all four of their optimal stopping problems indicates that these effects are not artifactually due to the single distribution used in Study 1. However, it appears several of the effects found in Study 1 were potentially weakened in Study 2 by this methodological change. For example, participants who saw more highly variable decision sets showed much less of an increase in delay compared to participants from Study 1 (Hypothesis 1c). Study 2 also did not fix the potential "ceiling effect" problem noted in the discussion of Study 1, above—because the numbers generated were actually random, it is possible that some participants in the "high mean" condition may have seen ideal or near-ideal options, for example "29 minutes of free time" when they knew the maximum to be 30. Similarly, in Study 2, participants were in fact swayed by the overall mean amount of free time offered—they chose faster when the mean was higher, whereas this was not the case in Study 1. This finding in Study 2, however, does not contradict Hypothesis 1, which stated that additional distribution information would be desirable. The fact that other factors (such as the overall mean quality of options seen) may contribute to how many options are viewed does not mean that people are not still interested in seeing more options when the decision appears more variable, or that people are uninterested in learning about the distribution overall.

Another notable failure to replicate Study 1's results was for parts of Hypothesis 2b. In Study 1, delaying choices showed a positive (but tapering off) effect for both objective and subjective quality (albeit occasionally marginal): Those who delayed made better choices, but especially long delays had a negative influence on choice quality. In Study 2, this effect was present for objective quality of the options, but not subjective quality. While many have argued that objective quality is really the "gold standard" of whether participants have made a "good decision," it is curious that the decision-makers do not appear to experience this "better choice." One explanation for this is taken from Hypothesis 2a: As time goes on, participants show a stronger correspondence between subjective and objective quality, indicating that this correspondence is *weaker* earlier in the decision task. It is possible that, by the time participants have viewed enough options to feel comfortable choosing, the point of diminishing returns has passed. This process would lead subjective quality to be less related to delay earlier in the decision (when the linear effect is possible), making this effect difficult to detect. Whether people learn fast enough to show such a subjective effect is likely to depend on the actual distribution they are faced with. Thus, for example, the use of random data in Study 2 may well have rendered these effects undetectable, whereas they were detected in Study 1 (which used only one distribution). Indeed, taking this into account, it is perhaps remarkable that the effect for objective quality, some tests of which were marginal in Study 1, replicated in Study 2.

Another methodological difference between Studies 1 and 2 was the introduction of an "efficient choice" manipulation: Some participants were encouraged to make more efficient choices. The effects of this manipulation were mixed: The "reward" manipulation failed to increase choice efficiency, though the "cost" manipulation did. In hindsight, this may be due to the competing demands placed on participants with the introduction of the "reward" manipulation: The instructions encouraged them to make better choices faster may, which may have simultaneously encouraged deliberation and information seeking as well as speed. Conversely, the "cost" manipulation may have increased speed at the cost of consideration. Further, I note that whenever "raw" and "discounted" objective quality (a separation that existed only for the "cost" manipulation) differed, in all cases the "raw" quantity provided stronger support for the hypothesis in question. This indicates that decision-makers were more focused on the objective quality of the choice itself, and less focused on the quality they would personally receive if they chose it. This also provides evidence for the weakness of the "cost" manipulation, suggesting that future research into motivating efficient choice should consider utilizing additional motivation strategies.

One effect of this weak manipulation was weak support for Hypothesis 2d, which predicted that motivation to choose would lead to choices being better, operationalized as finding stronger support for Hypotheses 2a, 2b, and 2c when choice was motivated, despite the fact that the choices were being made more quickly. There was only piecemeal support for these predictions: The "reward" manipulation led to a more rapid increase in the correspondence between subjective ratings and objective values, but the "cost" manipulation did not. However, the cost manipulation did lead to faster choice (Hypothesis 1e), and neither efficiency manipulation led to *worse* choices: The lack of interactions, while unfortunate, also indicates that the effects shown in Study 1 and replicated in Study 2 are robust against manipulations such as these. While the manipulations themselves may have been less than ideal, they undeniably added a cost to delay or a reward for choosing faster, and Hypotheses 1 and 2 were robust against these costs and rewards.

Hypothesis 2e, which predicted that the slope of the options viewed would lead to differences in the likelihood of an option being chosen, was not directly supported. This may be a methodological artifact of the design of the study: As participants had no idea what distribution of options to expect, those experiencing a positive (or negative) slope had no reason to expect that the slope would continue to rise (or fall), but may have worried (hoped) that eventually the trend would crash (bounce back). Prior research on slopes suggests that the point-slope observed for the chosen option would be a significant predictor of decisional satisfaction (Hsee & Abelson, 1991), but I did not examine decisional satisfaction directly in this study, so my results do not conflict with Hsee and Abelson's. My exploratory analyses, indeed, support their results: The subjective quality of options (appreciation of an option itself, which may or may not have been chosen) was higher for more-positively sloped distributions, and yet higher still when the positive slope had persisted for some amount of time. This extends Hsee and Abelson's results by suggesting an underlying process by which decisional satisfaction is increased: People are more satisfied with options drawn from a positive slope simply because options, chosen or not, are rated subjectively higher when the slope is higher, even controlling for the objective values of the options.

Finally, the individual difference measures once again failed to produce any reliable effects, as well as any reliable interactions with the efficiency manipulations. These hypotheses are given a third chance in Study 3, below, and will be reviewed at greater length in the General Discussion.

Overall, Study 2 did a good job of supporting Hypotheses 1 and 2, broadly. No evidence appeared that suggested that waiting to learn about the distribution of options was harmful to decision-makers, and many pieces of evidence (the replications of hypotheses also supported in Study 1) suggest that gaining distributional knowledge, even at some cost, is beneficial and desirable. Study 2 showed also that participants are sensitive to the slope of the options that they have seen so far (albeit not precisely in the hypothesized manner), and that the effects of Study 1 remain reliable even when the decision context is manipulated such that delay has a cost (or lack of delay is rewarded).

Study 2, however, raises some additional questions regarding the generalizability of Hypotheses 1 and 2. First, over these two studies, using a "reward" of spending study time as one wishes (rather than doing math problems) produced a set of participants whose preferences empirically indicated that their preferences were reversed and they would rather complete math problems. As discussed in Study 1, above, this could be for several reasons, not the least of which is that participants may have simply preferred to be helpful during their time as research participants than to fritter it away. While the method of reverse-coding objective value for these participants is robust so long as these correlations actually indicate a preference for completing math problems, this inconsistency with the expected effect of this paradigm requires replication using a different decision. Similarly, the decision itself, albeit "real" in the sense that participants actually received the option they chose, was not the sort of decision that these participants (college students) would likely face outside the laboratory. Even beyond trouble with the specifics of the "free time" versus "math" paradigm, it still provided only one dimension along which participants could evaluate options. Many real world decisions are "multi-attribute," meaning that a given option has scores in many different domains of interest that may contradict each other (Kramer & Hodges, in prep; Timmermans, 1993). Finally, while Study 2 improved upon Study 1 by allowing each participant to experience their own distribution (and thus allows generalization of the replicated effects across many random sequences), Studies 1 and 2 both made use of a normally distributed set of numbers, while decisions in the real world may be distributed in many ways. These concerns are addressed in Study 3.

CHAPTER IV

STUDY 3

Whereas the first two studies are geared towards testing whether people desire and utilize information about distributions of choices, the decisions tested in these studies are much simpler than most decisions made by real people in the real world. To address this, Study 3 extends the two prior studies in two ways. First, it uses a vignette describing a real-world optimal-stopping problem that many people in the college-student population have experienced or have had direct experience with: choosing a roommate. Second, Study 3 extends the prior two studies by allowing for more than one quality metric: Potential roommates are "good" if they are financially solvent, but also if they are personable. Many real-world decisions have multiple dimensions on which options can be evaluated, leading to additional complexity that bears study (Kramer & Hodges, in prep; Timmermans, 1993).

I still predict to find the same basic processes with this more ecologically valid vignette-based study: People will wait to gain distributional information, and will still make use of it. To see if other results from Studies 1 and 2 will replicate in Study 3 requires measurement of the same conceptual variables in Study 3: number of options seen, subjective ratings of each option, objective ratings of each option, correspondence of the subjective ratings to the objective ratings, and individual difference variables.

While most of these variables will be measured in the same manner in Study 3, measuring objective quality provides a unique challenge in a study designed to have no "objectively correct" choice (in the real world people may reasonably value personability of a housemate differently than they value the housemate's financial stability, or not). However, in Study 3, there are still only two cues to the quality of an option that participants are aware of: personability and financial stability. There is also a rather impressive literature on clinical versus actuarial decision-making that has shown that, across training sets, models based on individuals' actual behaviors (i.e., without reference to a training set) actually "...predict...the external criterion more accurately than [the individuals]; this was true for each [individual]...because they do not apply their own weights consistently" (Grove & Meehl, 1996, p. 316; see also Dawes & Corrigan, 1974). This is to say, by computing a linear model for each participant predicting their subjective quality ratings from the value of each cue, I should be able to predict the external criterion (in this case, an idiographic metric of "objective quality") better than the ratings themselves would for a single given option. As such, I predict that results found for objective quality in Studies 1 and 2 will replicate using idiographic model-implied quality (computed separately for each subject) in Study 3.

Of course, there is the distinct possibility that this model-implied choice will be more accurate for some people than for others. While my predictions regarding "objective quality" may appear to suffer for this lack of precision, I note that in the prior studies, "objective quality" was defined as the number of minutes of free time (versus math time) that participants would have as part of the study; in Studies 1 and 2, this objective quality metric may have varied in precision across participants (some may have had a stronger preference for one sort of time than the other; indeed, the fact that some preferred to complete math problems promotes this), the model-implied quality metric in this study may do so as well. In essence, this also constitutes an interesting metric for additional study: As participants make their decisions and gather information about the distribution of options (in Study 3, the joint distribution of the two cues), will they focus on one cue, or an amalgamation of the two cues into a single "quality" distribution, or does the tendency to do this depend on individual difference measurements?

There is a large body of literature on multi-attribute decision-making or "multiple-cue judgment" (see Karlsson, Juslin, & Olsson, 2008 for a review). Several process models exist for explaining how multiple-cue judgments are to be made. The most frequently discussed judgment process for this is referred to as "exemplar memory," in which a decision-maker recalls an exemplar (i.e., an item that would be judged as "ideal"), and then judges an item at hand not directly, but in terms of its similarity to the exemplar (Juslin, Olsson, & Olsson, 2003). Another, more compensatory strategy for multiple-cue judgment is a simple additive model: Quality ratings along all cues are summed, and then the judgment is made corresponding to the sum. People appear to utilize some modification of one of these two judgment procedures (exemplar or additive) depending on the task at hand (Juslin, Karlsson, & Olsson, 2008). However, there is still a rather marked lack of evidence regarding how people might weight cues when using an additive judgment strategy: Tversky (1972) suggested that "bad" (less important, less differentiating, etc.) cues would simply be dropped in order to clarify decision-making, and Dawes and colleagues have built a remarkable case nothing that human judgment underperforms statistical models in many contexts (e.g., Dawes, 1979; Dawes, Faust, & Meehl, 1989; Goldberg & Werts, 1966; Grove & Meehl, 1996). Given this, there is reasonable evidence to expect participants, when offered multiple cues that together require a single comparative judgment of "quality" be made, will use a compensatory (e.g., additive) strategy. However, this literature does not address how people seek information regarding how each cue is distributed across the judgment space: If cues are to be weighted against each other in any manner, judges likely develop and then utilize a notion of how "relevant" each cue is. Indeed, the literature on feature-matching in judgment suggests that people do just this: Any cue which is effectively "irrelevant" in the sense that options are more or less equivalent on the cue will be "matched" on and subsequently ignored (Hodges, 1997), indicating that (at least in this extreme case) the distribution of cues across options affects the weighting of the cues themselves. In sum, this literature does not address the desirability of "knowing the option space" for an option space defined by multiple cues.

The literature has, however, focused frequently on specific strategies for choosing among multiple cues in single-shot decisions about options that vary on multiple dimensions (rather than gaining information about a set of options that vary on more than one cue). Following the literature on multiple-cue judgment, which shows that people may undertake many strategies for amalgamating cues in order to make a choice (Karlsson et al., 2008), I predict that the extent to which participants utilize and aggregate multiple cues (as opposed to ignoring some) will be an individual difference. The key to this difference, I believe, has to do with the mathematical difficulty of a compensatory strategy: People are not good at weighting cues against each other, either in the abstract or in actual decisions; rather, this sort of cue weighting is best left to linear models (at least, when a training process is available; Dawes, Faust, & Meehl, 1989). Tversky's (1972) "Elimination by aspects" theory of choice suggests that people will simply drop cues from their mental representations of decisions until a dominant option rises to the top on the remaining cues. Tversky's data and theory suggest that people will attend to one cue specifically, rather than weighting the cues against each other, while Juslin, Karlsson, and Olsson (2008) suggest that if cues can be added together, they will be.

As such, this study addresses the possibility that the use of cues (in a decision where cues could be added) varies as an individual difference: the degree to which individuals take pleasure in mathematically difficult tasks such as this. Peters and colleagues' (2006) numeracy research suggests that some individuals are not only better at mathematical tasks, but also take pleasure in the processing and manipulation of numbers, suggesting that numeracy should predict the extent to which people actually utilize (one or both of) the cues provided to them, noting well that Studies 1 and 2 would caution against making predictions of individual differences predicting decisional behavior. That said, given the body of research suggesting that (even when they are unqualified) people prefer to mentally weight cues than to rely on actuarial tables (Dawes et al., 1989), numeracy should predict the extent to which people utilize the cues available to them, or utilize an aggregate of cues rather than just one cue.

I will test the numeracy prediction using the output from the idiographic linear model I have computed for each participant in order to generate model-implied choice: Along with calculating model-implied choice for each option viewed by each participant, I can also examine qualities of each participant's linear model: Specifically, the beta weights and the R^2. I will take the sum of the standardized beta weights for each cue (the stability cue and the personability cue) as an empirical individual difference variable: the extent to which participants are using multiple cues. I will also examine each individual's adjusted R^2 for predicting ratings from the value of each cue: This allows me to detect the extent to which participants' ratings are *determined* by the cues, which effectively includes people who underweight one cue to promote the other or use other noncompensatory strategies. I predict that, for both measures of "cue use," numeracy will be a positive predictor (showing that numeracy predicts cue use in a continuous manner).

Study 3 also drops the "mean" and "variance" within-subjects manipulations from Study 2, as the "roommate" vignette potentially lends itself to cross-decisional learning: As roommate searches are more "real" and emotionally tinged, it is harder to argue that participants will be able to "reset" their notion of what counts as a good roommate four times. Instead, participants are asked to make just two roommate selections, switching from random scoring of roommates on a normal distribution to the random scoring of roommates on a uniform distribution. However, Study 3 still allows for a test of whether distributions with higher variance lead to longer decision-making times, as uniform distributions have higher variances than normal distributions.

Hypotheses

Study 3 is expected to replicate the basic findings of Studies 1 and 2, despite the methodological differences between the studies. This includes all hypotheses which received support in either Study 1 or Study 2 (save those referring to differences due to the "mean" condition). I also retain the specifics of Hypothesis 3, constituting a second chance for reprisal for the individual difference predictions, which may be more clearly evidenced via a more naturalistic decision vignette. In sum, my predictions follow, using the same indices as in the prior studies for parallelism (even though some hypotheses, noted below, have been dropped or altered); new and heavily altered predictions are in boldface. I predict:

1. Gaining knowledge of the choice set is desirable.

1a. Quality metrics will not fully mediate the relationship between number of options seen and likelihood of choosing an option.

(1b. This study drops the prediction that option sets with a higher mean will show no difference in delay behavior, because the mean of the distribution is not varied.)

1c. Participants who view a uniform distribution of options will take longer to choose than those who view a normal distribution, as the uniform distribution will show a higher variance.

1d. Decision quality and number of options seen will interact positively to predict how long choice is delayed: Option quality will become a better predictor of choice as participants view more options (constituting a positive linear interaction), but this will taper off as choice is delayed further (constituting a negative quadratic interaction).

2. Gaining knowledge of the option set is useful.

2a. Subjective and model-implied quality metrics will become more similar over time.

2b. Delay will lead to better choices, but this effect will taper off eventually.

2c. Subjective quality will relate to the relative model-implied quality of proffered options, over and above the extent to which ratings track the absolute model-implied quality.

(2d. This study drops the prediction that motivation manipulations will enhance Hypotheses 2a, 2b, and 2c, as no motivation manipulation is used.)

2e. The experienced slope of the model-implied quality of options will predict whether people choose: Positive slopes will predict delayed choice, negative slopes will predict quicker choice.

2f. Point-slope will interact with model-implied quality in predicting subjective quality, showing that a positive slope makes good options seem even better.

2g. Effects previously shown for "objective quality" will replicate for "model-implied" quality.

3. Individuals will differ in how they utilize distributional information.

3a. People who are more averse to making decisions will have more extreme delaying behavior than those less averse to decisions.

3b. Vigilant decision-makers will delay longer than less vigilant decision-makers, and delay will be a better predictor of outcome quality for those who are more vigilant.

3c. Maximizers will delay longer than satisficers, but delay will not result in especially good decisions for maximizers than satisficers (and in fact, satisficers may make better decisions).

3d. Decision-makers who are more vigilant, less maximizing, and less decision averse will produce subjective ratings that are closer to the relative objective values of the option at hand than the less vigilant, more maximizing, and more decision averse, respectively.

3e. People whose subjective ratings more closely track the relative model-implied quality of the option at hand will obtain better outcomes.

(3f. This study drops the hypothesis that individual difference effects will be moderated by motivation to choose quickly, as no motivation is used.)

3g. Those higher in numeracy will be higher in "cue use," i.e., the extent to which they utilize both of the cues available to them.

3h. Those higher in numeracy will make more consistent choices, indicated as having a greater proportion of variance in subjective ratings of each item explained by the model-implied quality for that item.

Method

Participants

Participants were 55 college students, recruited for a 30-minute study in the same manner and from the same pool as participants in Studies 1 and 2.

Procedure

The study was advertised as a half-hour-long study. The study procedure was equivalent in almost all respects to the procedure used in Studies 1 and 2, with a few exceptions. First, the instructions were changed in order to explain the vignette. Participants were told that:

> **Your decision**, over the next several pages is to decide **which housemate you would like to have**. Imagine that you have a house, which you love, but that a housemate has moved out and you need someone else to move in so that you can cover rent. There are no restrictions from the landlord on who this can be--it's entirely up to you. Furthermore, you have several applicants for the room, all of whom are different. … These potential housemates will differ from each other in two ways--how financially stable they are (someone who is not very stable might miss utility payments or be late with rent) and how well you get along (or how "personable" they are).

Emphases were in the actual instructions to participants. Instructions were otherwise the same, though the display indicating the relative quality of the housemates was also altered to display housemate quality in a continuous manner, using sliders (see Figure 4).

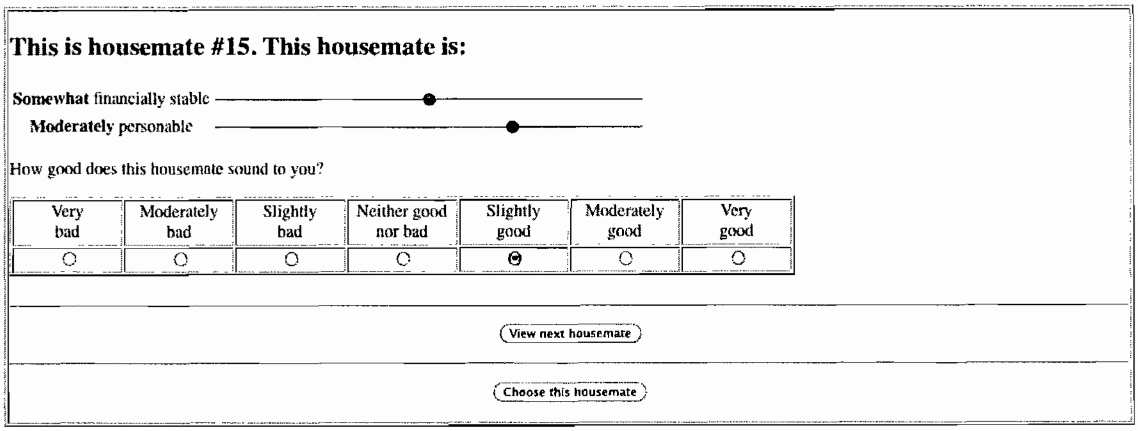

Figure 4. Participants' view of the decision they were making in Study 3.

As in Study 2, some participants (n = 5) completed the individual difference measures before participating in the experiment, while the others (n = 46) completed these measures after making both decisions, as in Study 2. This large difference in the number of subjects who completed the forms in the "general survey" as opposed to during the study was due to a lack of students participating in the general survey in the term during which data was collected. There was also an alert in between decisions (the first and second decision, constituting the switch from the normal to the uniform distribution of options or vice versa; these orders were counterbalanced) that read:

> Next, we would like you to imagine that you have graduated and have moved to a new city. You do not know anybody, but you are in the same boat: You have found a place to live, but need a housemate. Since you're in a new city, you don't have a good sense of the pool of housemates available. So, your knowledge of roommates in Eugene may not help you here. Click OK when you are ready to begin.[5]

---

[5] Interestingly, while collecting data for Study 3, the author was faced with the inverse decision: Though an excellent housemate had been found, the author was engaged in a

When participants had completed the optional stopping problems and the surveys (note that some participants may have completed the surveys prior to beginning the study; see above), the study was complete.

Results

The analytic strategy taken in Study 3 was similar to that taken in Study 2. Exploratory analyses showed that personability and financial stability were widely considered desirable qualities to have in a housemate, with 94% and 95% of participants respectively showing a positive relationship between subjective ratings and the roommate quality. For only one of the three participants who showed a negative relationship for either stability or personability was this negative relationship statistically significantly negative (for personability; this individual viewed only two options for each of his two decisions, making this significant effect quite untrustworthy; removing this participant did not change any of the reported effects). Due to the small number of participants who completed the surveys outside of the study ($n = 5$ is generally considered too small to produce trustworthy comparisons), differences due to survey administration procedure are not examined.

Participants' use of cues was fairly consistent (mean std. beta = 0.56 for financial stability, 0.58 for cue use); participants did not significantly prefer one to the other,

---

long-distance optimal stopping problem to choose an apartment to rent, effectively weighting commute time against cost.

$t(49) = -0.32$, $p = 0.75$. Cue use (sum of betas) had M=0.56, sd=0.25; cue use (adjusted R^2) had M=0.69, sd=0.16).

Hypothesis 1: Gaining Knowledge of the Option Set Is Desirable

Hypothesis 1a predicted that quality would not fully mediate the relationship between number of options seen and the likelihood of choosing an option. This was tested by predicting whether an option was chosen or not from the amount of delay and the two quality metrics: subjective rating and model-implied quality. This effect replicated, $B = 0.01$, $z = 2.08$, $p = 0.04$, showing that delay was still a positive predictor of whether or not an option was chosen, over and above the effect of quality.

Hypothesis 1c predicted that participants would take longer to choose when the distribution of qualities was uniform rather than normal, as an analog to the earlier findings regarding how increased variance in quality led to more delay in choice. This effect was not significant, $B = 0.38$, $z = 1.24$, $p = 0.21$.

Hypothesis 1d predicted that, as participants view more options, the quality of those options would become a stronger predictor of choice. This is operationalized as a positive linear interaction and a negative quadratic interaction between decision quality and the number of options seen in predicting how long a choice is delayed. For subjective decisions' quality, this effect was significant and positive, as predicted, for the linear effect of delay, $B = 0.01$, $z = 3.76$, $p < 0.001$, and negative and significant, as predicted, for the quadratic effect of delay, $B = -0.001$, $z = -4.61$, $p < 0.001$. For objective quality, Study 3, which uses model-implied quality, did indeed show this effect, both in terms of the positive linear effect, $B = 0.17$, $z = 3.42$, $p < 0.001$, and a negative quadratic effect, $B = -0.0002$, $z = -5.70$, $p < 0.001$, replicating what was found in Study 1, but not Study 2 (where these effects were not present for objective quality).

Hypothesis 2: Gaining Knowledge of the Option Set Is Useful

Hypothesis 2a predicted that subjective and model-implied quality measures would become more similar over time. Given the differences between "objective" and "model-implied" quality, which make model-implied quality much more similar to subjective quality, there may be much less variance in the interaction term (between subjective quality and amount of delay, in predicting model-implied quality) that tests this hypothesis. Indeed, in Study 3, this effect was not significant.

Hypothesis 2b predicted that the effect of delay on choice quality would be positive but taper off. Specifically, a positive linear effect of delay on individual model-implied quality was predicted, but with an additional negative quadratic effect. As in Studies 1 and 2, this effect was examined within levels of the distribution variables (i.e., separately for the uniform distribution, and then again for the random distribution). For model-implied quality, an analogue of "objective quality," for which support was found for Hypothesis 2b in Study 1 but not in Study 2, only one effect out of four (a linear and quadratic effect within each distribution type), was significant: Contrary to predictions, greater delay predicted *lower* model-implied quality (i.e., a negative linear effect) for the uniform distribution, $B = -0.003$, $t = -2.50$, $p < 0.05$. For the uniform distribution the quadratic effect was positive (also counter to predictions), but not significant. For the normal distribution, the linear effect was nonsignificant but in the right direction, $B = 0.00005$, $t = 0.07$, $p > 0.10$, and the quadratic effect was marginal and in the right direction, $B = -0.000002$, $t = -1.66$, $p_{mcmc} = 0.10$. The same four tests were run for subjective quality, which showed support for Hypothesis 2b in both Study 1 and Study 2. Both linear effects were nonsignificant but in the right direction (B's > 0.02, p_{mcmc}*'s > 0.29)*, and the quadratic effect was marginal but in the correct direction for the normal distribution, $B = -0.0003$, $t = -1.67$, $p_{mcmc} = 0.10$, and significantly in the right direction for the uniform distribution, $B = -0.0005$, $t = -2.12$, $p < 0.05$.

Hypothesis 2c predicted that subjective quality would be correlated with relative model-implied quality, over and above absolute model-implied quality. This hypothesis may seem odd or trivial based on the fact that my model-implied quality metric is generated from subjective quality ratings, however, I note that each option's model-implied quality is based on the entire *set* of options viewed and rated. So, for example, computing the relative model-implied quality for the third option viewed involves modeling quality based on all options viewed, generating the model-implied quality (based on all options) for the first, second, and third options, and then examining the z-score for the third option (relative only to the other two). As such, this still remains a test of whether the relative values (based on choice that has not yet been fully revealed) are predictive of ratings over and above the absolute values (which are currently being revealed). As above, because I use quality metrics that are derived from each other in my model (i.e., as "controls"), this effect should be more difficult to show in Study 3 than prior studies, however I did find a significant effect in this study, $B = 0.17$, $t = 2.12$, $p < 0.05$, despite the use of model-implied quality instead of objective quality.

Hypothesis 2e predicted that positive slopes of objective option quality would increase the likelihood of choice, while negative slopes would decrease this likelihood. Slopes were computed the same manner as in Study 2. As in Study 2, no main effect was found for point-slope in predicting whether or not someone chose, $B = 5.88$, $z = 0.84$, $p = 0.40$. However, Hypothesis 2e also predicted that point-slope would interact with model-implied quality in order to predict subjective quality ratings, as was shown via exploratory analysis in Study 2. This effect was significant in Study 3 as well, $B = 38.41$, $t = 2.33$, $p < 0.05$, confirming that positive slopes make good options seem subjectively better. The three-way interaction of point-slope, model-implied quality, and amount of delay in predicting subjective quality (marginal in Study 2), was significant in Study 3, $B = 15.51$, $t = 2.10$, $p < 0.05$, once again indicating that positive slopes make good options seem even better when more options have been viewed.

Hypothesis 2f predicted that effects involving "objective quality" in prior studies would replicate using my model-implied quality metric. Indeed, a the results reported above suggest that they have: Hypotheses 1d, 2b, 2c, and 2e showed the same effects for model-implied quality as Studies 1 and 2 had shown for objective quality, while only Hypothesis 2a did not.

Hypothesis 3: Individuals Will Differ in How They Utilize Distributional Information

Individual difference measures were given one final try for predicting decision-making behaviors in Study 3. Once again, the scales showed alphas and correlation patterns commensurate with prior work (as discussed in Studies 1 and 2 above), though with slightly lower alphas in some cases (likely due to the smaller sample size in Study 3). As in the earlier two studies, nearly none of the predicted effects were present: Decision aversion, maximizing tendency, and vigilance did not predict choice quality or delay (by either measure of decision aversion, either measure of maximizing, either measure of quality, or any measure of delay). While some of these effects are occasionally marginal or significant, fewer than 1 in 20 predicted values were significant (consistent the significant effects being due to type-I error), none of the significant or marginal effects replicated across measures of the same construct (e.g., they would be present for only Schwartz et al.'s 2002 maximizing scale but not Diab et al.'s 2008 scale), nor did they replicate the seemingly spurious results from Study 1 or Study 2.

The additional hypotheses involving individual differences from Study 3, however, deserve additional attention: I predicted that those higher in numeracy would make greater use of the cues available (showing a higher mean or sum of the standard betas extracted from the model-implied quality model; Hypothesis 3g). This was tested via linear regression using numeracy to predict the sum of the standardized betas from a linear model predicting each individual's subjective ratings from the personability and financial stability ratings for each viewed option. This result was not significant,

$B = 0.007$, $t(48) = 1.12$, $p = 0.27$. However, my final prediction was that the more numerate would show more *consistent* preferences—rather than just showing a stronger relationship between the cues and the criterion, their individual model-implied choice models would show less error. I tested this by taking the adjusted R^2 (adjusted, rather than raw R^2 was used due to the fact that some models had very few options, indicating a likelihood of overfitting) from each individual's linear model and predicting this value from their numeracy score, using linear regression. Indeed, numeracy was a positive predictor of adjusted R^2, $B = 0.03$, $t(47) = 2.31$, $p = 0.02$, and remained significant and positive even when amount of options viewed was added as a control.

Discussion

Study 3 provided support for the two basic hypotheses of this thesis, replicating the results found in Studies 1 and 2. Study 3's extension of Studies 1 and 2 indicates that people both seek and benefit from gaining knowledge of a distribution of options in contexts with real-world appeal, such as finding a roommate. Study 3's results also show that distributional knowledge is desirable and helpful for a variety of distributions, for example, normal and uniform distributions. These distribution types were not observed to differ from each other in any hypothesized way (Hypothesis 1c), though they did show slightly different patterns of support for Hypotheses 1 and 2. For example, Hypothesis 2b, which examined the effect of delay on quality separately within each distribution, showed a significant negative linear effect for the uniform distribution for objective quality (though there were no effects of objective quality in Study 2), whereas the normal distribution did not show this counter-hypothesized effect.

Study 3 also replicated the results of Studies 1 and 2 even though Study 3 utilized quality metrics on two dimensions (both personability and financial solvency). In other words, we find that choices are still (marginally) better, according to direct ratings by participants and linear models built from the participants' many ratings, when participants delayed more (Hypothesis 2b), even though participants had to attend to two distributions of option quality. This means that the desirability and utility of gaining distributional information is present even for multi-attribute choices.

Not everybody makes the same use of the multiple cues, however: Those higher in numeracy made better use of the cues, as their subjective ratings were more accurately predicted by their model-implied choice than those who were less numerate. It is possible that this could explain the failure in Study 3 to replicate the "learning effect" whereby subjective ratings grew closer to objective quality in Studies 1 and 2. To explore this, an additional analysis added an interaction of numeracy with Hypothesis 2a, allowing a test of whether the relationship between time and the correspondence between subjective rating and model-implied quality was higher for the more numerate. However, this interaction was also not significant.

Study 3's results also extend many findings of Studies 1 and 2 by showing that the result holds for hypothetical tasks of greater ecological relevance: Study 3 examined a task participants may have had experience with (roommate choice); at the very least, it is likely that most participants (who were college students) would have at least discussed the perils of roommate choice with peers. This is notable because it extends the findings of Studies 1 and 2 not only in a more ecologically relevant direction, but also replicates the findings using a completely different quality metric: These effects are present when participants are judging and deciding among hypothetical roommates as well as when they are judging and deciding among how much time to actually (non-hypothetically) spend completing math problems.

Study 3 also replicated the many results of Studies 1 and 2 that dealt with objective quality by using a model-implied quality metric (Hypothesis 2g) —not only are participants obtaining higher-quality outcomes when they delay choice somewhat, but these outcomes are more consistent with the preferences they reveal over multiple ratings (whether or not these ratings are consistent or whether they are aware of them), by examining idiographically defined quality metrics (i.e., model-implied quality).

Study 3 replicated results from Study 2 regarding the relationship between the slope of option sets and peoples' perceptions of quality (Hypotheses 2e and 2f): When options are "getting better," people like the options more, over and above the effect of the options being "good" in an absolute sense.

Individual differences in Study 3, however, found largely the same fate as they did in Study 2: inconsistent and sporadic support (if any). This third failure to appear in three separate data sets across both hypothetical and non-hypothetical decisions about multiple topics shows least consistency of the unreliability of these relationships. This is discussed in greater detail in the General Discussion below.

CHAPTER V

GENERAL DISCUSSION

Overview

Over the course of three studies, I have shown consistent support for two basic hypotheses regarding the manner in which human decision makers form judgments regarding the options available to them: That people delay choice in order to gain knowledge of how the options available to them are distributed, and that doing so is potentially advantageous. People seek to view additional options when they are available, and this viewing appears to lead to learning distributional qualities of the options which people then appear to utilize in their evaluations of each individual option to which they are exposed: They continue to view options early on, not solely relying on the quality of the options they see (Hypothesis 1a, supported in all studies), they view more options when the option set is variable (Hypothesis 1c, supported for the low-mean condition in Studies 1 and 2), and as more options are viewed, the quality of the options becomes a stronger determinant of whether options are chosen (Hypothesis 1d, for subjective quality of options in all studies, objective quality of options in Study 1, and model-implied quality of options in Study 3).

People also gain a better understanding of the distribution of options as they view more options, as evidenced by a greater correspondence between subjective ratings and objective quality as additional options are viewed (Hypothesis 2a, supported in Studies 1 and 2 but not using model-implied quality in Study 3)—this provides evidence that people are indeed learning the distribution as they view more options (whether or not learning the distribution is an explicit goal). This provides a nice parallel to Hypothesis 2b, which predicted that those who view more choices (up to a point of diminishing returns) would end up with better outcomes, from both an objective and subjective perspective. For objective quality, this result was weakly supported in Study 1, replicated in Study 2, and conceptually replicated in Study 3 (using model-implied quality). For subjective quality, this result was supported in Studies 1 and 3 (but not in Study 2).

As additional evidence for the claim that distribution learning leads to better outcomes, Hypothesis 2c predicted that the relative value of viewed options would be a determinant of subjective quality over and above the effect of objective quality. This result was significant in all three studies, indicating that participants were sensitive not only to the ground-zero objective quality of each choice, but also to the relative objective values. This result dovetails with Hypothesis 1b (supported in Study 1, but not Study 2, and untested in Study 3), which predicted that the absolute objective value of the options offered would not even matter, though Study 2 suggests that at least in some cases the objective value does matter (such that people delay more when the mean is higher). If participants are more likely to choose items that are excellent relative to the history of items seen, it is difficult to argue that the history of options viewed is having absolutely no effect on participants. Finally, I showed that participants whose eventual choices were objectively better were affected by the slope of the options they had seen—options that were good and came from a positively-sloped distribution were rated as subjectively yet-better than options that were less good or which came from a less-positive (or negative) slope (Hypothesis 2e, tested and supported in Studies 2 and 3). These effects are present both within and between subjects (the hypotheses were supported between-subjects in Study 1 and replicated within-subjects in Studies 2 and 3), using both hypothetical and real decisions (the hypotheses were supported using real decisions in Studies 1 and 2 and replicated using hypothetical decisions in Study 3). In this chapter, I will discuss the implications of this research, limitations of this research, and proposed future work to extend this research.

Implications

The most notable implication of the work described here is that delay, specifically decisional delay, may in some cases be wise. The literature on procrastination and decisional procrastination has illustrated the dangers of delaying actions and decisions, but to date without addressing the other possibility—that it may sometimes be wise to do so. One instance in which it may be wise to do so is in order to gain a better understanding of the distribution of possible options. In many real-world decisions, decision-makers are posed with the option of waiting to view additional options, perhaps at some cost (e.g., an opportunity cost such as risk of losing currently available options). This dissertation suggests that before the distribution of options is known, people will seek to learn it, and this delay may (for some decisions) lead to better outcomes, ostensibly due to a more nuanced notion of what counts as a "good" option.

In the studies in this dissertation, I generated random sequences of options pulled from various distributions, including normal and uniform distributions, with varying means and variances. Evidence that participants effectively "learned" the distribution comes from examining the extent to which their subjective ratings of how good the options were vary as they view additional options. For example, I show that these subjective ratings corresponded more to objective values as additional options are viewed, and to relative objective values over and above the absolute objective value, derived either from the structure of the decision task or from a model of the participants' preferences revealed by rating all of the options each participant viewed. As participants viewed more options, their subjective ratings more closely approximated the objective values, indicating that they were internalizing the relative quality of the options—effectively, learning how "good" an option is relative to the other options they saw available to them.

This work extends several threads in the decision-making literature. Corbin and colleagues (1975) showed some tentative effects of distributional slope and variability—that people are more likely to wait for a third option if the first two options are similar than if they are dissimilar. My results, in contrast, show that more variability of viewed options leads to more delay. This could be for two reasons: My participants had up to 100 options to view, while Corbin and colleagues' participants had only five; Corbin and colleagues also rewarded their participants only for choosing the best option, while my participants (in Studies 1 and 2) received the option they selected, thus receiving simply the reward associated with that particular option (as is more frequently the case in everyday decisions). Corbin and colleagues' results may suggest, however, that my participants were sensitive only to big jumps, or choice patterns in which one option greatly exceeds the prior option. However, this process would produce a choice pattern fully determined by option quality, and not the amount of delay—evidence in favor of Hypothesis 1a contradicts this. While direct comparisons of the success of my participants to those in Corbin and colleagues' research are not possible, the fact that participants who delayed some (but not too much) attained better outcomes in the current studies suggests that decisional delay is fruitful for more ecologically valid decisions.

Dhar (1997) also discussed waiting in order to view more options. Dhar's studies involved having participants view several options, and examined what composition of options led participants to wait for more options at greater rates. Dhar's explanations, however, operated at the level of the options available to participants—he did not address the possibility that participants were viewing the options available as markers of the sort of options one could hope for by delaying. For example, when participants viewed one superior option and one inferior option, delaying choice was less likely than when two equal (but superior) options were available. As such, my results extend his by offering a broader explanation for why people seek a "no-choice option": A high-variability option set leads to delayed choice, because it indicates to participants that there may be better things to come. Like Corbin and colleagues' study, a simple "one option stands out" explanation for Dhar's results does not explain the results shown in Hypothesis 1a—that amount of delay is significant over and above the effect of objective and subjective quality.

In a broader sense, I have shown that people desire more information about the options available to them. This is consistent with Shafir's work on framing effects (Shafir, 1993; Tversky & Shafir, 1992): Shafir shows that options with reasons for selecting them (and rejecting them) are more likely to be selected (and rejected) than less "enhanced" options. This suggests that people prefer options about which they have more information. An interesting test of the hypothesis that information (and perhaps feeling informed) is the key would be to compare Shafir's "impoverished" option to an option "enhanced" with irrelevant information.

The work of Patalano and Wengrovitz (2007) could also be informed by these studies: They examine how indecisiveness interacts with the riskiness of the choice set in order to predict decisional delay and quality of choice. These results could be explained in delay-for-information terms by simply noting that those who delayed choice ended up with better options, as I have shown for additional decision contexts. Patalano and Wengrovitz showed that indecisive individuals behaved the same in "risky" and "no-risk" conditions, while more decisive individuals modulated their decisions in accordance with the risk of losing a good option. When Patalano and Wengrovitz tailored their decision to have the ideal choice occur later in the choice set, effectively rewarding risk-taking, they found that less indecisive individuals benefitted. They used the same individual difference measures of decisiveness as the current studies, but with a median split to define groups as "decisive" and "indecisive." This could partially explain their success at finding that individual differences exist, although in most cases median splits lead to spurious results when they conflict with continuous analyses (c.f. Howell, 2008; Judd, McClelland, & Ryan, 2009). Even so, an exploratory analysis examining median and lower-quartile splits on decision aversion (to mimic Patalano and Wengrovitz's technique) did not produce a difference in decision delay for those who were more decision-averse (by either measure).

As discussed in the Introduction, the Indecisiveness scale (Frost & Shows, 1993) is also arguably a measure of compulsive decision aversion rather than simple inability to decide: Patalano and Wengrovitz's results may indicate that those who are unable to decide simply *don't* decide, and that for those above the median on this scale, the "simply don't" was not affected by the risk manipulation in their course-choice paradigm. Thus, the indecisive people fared better, because the best option was designed to appear at the end. If there is more at work, however, my research suggests a further question: What leads people to delay? Patalano and Wengrovitz's research (more so than mine) was able to predict likelihood of delaying from some (but not all) measures of decision aversion as an individual difference variable (see Limitations, below), but an articulation of two kinds of indecisiveness—decision avoidance versus information seeking—may provide a more nuanced framework for their paradigm (see also Anderson, 2003). Perhaps some of the "indecisives" who achieve better outcomes do so because they are seeking information, rather than avoiding choice. Whether this practice is wise would then depend on factors regarding the precise decision being made: what the cost of waiting is, whether the decision-maker learns the distribution quickly (or correctly), and of course whether better options would be available to those who wait.

This work also bears some implication for researchers of delay in general. Taken in a broad light, my studies have shown that participants seek (and may benefit from) delaying choice in order to gain additional information about the options available to them. While my study focused specifically on gaining information about the distribution of options, any scenario in which gained information could be used or useful could constitute a similarly wise reason for delay. Some researchers have examined delay as a means for deciders to gain more information about already available options (rather than to explore *additional* options), suggesting that not all information that might be sought is necessarily wise: Bastardi and Shafir (1998) showed that the mere act of delaying led delayers to over-rely on the information gained via delay. However, this finding suggests these participants are not just "spinning their wheels" or looking for reasons to wait. Whereas Tykocinski and Ruffle (2003) show evidence that people do indeed simply prefer to wait, Bastardi and Shafir's results indicate further that participants rely on the information gained via delay. This information affects their choice inappropriately, as participants given this information up front (i.e., without having to delay, and also without being given the option to delay) make choices that are less based on the information. Other research has suggested the same thing: Gaining (or just having) information leads people to rely on it to a fault (Dawes, 1975; 1994). Together, this suggests that the "wisdom of waiting" may depend on whether the information gained is appropriate or wise to rely upon, and also how well it is used.

The literature on procrastination may also need to be updated to account for these results. Notably, there may be reasons for delaying action that are not inherently bad, and good things (in some cases) do come to those who wait. Definitions of procrastination that inherently presume that delaying action (of any sort) is unwise should take heed of the case of delaying decisions, which counts as an instance of delaying action, and thus sometimes counts as at least one instance of beneficial "delay." If delay is not assumed to be inherently negative, then the question for procrastination researchers would be how to define "procrastination" in a non-circular manner: One cannot call an action "procrastinated," for the purposes of research, without the person delaying action believing, before the delay, that the delay will lead to no good. To do otherwise is to call wise delays "procrastination" until they have proven themselves beneficial (or to be unable to call anything "procrastination" *a priori*). Many procrastination researchers have avoided this pitfall by studying only actions that common wisdom shows (and thus the participants themselves likely believe) to be a bad idea to delay, such as signing up for classes (Ariely & Wertenbroch, 2002; Tice & Baumeister, 1997). Requiring the expectation that delay lead to poorer outcomes was made explicit by Steel (2007), although even his explicicitness also poses problems: Many so-called "procrastinators" and those who delay to their detriment do not, in fact, expect to be worse off for delay. The planning fallacy (Buehler et al., 1994) illustrates one example of people in general over-delaying in a manner often called "procrastination" but under circumstances where they have no expectation of the delay leading to poorer outcomes. Similarly, many people may believe (or fear) that delaying a certain action will be harmful, even when there is no evidence that delay will hurt the particular action they must complete. In my studies, participants were not normatively worse-off for waiting (because option quality was randomly generated), but I still showed a positive linear effect for delay on quality (indicating that delay is overall helpful) as well as a negative quadratic effect (indicating that delay becomes less helpful eventually). While I do not argue that delay is an inherently a good idea, I argue that it is not an inherently bad one; this means that the findings in the procrastination literature may not apply to the delay of action in general. Rather, if the study of procrastination is limited to actions that participants believe to be disadvantageous to delay, then it cannot be generalized to delaying action in general, which may show that a "wait a bit…but not too long" strategy would lead to the best outcomes. Prior procrastination research has selected methodologies almost entirely designed to produce poorer outcomes when participants wait (c.f. Ariely & Wertenbroch, 2002; Tice & Baumeister, 1997), and has not analyzed "amount of delay" either in a continuous manner or with a quadratic component.

Another implication is the inconsistency of results found for subjective and objective measures of option quality. Though highly correlated, these measures occasionally behaved differently in terms of predicting choice processes and choice quality. For example, Hypothesis 1d predicted that "quality" would become a stronger predictor of whether an option is chosen as more options had been viewed (and then taper off); this was supported for subjective quality in all studies, but not so consistently for objective quality. Hypothesis 2b, which predicted that quality of the option eventually chosen would increase for participants who delayed (and then taper off); this was well-supported for objective quality in Studies 1 and 2 (and untested in Study 3), but received the weakest support for subjective quality, which was marginal in Study 1 and nonsignificant in Studies 2 and 3. These inconsistencies are likely due to an effect observed in Hypothesis 2a: That subjective and objective quality are collinear, but that they become more highly related as additional options are viewed. Hypotheses 1d and 2b (mentioned above) both examine linear and quadratic effects of delay on quality, which might explain the inconsistency: The act of choosing is, perhaps unsurprisingly, more strongly influenced by subjective quality than by objective quality. However, it is objective quality that differentiates those who have waited to choose from those who chose early. This can be explained by noting that the two models were different: Hypothesis 1d looked at every option chosen or rejected, in order to model what predicts choosing. This analysis was conducted within subjects, indicating that when an option has a higher subjective quality for the individual in question (eliminating noise due to the fact that people use Likert scales differently), choice is more likely; however, as objective quality varied randomly within subjects, some objectively excellent options may have been passed over during the learning phase (i.e., before subjective and objective quality had grown close together; c.f. Hypothesis 2a), reducing the predictive power of objective quality in the model for Hypothesis 1d. In the results for Hypothesis 2b, which examined quality of only the chosen option, more variance was present in how the scale was used (as this was measured between-subjects in Study 1 and with much less within-subjects data in Studies 2 and 3), reducing the predictive power of subjective quality across people. Objective quality was measured in the same units for all participants, so it did not become more "noisy" for the same reason. In fact, the opposite may have occurred: Because only chosen options were modeled in my tests of Hypothesis 2b, the within-subjects variance in objective quality was also ignored—if a person ignored an option of "25 minutes of free time" three times before eventually realizing that 25 is a relatively "good" number, and then choose it the fourth time it came up, this would show up only as a "25" in a model of "quality of chosen options," rather than showing up as an effect akin to rejecting a "25" three times then choosing it once in a model of "whether an option will be chosen."

A larger question, however, is what conclusions one is to draw from the apparent lack of individual differences in predicting many aspects of the delay of decisions. Across three studies, results for individual differences were inconsistent or entirely absent. One possible explanation for this is that the individual difference measures I used did not function as expected (i.e., that they failed to measure the construct in question in a valid manner). Notably, these measures did show the same intra- and inter-measure relationships as they had in prior work, suggesting that they were at least reliable: The two maximizing scales correlated significantly, the two decision aversion scales correlated significantly, and these scales all had internal consistencies consistent with the papers that presented them. This correlation structure indicates that, in my study, these scales were similar constructs to the ones measured in the original papers.

However, the scales might still lack some validity in terms of being general measures of the construct of decision aversion, indecisiveness, or decisional procrastination. Frost and Shows (1993) designed their scale in order to measure procrastination, sub-clinical levels of obsessive-compulsive disorder, and hoarding tendency; they then argued that the scale was valid on that basis. The scale was also shown to predict how long people spent making decisions; the metric for "taking longer" to make a choice was a reaction-time measure of choice between various things: clothing articles, courses to take, restaurant menus, or free-time activities. Over all comparisons, indecisive people took 60% longer to decide. However, as discussed in the Introduction above, taking more time to decide may involve considering more information or considering that information more carefully – simply taking more time is not the same as procrastination, and is not necessarily negative.

Furthermore, Frost and Shows' (1993) scale items conflate information seeking with decision aversion: those who self-identify as "putting off" making decisions (item 1 on the scale) are likely owning up to a decisional procrastination tendency, with those who disagree that they "…always know what they want" (item 2) or may do so because they are careful choosers. Item 3 asks participants whether they find it "easy" to make decisions, conflating the scale with a third construct: ease of making decisions. Some people may simply find decision-making difficult because they carefully consider every decision they make, without necessarily delaying choice in order to do so.

My studies also use this scale to predict delay of choice when participants were given a reason to delay—to seek better options. This sort of delay is very different from the seconds-spent-choosing method used to show validity of the Indecisiveness scale. In sum, it is possible that validity concerns due to undermeasurement of the "Indecisiveness" construct could have led to my failure to find relationships between indecisiveness, decision delay, and outcome quality in an optimal stopping paradigm. I also note that this measure, which purports to measure a trait, showed mean-level differences between participants who completed the questionnaire as part of my study and participants who completed the questionnaire as part of the unrelated General Survey, with participants self-reporting themselves as less indecisive in the "study" version of the questionnaire than in the General Survey. This inconsistency did not contradict any of my hypotheses, nor did it interact with any relationships I hypothesized, but this nevertheless provides evidence that the construct itself is measuring more than trait-level indecisiveness.

The other measure of decisional procrastination, the "Decisional Procrastination" subscale of the MDMQ (Mann et al., 1997, see p. 12 for scale items) is much shorter, and used differently troublesome items. Though this scale honestly admits to being designed to measure decisional *procrastination* (and thus excuses itself from failing to measure beneficial delay), 40% of the five scale items pertain to procrastination of action rather than decisions ("Even after I have made a decision, I delay acting upon it") or judgments regarding the effect of delay rather than delay itself ("I delay making decisions until it is too late,"). The six-item "vigilance" subscale of the MDMQ, rather, attempts to model careful delay directly, but leaves aside the question of delay: People may "like to consider / find out the disadvantages of all of the alternatives" (items 1 and 2) or "like to collect a lot of information" (item 4) but be unwilling to commit time in order to examine alternatives; others may "try to be clear about their objectives before choosing" or "take a lot of care before choosing" (items 5 and 6), but whether this requires any marked time commitment likely varies by person and by decision topic. This scale, like the decisional-procrastination scale, also suffers conflation of the decision task and judgments regarding the process: Careful *people* may "consider how to best carry out the decision" (item 3), but not delay once it is time to decide. Indeed, the fact that these two scales both purport to measure subconstructs of decision delay but are reported in Mann and colleagues' (1997) paper as correlating significantly at $r = -0.32$ raises some suspicion.

One could also argue that decision-aversion scales may have been poorly chosen as they only measure the negative part of decisional delay, however, the result that these were nonpredictive of how long decisions took is still fairly concerning: Those who are high in decision aversion should have found choosing aversive in my studies (which undeniably involved making "decisions"), and thus would be expected to avoid or procrastinate choice (either by choosing hastily or by dawdling). Together, this suggests that these scales may not yet be ideally suited for measuring decision aversion or delay for a paradigm such as mine.

If these constructs were correctly (or at least adequately) measured, then the question of why they failed to relate in the predicted way is harder to answer. Patalano and Wengrovitz (2007) showed no significant main effect of indecisiveness on decision delay, but they did show an interaction with their "risk" variable: Indecisive individuals were not affected by their "risk" manipulation, while decisive individuals were. I take special notice of the first result: that the indecisive did not, overall, delay longer than the decisive. This is the same null result that I found in all three of my studies, indicating a fourth set of participants and a third decision-delay task in which the indecisiveness and decisional procrastination scales did not directly predict delay of decisions. Patalano and Wengrovitz's (2007) results also used a median-split approach, treating the 50% of participants with the highest indecisiveness scores as "indecisive" and the others as "decisive," which is widely considered to be an analytic technique that reduces power overall while increasing the likelihood of "capitalizing on chance" (Judd, Ryan, & McClelland, 2009). I note, however, that Hypothesis 3f from the current Study 2 was based on their findings, and designed to produce a conceptual replication of their work: Individual difference effects (such as decision aversion) would be enhanced when decision-makers were motivated to choose quickly (an analog to the "risk" manipulation of Patalano and Wengrovitz). Though this hypothesis was considered to lack support due to widely inconsistent findings, one of the few "significant" results (rendered insignificant once controlling for family-wise alpha error) found was the exact opposite of Patalano and Wengrovitz's: The more decision-averse were marginally *more* affected by the "cost" manipulation than those who were less decision-averse, whereas Patalano and Wengrovitz's result which showed that indecisive individuals were *less* sensitive to their risk manipulation. Further, this result in my study was marginal only for the MDMQ's Decision Procrastination scale (Mann et al., 1997)—not Frost and Shows' (1993) indecisiveness scale, whereas Patalano and Wengrovitz showed their result for the latter scale and a null result for the former. This result is consistent only with the inconsistency that seems to occur more often than would be desired with this set of individual difference measures.

A similar set of arguments can be made for the measures of maximizing tendency that I used: Diab and colleagues (2008) provide an eloquent (albeit scathing) critique of the classic Schwartz and colleagues (2002) maximizing scale, and do a considerable amount of work to produce a more psychometrically sound scale; however, evidence that their scale is more reliable does little to make a case for validity. In fact, their scale correlated with only one of the scales that the original scale did (suggesting only divergent validity), and with only two of five behavioral measures of maximizing; it also showed mean-level differences between participants who completed the measure as part of my study versus the unrelated "General Survey" (those who took the General Survey were more oriented towards satisficing than those who completed the measure during the study, according to Diab and colleagues' scale but not Schwartz and colleagues' scale), raising the same validity questions as discussed above for the Indecisiveness scale. Further, though the construct of "maximizing" versus "satisficing" is easy to define, it is not clear how exactly the construct of maximizing should be applied. Patalano and Wengrovitz (2007) point out some trouble with use of the construct in the realm of "decision delay" quite nicely (albeit unintentionally): They argue that one reason that maximizers delay less than satisficers in their study is that the maximizers have a low notion of what an "optimal" choice is, while the satisficers have an oddly high "good enough" cutoff. Indeed, someone anywhere on the maximizing/satisficing continuum could easily make the same choice (choose the same option, after the same amount of delay) as someone on the opposite side of the continuum, with one person arguing that they believed the option chosen to be "maximal" while the other argues that he or she chose it because it was simply "good enough." The theory behind satisficing as a strategy, rather, relies heavily on satisficers determining a cutoff value *a priori* and then following it: Simon's (1955; 1956) work utilized a formal set of rules and showed that these rules were rational within reasonable bounds, while other researchers have shown that human judges are loath to trust anything important to a formalized rule (Dawes, 1994). Validity studies involving these scales have never, to my knowledge, examined the extent to which those who score in the "satisficing" range have such cutoffs in mind while deciding, nor whether satisficers habitually choose the first option that exceeds this cutoff (choosing this option even if it is the first option, choosing this option across all domains from purchasing new homes to purchasing sandwiches, etc.). In sum, it is possible that these scales (and the decision-aversion scales) failed to predict choice delay or quality in my studies because the scales lacked some validity.

Another possible explanation for the inconsistency of these results is one pertaining to the experimental methodology. Given that the studies at hand were conducted in a laboratory environment with clear instructions regarding study methods, it could be that these methods left little room for individual differences to affect choice behavior. This explanation is plausible, though the three studies presented here do produce a reasonable amount of variation in study methods. However, the optimal stopping paradigm or range of options may also have reduced the ability for individual differences to express themselves. For example, one could view the studies reported here as "guess the best" challenges, rather than decisions, which may have circumvented individual differences regarding decision-making style. If these individual differences are only expressed when the decision-maker has the subjective experience of "making a decision," then they would only be expressed in instances of "choosing," rather than "rejecting." Similarly, the "choose or reject" methodology, which required that participants actively push a button labeled with their choice, may have more closely resembled a series of decisions (one choose/reject decision for each option) than a single decision (of which option to choose). If so, then individual differences expected to have an effect on *each* choice, which may not necessarily have translated into delay for the larger meta-decision. While these possibilities do not challenge the results shown in Hypothesis 1 and 2, they may serve to explain why the individual difference effects were so weak.

A final, and perhaps more likely explanation for the consistent inconsistency of my individual difference results is drawn from the specific predictions I have made: I am predicting behavioral outcomes based on individual differences. This is exactly the sort of prediction that has led many researchers to malign the very study of individual differences: Trait-like measures of personality do a very poor job of predicting single behaviors (Mischel, 1968), with correlations frequently being less than .2 in magnitude. This implies that the correlations I observed, though small, may be reliable—this effectively implies that sample size may be a concern. Studies 1 and 2 had an n of around 100, while Study 3 had an n of only 55; for $n = 108$, a correlation would only be significant if it exceeded 0.18. As such, I conducted a set of exploratory analyses in which I combined the data from all three studies to re-examine these individual differences: Subjective ratings, objective quality, and amount of delay were standardized within each experiment /condition cell, and hierarchical models were used to account for the fact that some participants (those in Studies 2 and 3) completed more than one decision; this resulted in a total of 624 decisions made by 144 participants.

This aggregated data set was then used to re-test some of my individual difference hypotheses. Hypothesis 3a, that those averse to decision-making will be more extreme in terms of delay, found support: This was significant using the MDMQ Decision Procrastination scale, $B = 0.23$, $z = 2.08$, $p = 0.04$, and marginal for the Indecisiveness scale, $B = 0.10$, $z = 1.92$, $p = 0.055$. Additional analyses examining length (rather than extremity) of delay showed positive but non-significant results, indicating that extremity is the variable of interest, which in turn indicates that hasty and prolonged choice are both possible outcomes of indecisiveness: Those averse to decisions either delay a long time, or they choose hastily. Hypothesis 3b predicted that more vigilant decision-makers would delay longer or attain better outcomes. When controlling for decision aversion, this hypothesis was not significant using the aggregated data set. Hypothesis 3c predicted a null or negative effect of maximizing on outcome quality and a positive effect of maximizing on delay; out of the eight analyses that tested this hypothesis (due to multiple measures of maximizing, quality, and delay), only two were significant. Those higher on the Diab and colleagues (2008) measure of maximizing showed significantly higher delay and also significantly higher extremity of delay (although this effect was positive only in Study 1, slightly negative in Study 2, and very close to zero in Study 3), lending tentative support to this hypothesis and supporting Diab and colleagues' arguments that Schwartz and colleagues' (2002) scale may not be a very good measure of maximizing tendency (results on the Schwartz scale were in the same direction as the Diab scale, but non-significant; p's > 0.30 in all studies and in aggregate). Hypothesis 3d and 3e concerned predictions of sensitivity to the distribution, though I was unable to aggregate this variable across studies as the differences in distribution indicate that the "sensitivity" variable has a fundamentally different meaning across distributions and studies.

Thus, in some cases, my hypotheses obtained support, but only when the sample size was increased to a level able to detect weak effects. There is some danger, however, in the process of combining data from several studies into a single analysis—the studies differed in terms of their methods, the decisions made by participants, and the distributions viewed, though this aggregation treated these potential effects as either random or nonexistent. As such, the question of whether the metrics from the diverse studies are inherently comparable is a valid one; future work on the topic is encouraged (especially for those who have the ability to recruit many subjects or to examine many decisions for each)

My additional hypotheses regarding individual differences included a prediction that more numerate individuals would make better choices—this did not pan out, though I did show evidence that numerate individuals had a significantly higher adjusted R^2 for predicting subjective ratings from model-implied quality than the less numerate. This indicates that more numerate individuals' subjective ratings are more strongly determined by their underlying (i.e., revealed) preferences than the subjective ratings of those who are less numerate. This is consistent with prior research on numeracy (e.g., Peters and colleagues, 2006), which shows these individuals to be more comfortable with (and competent at) making use of numbers. This result has implications for anybody who studies decision-making or decision satisfaction using subjective ratings of decision quality. Ratings produced via Likert-scale "subjective quality" queries such as those that I used are noisy measures of subjective quality, and further, the precision of these estimates varies as a function of numeracy: Subjective quality measures are predictably heteroscedastic, such that the error is not random, but rather is a function of numeracy.

In sum, the individual differences hypotheses were not well supported, though an aggregation across all three studies provided a large enough sample size to suggest that some may indeed be valid. Further research is needed to test these hypotheses via methodologies that allow for higher power. None of the sub-hypotheses of Hypothesis 3, that individuals will differ in terms of their distribution-learning processes and their use of distributional information, were actively disproven; and only more research can answer whether the hypotheses involving individual differences are correct.

Together, the results of this set of studies imply that, when humans can delay choice in order to gain a better knowledge of the set of options from which they are choosing, humans (in some contexts) do indeed wait. I have also shown evidence that this sort of decisional delay is sometimes adaptive, in the sense that some amount of delay leads to better outcomes for some decisions, but also disconfirmed the implicit hypothesis in much of the procrastination literature (decisional and otherwise) that delaying choice is inherently dangerous or harmful. These results suggest that whether to wait before choosing is a real and valid question—almost everyone does, and it sometimes helps.

Limitations

This set of studies does contain some flaws, however. First and foremost, it utilized only one decision-making paradigm: optimal stopping. While I believe that this paradigm offers a great amount of insight into decision-making processes through its apparent relation to many real-world decision tasks, multi-trait and multi-method approaches are inherently more desirable.

The first limitation of my paradigm is that the methods used here never had participants choose among several options that were all simultaneously available: The only options available to participants at any one point were to "choose" or "reject," whereas many real-world decisions involve a choice among many options at each time point. For example, in a more ecologically valid roommate-choice study, one might interview roommates all week, telling them all, "I'll let you know on Friday." This sort of approach would be better represented by an optimal stopping paradigm in which options become unavailable stochastically. Patalano and Wengrovitz (2007) do this in their "course sign-up" paradigm, in an effort to more closely model the way that courses are chosen. However, their study used only one fixed set of courses at each timepoint—and the set of courses was designed beforehand to produce certain outcomes, while their participants (students, all at the same school) likely had experienced the same course sign-up procedure that might lead them to expect something else. In other words, they had knowledge of the distribution of courses and how courses "fill up," based on semesters prior. My first study also utilized only a single ordering of available options, though I note that Studies 2 and 3 did not; similarly, the ordering used in Study 1 was random, and thus not manipulated to test a specific hypothesis, rather, it was generated randomly.

Similarly, Study 3 extended Studies 1 and 2 by adding a second attribute of "quality," or desirability of an option. However, two indicators (while better than one) still vastly underrepresents the number of variables simultaneously determining "quality" in everyday decision-making (Kramer & Hodges, in prep), and additional research will also be necessary to flesh out the role of many quality indicators, which may be correlated (either positively or negatively) depending on the choice context.

These paradigmatic issues may be partially responsible for some of the inconsistencies in the findings for several hypotheses. The lack of support for Hypothesis 3 has already been discussed, but it must be noted that Hypotheses 1 and 2 also did not find uniform support. Hypothesis 1c predicted that more variable option sets would lead to longer choice was significant only in the "low mean" condition in Studies 1 and 2. As discussed in Study 1, above, the best reason for this is that ratings in the high-mean occasionally "hit ceiling," or were in a sense the best possible outcome (for example, 30 minutes of free time and 0 minutes of math problems). The methodology in Study 3, in which there was no "ceiling effect," nearly provides for the testing of this explanation, but did not produce a significant result. However, Study 3 did not manipulate variability of option sets directly, rather it did so by varying the actual distribution (to be either normal or uniform; the uniform distribution had theoretically a bit over twice the variance), though the variance manipulation could have been either too weak to be detected or the distribution shape could have rendered this effect undetectable. Future work might consider comparing two normal distributions (or two uniform distributions) which have the same mean and different variances to see when and how this effect appears, using an unbounded value metric (such as money gained or lost), rather than the bounded one used in Studies 1 and 2 (number of minutes, range between 0 and 30).

Hypothesis 1d also received mixed support across studies: It predicted that quality metrics would become stronger predictors of choice as more options were viewed. This was supported in Study 1 (for subjective quality and objective quality) and Study 3 (for subjective quality and model-implied quality), but only marginal in Study 2 (for subjective and for "raw" objective quality, but not for discounted objective quality). This pattern of "weak results" also applies to Hypothesis 2b, which predicted that delayers receive better outcomes: This effect was significant or marginal for both subjective and objective quality in Study 1, and significant in Study 2 for objective quality (but not for subjective quality), and marginal or non-significant in Study 3.

Some of these inconsistent results may well be due to the methodological differences among studies. For example, in Study 1, a single distribution was used, so effects dependent on delay were affected by the same order of options' qualities for each participant, making these results much easier to detect across people. The fact that these effects replicated only inconsistently could indicate that these effects depend on the actual distribution of options seen by the decision-maker. This may be due to the interaction between distribution learning and choice: As people learn the distribution, they are effectively rejecting options. When people delay decisions to see more options, the first part of that delay may be to learn the distribution, but the latter part of the delay may be to use that learned information to evaluate options in order to pick a superior one: People may gain information about the distribution, and then at some point begin evaluating choices relative to those they have already seen. Future research should examine whether there is a "cutoff point" (or range) where the decision-maker has started to feel comfortable choosing, perhaps by examining confidence in subjective ratings as an additional variable. If such a point exists, it would also be interesting to explore whether the relationship between delay (specifically during the learning period) is more clearly related to quality of the chosen option.

Above research considerations aside, I also do not claim that the results in this dissertation are sufficient to suggest delaying decision-making for any decision in any context. For example, many decisional contexts carry sufficiently steep costs for waiting that any delay (to see more options, to further deliberate, or even to blink an eye) may be unacceptable. Similarly, while I have shown distributions of available options to have been learned by those who delay choice, I have not undertaken any analysis of the accuracy of this learning. Some distributions may be inherently difficult to learn, while other distributions may learned incorrectly: For example, if humans are used to dealing with normal distributions, then distributions that appear to be normal but deviate in some uncommon way (for example, having odd skew or kurtosis) may be incorrectly learned as "normal," thus leading to improper judgment of the likelihood of options falling in the tails. In general, it is difficult to argue for or against delay in a general-case decision based on the results in this thesis; rather, the utility of delay is likely to vary across decisional contexts. My results, rather, serve to motivate the question and to suggest that delay is not an inherently bad strategy.

This set of studies also lacks some ecological validity in the realm of reality of the quality of options: The number of minutes used for free time was "real" in the sense that participants received the option they chose in Studies 1 and 2, and the roommate choice paradigm was "real-world" in the sense that the decision was designed to represent that which participants may face in everyday life, but these vignettes may be too "cold" to generalize very far from. For example, "personability" of a person one has just met is an inherently visceral or emotional reaction to that person (the literature on person perception is based largely on first-impression short meetings, see Ambady & Rosenthal, 1992, for a review); simply saying that it's "high" in a vignette may lead people to mis-weight its worth relative to how they would weight it in an actual person-to-person roommate interview with someone they would then rate at the same level.

Just as real-world decisions are multiattributive, the set of options available in each distribution will also follow many distributions not tested in this thesis. Broadly, one would expect that people would be best at learning (and choosing good options) when the distribution of options is familiar to them; Stewart, Chater, and Brown (2006) show a similar effect linking real-world interactions with money to the well-known "value function" that represents peoples' valuation of changes in value established in Kahneman and Tverksy's (1979) seminal work. This suggests that the experience of a distribution and familiarity with the distribution is very likely to be based on prior experience. Indeed, most participants have some experience with choosing roommates or at least thinking about roommate choice. The fact that Study 3's participants were college students (most of whom had likely had at least one randomly-assigned roommate as part of their college experience) may have predisposed the participants to presume a certain distribution of roommates, which may or may not have been normal or uniform (the two distributions used in Study 3). This presumption may have led participants to spend less time learning the distribution at hand or to use their knowledge of an alternative distribution in judging the relative worth of roommates: For example, if they believed most roommates to be poor and thus the distribution to be positively skewed, a middle-of-the-road (i.e., median) roommate presenting in the normal-distribution condition of Study 3 may have been chosen prematurely. This could explain why participants did not delay longer for the uniform distribution than for the normal distribution in the roommate-choice study, even though there is evidence that higher variability of option sets leads to greater delay in choosing (as shown in Study 1 and partially replicated in Study 2). However, this may not be a fatal flaw for the theory, so much as my choice of methods: Many real-world optimal stopping decisions are made in contexts in which the decision-maker is not completely new to the task. Most purchasers of airline flights, for example, are people who have purchased other airline flights in the past, and even the businessman of the classic vignette was purported to be choosing a secretary in order to replace (at least) one with whom he was already acquainted. Future work should consider using contexts for which known distributions exist, or empirically determining these distributions, for example, examining distributions of flight prices from one city to another, and then examining optional stopping behavior in a hypothetical "flight choice" decision among individuals who purchased such a flight, or by systematically evaluating the personability and financial stability of people who post to Craigslist looking for somewhere to live and then using the observed distribution in a replication of Study 3.

Indeed, if people are better-acquainted with some distributions rather than others, then the extent to which studies utilizing normal (or uniform) distributions can be used to illustrate the underlying processes of learning distributions is unclear. As we know from introductory statistical literature, assuming that data follows a normal distribution (when it doesn't) does not always lead to tragedy (c.f. Howell, 2008); this suggests that, even if human decision makers naturally presume that the options offered to them come from a non-normal distribution, presenting normally distributed options should still offer a (tinted) window into the processes at work. Indeed, this sort of expected/actual distribution mismatch could explain some of the weaker or more confusing effects illustrated by my studies.

Another limitation has to do with the fact that the randomness of the options shown to my participants may be inherently ecologically unsound, as very few sequences of options available in real-world decisions are actually generated by *random* processes. Even very hard-to-predict outcomes, such as airline prices or the stock market, are not generated randomly, but rather are generated in a manner unknown to the human judge. This is consistent with the noted human inability to identify or generate random sequences (c.f. Wagenaar, 1972) and the consistent search for patterns even in random data (Wolford, Newman, Miller, & Wig, 2004), perhaps synergistically so (one could argue that all of science is a search for patterns among a chaos that we believe to contain patterns). The human trouble coping with randomness (as well as other biases in judgment) suggest that distributions that are unbalanced with regard to option quality may also be unbalanced in terms of how (normatively) desirable learning the distribution would be, how precisely people prefer to learn the distribution before feeling comfortable choosing, how successfully people can learn the distribution, and how well they do with their choice.

Such ecological invalidity might also explain the inconsistency of the effectiveness of my "mean" manipulation in Study 2: In Study 1, the "low mean" condition may not have been considered objectively "bad," while in Study 2, the within-subjects manipulation may have led participants to view the "low mean" conditions as objectively "bad" when they had seen the "high mean" condition previously (due to random ordering of conditions within subjects, 82% of subjects would have seen one "low mean" condition after at least one "high mean" condition. While two types of distributions, normal and uniform, were examined in this paper, future research should consider additional distributions or examine the frequency of distributions in delayed-choice paradigms in everyday life, as well as distributions for which subjective quality bears a nonlinear relationship to objective quality (e.g., money; Kahneman & Tversky, 1979).

An additional limitation to this work has to do with the discussion of "learning" a distribution, or delaying "in order to" learn about a distribution of options. There is consistent evidence that people do indeed delay—Hypothesis 1 shows just this. Further, there is consistent evidence that delaying leads to knowledge of the distribution—Hypothesis 1d (that quality is a stronger predictor of choice when participants have delayed) and Hypothesis 2a (that subjective quality grows closer to the actual objective quality of an option as participants view more options) support this. Hypotheses 1c and 2e also provide concurrent evidence, as they show that delay behavior is based on the distribution (or slope) of options observed. However, using the term "learning" throughout this paper may suggest an intentionality that has not been directly tested. I have not shown evidence that distribution learning is an active or conscious determinant of delay, nor is there evidence that participants know the mean, variance, skew, or slope of the options they have seen. Clearly, some sort of learning does occur, as shown by the results discussed above, but the conscious, intentional processes by which the learning itself occurs, or why people believe themselves to be delaying, has not been addressed.

Future Work

As I have alluded to above, there are several ways to extend this work, involving examination of additional distributions, additional choice frames, additional modifications of the optimal stopping paradigm, and additional decision scenarios. First, prior knowledge of a distribution should greatly affect the processes by which people seek additional information. For example, when attempting to purchase an airplane ticket, most people hold knowledge of how much tickets cost, as a function of the many attributes (coach vs. first class, length of flight, number of stops, etc.) before they even begin seeking options. One interesting extension of this work would be to examine how these choices unfold from a distribution-learning perspective. Memory and recall of distributional information may take several forms, including recall of distribution-level qualities (such as the mean and variance), best-case and worst-case scenarios (i.e., the range of the distribution, akin to Karlsson, Juslin, and Olsson's 2008 discussion of exemplars), or some other sentinel value.

Another extension of the current work could address the rate at which prior distribution information goes stale: Even though distribution information may be correctly recalled, some time may be taken to re-learn or re-confirm the distributional qualities of a similar option set. Indeed, the question of how people apply distribution information gained from one option set to others is an interesting future direction: for example, translating a known distribution of airline prices from Eugene to San Francisco to flights from Boston to Chicago. While the mean price will certainly differ, people may presume that the variation around the mean would be similar, empowering them to make decisions quicker. Further, are those who delay more initially able to make faster decisions (delaying less) in a second or far-off decision because the information initially gained remains salient? When multiple cues are present, do people learn their correlations as well, building up cues that allow them to evaluate less information in the future such as "direct flights are more expensive?"

Altering the optimal stopping paradigm used is also an important direction for future research. As mentioned above, the current paradigm fails to capture how decision delay functions when several options are simultaneously available. A future study could address this by adding and/or removing several options simultaneously in each step. For example, in Study 3, this could be accomplished by suggesting that the participant should imagine interviewing several roommates today, after being shown each potential roommate's personability and financial stability, and then note that some other roommates (interviewed previously) had found places to live and withdrawn their application. This would lead to several roommates being available at each decision point, such that delaying choice would not necessarily (but instead, probabilistically) eliminate any current option. Effectively, this would be a modification of Patalano and Wengrovitz' (2007) paradigm, altered to use a variable distribution of qualities and likelihoods of options becoming unavailable. This would allow the testing of several additional hypotheses: For example, if delay did not increase in such a study (despite the fact that good options would stick around), that would provide additional evidence that the observed distribution is indeed being learned or somehow internalized.

Extending these methodologies to paradigms other than optimal stopping problems or variations on them may also lead to greater knowledge about decision delay and decision making in general. For example, when researchers have participants choose one out of many options, they could consider qualities of the set of options as well as pairwise comparisons. For example, researchers using Dhar's (1997) paradigm could consider the variance and skew of his set of two superior and one inferior options, and those using Botti and Iyengar's (1994) paradigm could consider the variance in quality of the jams they have participants sample. It may well be that Dhar's conclusion of a preference reversal (for superior options versus waiting when another superior option is added) may be better explained as an example of more variable options sets engendering delay. Similarly, Botti and Iyengar's conclusion that it is aversive to have "too much choice" may also be better-explained by the theory that less-variable option sets lead to less satisfaction and/or a more aversive decision process. Broadly speaking, this research should lead researchers to question whether an assumption of "regularity," as Corbin (1980) puts it, is ever appropriate: How exactly should a comparison of option A to option B should be interpreted when option C is added to (or removed from) the choice? Expecting this added context to be ignored in the name of "rationality" may not, in fact, be rational.

Procrastination research can also be informed by the methods and results of this paper. As discussed in "Implications," above, there is now a strong case to be made for examining delay of action independently from an expectation (on the part of researchers or participants) that delay will lead to worse outcomes. Further, even researchers of the willful off-putting of actions by people who expect to suffer for their delay may be able to fruitfully use the methods found in the current studies. Modeling a linear relationship between delay and option quality in these studies did not tell the whole story: A positive linear effect with a negative quadratic effect showed that some delay was beneficial, although this benefit tapered off (and eventually became negative). This initial positive effect of delay may be important in explaining the relationship between procrastination and stress: Tice and Baumeister (1997) showed that procrastinators suffered less stress while delaying action (and less overall), but much more stress when the deadlines hit. Together, this suggests that the optimal strategy from a decision-making and health perspective is not the "immediate action" as many procrastination researchers suggest, but rather, an amount of delay appropriate to the decision context.

Summary

I have provided evidence that people both desire distributional information when making a decision among options (and will delay decisions while gaining this information), and those who have such viewed such information learn it and use it—and in some cases, obtain better outcomes. Though I failed to show any individual differences in how these delay processes work, the results discussed do have some bearing on the research and practice of human decision-making, bridging conflicting models of procrastination, at least in the realm of decision-making: Haste may indeed make waste, but he who hesitates (too long) may also be lost.

REFERENCES

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256–274.

Anderson, C. J. (2003). The psychology of doing nothing: Forms of decision avoidance result from reason and emotion. *Psychological Bulletin, 129*, 139–167.

Atkins, D. C., & Gallup, R. J. (2007). Rethinking how family researchers model infrequent outcomes: A tutorial on count regression and zero-inflated models. *Journal of Family Psychology, 21,* 726–735.

Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making, 19,* 87–98.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science, 13*, 219–224.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology, 75,* 19–32.

Baayen, R. H. (2010). languageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics." (R package version 1.0) [Computer software]. Available from http://CRAN.R-project.org/package=languageR

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Bates, D., & Maechler, M. (2010). lme4: Linear mixed-effects models using S4 classes (R package version 0.999375-33) [Computer software]. Available from http://CRAN.R-project.org/package=lme4

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 75*, 1252–1265.

Botti, S., & Iyengar, S. S. (2004). The psychological pleasure and pain of choosing: When people prefer choosing at the cost of subsequent outcome satisfaction. *Journal of Personality and Social Psychology, 87*, 312–326.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology, 67,* 366–381.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy, 94,* 1232–1254.

Clore, G. L., Gasper, K., & Garvin, E. (2001). Affect as information. In J. P. Forgas, (Ed.). *Handbook of Affect and Social Cognition* (pp. 121–144). Mahwah, NJ: Lawrence Erlbaum Associates.

Corbin, R. M. (1980). Decisions that might not get made. In T. Wallsten (Ed.). *Cognitive Processes in Choice and Decision Behavior* (pp. 47–67). Hillsdale, NJ: Lawrence Erlbaum Associates.

Corbin, R. M., Olson, C. L., & Abbondanza, M. (1975). Context effects in optimal stopping decisions. *Organizational Behavior and Human Decision Processes, 14,* 207–216.

Dawes, R. M. (1975). Graduate admission variables and future success. *Science, 187,* 721–723.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582.

Dawes, R. M. (1990). The potential nonfalsity of the false consensus effect. In R. M. Hogarth (Ed), *Insights in decision-making: A tribute to Hillel J. Einhorn,* (pp. 179–199). Chicago, IL: University of Chicago Press.

Dawes, R. M. (1994). House of cards: Psychology and psychotherapy based on myth. New York: Simon & Schuster.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research, 24,* 215–231.

Diab, D. L., Gillespie, M., & Highhouse, S. (2008). Are maximizers really unhappy? The measurement of maximizing tendency. *Judgment and Decision Making, 3,* 364–370.

Diefendorff, J. M., Hall, R. J., Lord, R. J., & Strean, M. L. (2000). Action-state orientation: Construct validity of a revised measure and its relationship to work-related variables. *Journal of Applied Psychology, 85,* 250–263.

Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology, 61,* 467–490.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311,* 1005–1007.

Duncan, S. M., Wengrovitz, S. M., Sedlovskaya, A., & Patalano, A. L. (2007). Weighing waiting: The influence of information certainty and delay penalty on waiting for noninstrumental information. *Judgment and Decision Making, 2,* 351–358.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2,* 312–329.

Effert, B. R., & Ferrari, J. R. (1989). Decisional procrastination: Examining personality correlates. *Journal of Social Behavior and Personality, 4,* 151–156.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science, 4,* 282–296.

Ferrari, J. R., & Dovidio, J. F. (2000). Examining behavioral processes in indecision: Decisional procrastination and decision-making style. *Journal of Research in Personality, 34,* 127–137.

Finucane, M. L., Peters, E., & Slovic, P. (2003). Judgment and decision-making: The dance of affect and reason. In Schneider, S. L., & Shanteau, J. (Eds.), *Emerging perspectives on judgment and decision research* (pp. 327–364). New York: Cambridge University Press.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288–299.

Forgas, J. P., & Bower, G. H., (1987). Mood effects on person-perception judgments. *Journal of Personality and Social Psychology, 53,* 53–60.

Fried, L. S., & Peterson, C. R. (1969). Information seeking: Optimal versus fixed stopping. *Journal of Experimental Psychology, 80,* 525–529.

Frost, R. O., & Shows, D. L. (1993). The nature and measurement of compulsive indecisiveness. *Behaviour Research & Therapy, 31*, 683–692.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17,* 295–314.

Goldberg, L., & Werts, C. E. (1966). The reliability of clinicians' judgments: A multitrait-multimethod approach. *Journal of Consulting Psychology, 30,* 199–206.

Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. L. Morgan (Eds), *Syntax and Semantics Volume 3: Speech Acts.* (pp. 41–58). New York: Academic Press.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2,* 293–323.

Harriott, J., & Ferrari, J. R. (1996). Prevalence of procrastination among samples of adults. *Psychological Reports, 78,* 611–616.

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world.* Thousand Oaks, CA: Sage.

Hodges, S. D. (1997). When matching up features messes up decisions: The role of feature matching in successive choices. *Journal of Personality and Social Psychology, 72,* 1310–1321.

Howell, D. C. (2008). *Statistical methods for psychology.* Belmont, CA: Wadsworth Publishing.

Howell, W. C. (1966). Task characteristics in sequential decision behavior. *Journal of Experimental Psychology, 71,* 124–131.

Hsee, C. K., & Abelson, R. P. (1991). Velocity relation: Satisfaction as a function of the first derivative of outcome over time. *Journal of Personality and Social Psychology, 60,* 341–347.

Hsee, C. K., Salovey, P., & Abelson, R. P. The quasi-acceleration relation: Satisfaction as a function of the change of velocity of outcome over time. *Journal of Experimental Social Psychology, 30,* 96–111.

Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York: Free Press.

Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple-cue judgment: A division of labor hypothesis. *Cognition, 106,* 259–298.

Juslin, P., Olsson, H., & Olsson, A-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General, 132,* 133–156.

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

Karlsson, L., Juslin, P., & Olsson, H. (2008). Exemplar-based inference in multi-attribute decision-making: Contingent, not automatic, strategy shifts? *Judgment and Decision Making, 3,* 244–260.

Kramer, A. D. I., & Hodges, S. D. (unpublished manuscript). Everyday decision-making: Decision perception across contexts.

Kuhl, J. (1981). Motivational and functional helplessness: The moderating effect of state versus action orientation. *Journal of Personality and Social Psychology, 40,* 155–170.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21,* 37–44.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology, 81,* 146–159.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science, 15,* 337–341.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65,* 272–292.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, H. Goldsmith, & K. Scherer (Eds.), *Handbook of Affective Science* (pp. 619–642). New York: Oxford University Press.

Mann, L., Burnett, P., Radford, M., & Ford, S. (1997). The Melbourne decision-making questionnaire: An instrument for measuring patterns for coping with decisional conflict. *Journal of Behavioral Decision Making, 10,* 1–19.

Mehle, T., & Gettys, C. F. (1979). Optimal stopping strategies for information purchase in multiple-stage inference tasks. *Organizational Behavior and Human Performance, 24,* 333–353.

Milgram, N. A., Sorloff, B., & Rosenbaum, M. (1988). The procrastination of everyday life. *Journal of Research in Personality, 22,* 197–212.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126,* 247–259.

Patalano, A. L., & Wengrovitz, S. M. (2007). Indecisiveness and response to risk in deciding when to decide. *Journal of Behavioral Decision Making, 20,* 405–424.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17,* 407–413.

R Development Core Team (2010). R: A language and environment for statistical computing. (versions 2.9.1 through 2.11.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.r-project.org

Rapoport, A., & Tversky, A. (1966). Cost and accessibility of offers as determinants of optional stopping. *Psychonomic Science, 4,* 145–146.

Rapoport, A., & Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance, 5,* 105–120.

Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). "I know, you know": Epistemic egocentrism in children and adults. *Review of General Psychology, 7,* 38–65.

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,* 1178–1197.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of personality and Social Psychology, 45,* 513–523.

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "secretary problem." *Organizational Behavior and Human Decision Processes, 69,* 221–236.

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making, 13,* 391–411.

Shaklee, A. B. (1958). Optional stopping and the theory maker. *Psychological Reports, 4,* 17.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*, 11–36.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.

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

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133,* 65–94.

Steel, P., Brothen, T., & Wambach, C. (2001). Procrastination and personality, performance, and mood. *Personality and Individual Differences, 30,* 95–106.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1–26.

Tice, D. M., & Baumeister, R. F. (1997). Longitudinal study of procrastination, performance, stress, and health: The costs and benefits of dawdling. *Psychological Science, 8,* 454–458.

Timmermans, D. (1993). The impact of task complexity on information use in multi-attribute decision making. *Journal of Behavioral Decision Making, 6*, 95–111.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–209.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1130.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3,* 305–309.

Tykocinski, O. E., & Ruffle, B. J. (2003). Reasonable reasons for waiting. *Journal of Behavioral Decision-Making, 16,* 147–157.

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin, 29,* 1159–1168.

Van Putten, M., Zeelenberg, M., & van Dijk, E. (2009). Dealing with missed opportunities: Action vs. state orientation moderates inaction inertia. *Journal of Experimental Social Psychology, 45,* 808–815.

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making choices impairs subsequent self-control: A limited-resource account of decision-making, self-regulation, and active initiative. *Journal of Personality and Social Psychology, 94,* 883–898.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin, 77,* 65–72.

Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology, 58,* 221–228.

Yates, J. F., Veinott, E. S., & Patalano, A. L. (2003). Hard decisions, bad decisions: On decision quality and decision aiding. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 13–63). New York: Cambridge University Press.