May 2020

Preference Construction in Difficult Multi-Attribute Health Decisions: Effects of Affect and Coherence Shifting

Michael Shreeves
Clemson University, michael.shreeves@asu.edu

PREFERENCE CONSTRUCTION IN DIFFICULT MULTI-ATTRIBUTE HEALTH DECISIONS: EFFECTS OF AFFECT AND COHERENCE SHIFTING

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Human Factors Psychology

---

by
Michael Shreeves
May 2020

Accepted by:
Dr. Leo J. Gugerty, Committee Chair
Dr. Kaileigh Byrne
Dr. James McCubbin
Dr. Fred Switzer

ABSTRACT

In shared decision-making, doctors provide patients with information about difficult trade-off treatment decisions so the patient can make an informed choice. Many models of decision-making assume that patients make decisions based on long-term, stable preferences, but research suggests that people dynamically construct preferences for each decision. Affect plays at least two roles in preference construction. First, coherence shifting, or altering preferences prior to choice to make one alternative more attractive, may regulate emotion. Difficult decisions, imagining unpleasant outcomes, and threats to closely held goals produce general negative affect, and coherence shifting may reduce this. Second, preferences for alternatives may be constructed from immediate affective reactions, driving choice.

Two dichotomous trade-off health decision scenarios were produced that are highly conflicted on outcome unpleasantness. Experiment 1 compared a serious disease trade-off decision with a job selection task used in prior research on preference construction. Experiment 2 compared decision-making between serious and mild disease treatment decisions differing in outcome severity, also including a physiological affect measure. In both experiments, choice was best predicted by a model including only affect towards alternatives within a decision context. Prediction was not improved by including outcome and attribute ratings independent of decision context, providing support for preference construction over revealed preferences. Coherence shifting of outcome affect and attribute importance ratings was fully or partially supported in all four tasks. Tasks

with more severe outcomes or threatening higher-level goals (e.g., survival) produced more aversive feelings but did not lead to stronger coherence shifting.

Keywords: affect heuristic, preference construction, decision-making, health

DEDICATION

This dissertation is dedicated to Dr. Sandra Carpenter, who taught me the meaning of ambition; To my fellow graduate students for making my time at Clemson some of the best years of my life; and to Dr. DeWayne Moore for telling me to "Stop thinking like an engineer." I'm trying.

ACKNOWLEDGMENTS

This dissertation could not have been completed without the continuing and substantial support of a community of colleagues, experts, and friends. First, thanks to Clemson University's Graduate School, the Office of the Vice President for Research, the Creative Inquiry and Undergraduate Research Program, and the Human Factors Institute for providing critical funding and equipment in support of this research program.

Thanks to my advisor, Dr. Lee Gugerty. I am profoundly grateful for your mentorship, for never failing to challenge me to do everything in the right way and for the right reasons, for believing in me as a researcher, and for giving me the opportunity to pursue my academic career here at Clemson. Thanks to my committee members, Dr. Kaileigh Byrne, Dr. James McCubbin, and Dr. Fred Switzer, for their extensive discussions and effort in this project and for your mentorship in difficult new skills. Your efforts have made this dissertation not only possible, but something I am proud of.

Thanks to my undergraduate assistants - Brianna Carmona, Mary Grubbs, Olivia Siciliano, and Paige Lawton - for their enthusiasm and their substantial involvement in the design and implementation of this research. I wish you only the best.

Finally, thanks to Dr. Eric Muth, Dr. Patrick Rosopa, Dr. Sheri Webster, and Linda Gugerty for their expert advice and assistance with the design of this experiment. Thanks to Dr. Dan Simon for his kindness and for providing his experimental stimuli. Thanks also to Nathan Dumessa for volunteering his efforts when I was in need and to all of our Creative Inquiry students (including but not limited to Shane Donahue, Keri Fallaw, Sydney Slattery, Spencer Tabberer, and Delaney Wallace) for their hard work.

TABLE OF CONTENTS

Table of Contents (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

List of Tables (Continued)

Table Page

List of Tables (Continued)

Table Page

List of Tables (Continued)

Table Page

INTRODUCTION

Shared decision making is an increasingly prominent paradigm in health care, including academic and medical research (Makoul & Clayman, 2006) and practical integration in large-scale government health programs (Elwyn et al., 2010). Elwyn et al. define shared decision making as "an approach where clinicians and patients share the best available evidence when faced with the task of making decisions, and where patients are supported to consider options, to achieve informed preferences" (p. 971). Communication tools known as decision aids have been developed for a variety of health contexts to aid clinicians and patients in shared decision-making scenarios (Stacey et al., 2011). Health decisions that patients may face in shared decision making, such as choosing to modify a treatment for cancer, often involve weighing costs and benefits on a variety of dimensions such as out-of-pocket cost, toxicity (side effects), and efficacy (e.g., Wong et al., 2013; Stacey et al., 2011). These types of multi-attribute decisions are commonly studied in psychology, and there are many well-documented strategies and processes for choosing between alternatives. These often involve subjective appraisals of the attributes on which alternatives can be compared and the specific outcomes of the alternatives. Elwyn et al. (2012), for example, suggest that clinicians should ask patients what (attribute) is most important to them in order to guide them to a preference.

There is no guarantee, however, that a patient's subjective appraisals will be relevant towards their long-term goals or that they will base their decision on the maximum amount of evidence presented to them. This may be the case even when decision aids use design elements that reduce the need for numerical or health literacy,

such as evaluative categories (Peters et al., 2009). To provide information in a manner that is useful and beneficial to patients, we must understand how participants arrive at preferences and how they use these preferences to reach a decision. Within the large field of decision-making, this dissertation will focus on the issue of revealed versus constructed preferences. Namely, do decisions reveal underlying preferences, or are preferences constructed in the context of that decision?

Several authors have noted the important role of emotions in decision-making. In a review of health decision-making literature, Carpenter and Niedenthal (2017) provided a number of ways in which decision-making generates aversive feelings and how such feelings can guide or divert the decision-making process, and further proposed that the processes of generating preferences and controlling emotion are interlinked. In particular, they emphasize findings by Carpenter, Yates, Preston and Chen (2016) suggesting that difficult decisions produce aversive feelings and that the process of altering preferences across the course of decision-making to support a single alternative (e.g., Simon, Krawczyk, & Holyoak, 2004) serves to regulate these emotions. Alternatively, Slovic, Finucane, Peters, and MacGregor (2004) and Bechara and Damasio (2005) suggest that emotions themselves are often the method by which a decision is reached, and several later studies suggest that choice is best predicted by emotion (Charpentier, De Neve, Li, Roiser, & Sharot, 2016; Schlösser, Dunning, & Fetchenhauer, 2013). It is also likely that health decisions will produce stronger negative emotions because they involve imagining visceral physical consequences that threaten higher-level goals like survival (Loewenstein, Weber, Hsee, & Welch, 2001; Luce, 1998).

The goal of this research project is to examine the role of emotions in a complex, multi-attribute health decision trade-off. In particular, this project was focused on the following questions. Will people show higher aversive feelings and more negative affect towards decision aspects (e.g., outcomes and alternatives) when facing a more serious health-based decision? Does affect towards aspects of a decision predict choice of a medical treatment better than other models of decision making, including more cognitively costly ones? Will people faced with a health decision engage in the shifting of preferences to support their final decision, and will this shifting reduce aversive feelings arising from a difficult decision? Finally, does the decision context impact this shifting?

In this dissertation, I explored these concepts using three difficult trade-off decision tasks in two experiments, manipulating the emotional salience and context of the tasks. In the first experiment, affect was measured across the course of decision making in a previously developed job-selection task and a newly developed shared-treatment decision scenario for a serious disease. The second experiment compared two novel shared-treatment decision scenarios of varying disease severity (serious and mild) with the addition of an objective physiological measure of affect.

The Construction of Preferences

The goal of presenting patients with as much evidence-based information as possible is to allow them to make an informed, rational decision about their health (Elwyn et al., 2012). One of the most comprehensive and cognitively difficult methods of decision-making, which is designed to use all possible evidence, is the weighted additive strategy. Under this process, patients would rate the utility of all possible outcomes, rate

the importance of the attributes under comparison, and linearly multiply importance by utility before summing scores for each alternative. The alternative with the highest score would be selected. Simplified, qualitative versions of this method exist that account for the lack of ability of non-experts to give meaningful utility or importance ratings (Hastie & Dawes, 2010). Weighted additive effectively uses all information in difficult, conflicted decisions (where there is not one alternative that is superior on all outcomes) because positive outcomes on one or more attributes can compensate for negative outcomes on other attributes.

Other, less cognitively intensive processes focus only on a single attribute. In lexicographic decision making, patients select the alternative with the best outcome on the most important attribute, such as picking the treatment with the highest survivability (Hastie & Dawes, 2010). Patients the using elimination by aspects strategy will reject any alternative with unacceptable outcomes on the most important alternative (e.g., rejecting all treatments with side effects that are too severe; Tversky, 1972). These two methods align with instructions for patients in shared decision making to think about what is most important to them, although a well-informed patient would have enough information to use a strategy like weighted additive (Elwyn et al., 2012).

The ability of weighted additive and the other processes specifically mentioned above to consistently produce the maximum possible *utility* (or personal value) across many decisions, however, relies on the basis that people's preferences about outcomes and attributes are invariant across both decision context and time. A preference in one decision must represent some stable, underlying or revealed preference (von Neumann &

Morgenstern, 1947). Research in the field of preference construction has brought this assumption into doubt. In one paradigm, Simon and colleagues (Simon, Pham, Le, & Holyoak, 2001; Simon et al., 2004) conducted a series of studies investigating dynamic preference construction, i.e., how people change underlying preferences until one alternative can meet their needs. Under this model, rather than accessing stable and invariant preferences and using them for a decision process in a conflicted decision, people alter (i.e., construct) their preferences prior to making a choice in order to make one alternative the most attractive. For an extreme judgement-based example, members of a jury given ambiguous evidence may alter the perceived strength of evidence for or against a crime based on a simple first impression of the defendant, continuing until the evidence for that initial leaning appears overwhelming and the evidence against it seems negligible (Simon, 2004). Simon et al. (2004) explain this process using connectionist constraint-satisfaction networks, suggesting that attributes and outcomes that are similar (i.e., are attractive for the same alternative) will receive increasing activation while outcomes supporting other alternatives are inhibited. This leads to altered preferences that are uniformly high for one alternative and uniformly low towards the others, thus strongly supporting a single alternative. Simon et al. refer to this process of changing preferences aligning over time in order to support a single alternative as *coherence shifting*.

Simon et al. (2004) examined coherence shifting in difficult multi-attribute decisions using a job offer scenario, measuring preferences for alternatives and outcomes (which are the inputs to the weighted additive model) at three separate times in one

decision. Participants were first shown attributes (e.g., salary, commute) and specific outcomes (e.g., salary $600 below industry average, 18 minute commute) related to job offers but outside of the context of any specific job (pre-choice). Participants rated outcomes in terms of desirability and attributes in terms of importance. After a distractor task, participants were presented with two hypothetical job offers using four attributes and eight outcomes they had previously rated. This job decision was conflicted and represented a difficult trade-off, with each job superior on two attributes and inferior on the other two. Differences balanced to provide a similar overall utility, but a fifth attribute was manipulated to give one job or the other a small advantage. Participants were told to delay their decision due to a possible job offer retraction, but rated outcomes and alternatives a second time (mid-choice). Participants were then allowed to continue with their decision and provided final (post-choice) ratings. The authors found that participants significantly shifted their importance weights and outcome ratings between the pre-choice and mid-choice measures such that they were more positive towards the winning traits of the offer they selected and more negative towards traits favoring the offer they had rejected. This suggests that preferences related to attributes and outcomes can change during the course of making a single decision and that the direction of these choices depends on the final choice, which Simon et al. refer to as coherence shifting. (These shifts are not permanent. Simon & Spiller (2016) found preferences returned to per-choice levels within 6 weeks. Simon, Krawczyk, Bleicher, and Holyoak, (2008) found that preferences returned within one week, and perhaps as quickly as 15 minutes.)

Emotions from Decision-Making

Carpenter et al. (2016) believed making difficult, conflicted, multi-attribute decisions produces negative emotions, and that coherence shifting might serve as a process to regulate that emotion. The authors performed three conceptual replications of Simon et al. (2004). First, they found that increasing the level of conflict in a difficult multi-attribute decision (e.g., higher distance between outcomes on attributes, such as a larger difference in salaries) caused increasing aversive, negative feelings in participants. Second, they replicated the job offer task and three measurement times, but also included self-report measures of decision difficulty and a measure of physiological arousal (skin conductance response). Stronger shifting in preferences was associated with higher self-reported ease of decision-making, and those who strongly shifted preferences were the only ones to show a reduction in physiological arousal between the mid- and post-choice measures. In the third experiment, a manipulation that reduced available cognitive resources resulted in less coherence shifting. This aligns with findings that emotional regulation is a cognitively intensive process that leads to depletion of cognitive control resources (Hobson, Saunders, Al-Khindi, & Inzlicht, 2014).

Carpenter et al. (2016) took these findings as evidence that negative emotions arise from decisions in proportion to their difficulty (level of conflict between outcomes), and that coherence shifting serves to regulate these negative emotions. Coherence shifting reduces the perceived difficulty of the decision by making one option more attractive. This results in less internal conflict and decision difficulty than deciding based on stable, highly conflicted preferences from the pre-decision time. General negative emotions are

important in a health context because they can lead to biases or deficiencies in decision making. Kuykendall and Keating (1990) found that inducing a negative mood unrelated to a critical reasoning task (i.e., reading an unrelated unpleasant article) led to more systematic thinking, but in other studies negative affect was associated with reduced performance in perceptual decision tasks, a tendency to delay decisions, or (most importantly) a preference for less risky or threatening alternatives (Byrne, Peters, & Willis, 2018; Lerner & Kentner, 2001; Loewenstein et al., 2001; Luce, 1998).

Importantly, Carpenter et al. (2016) did not address the context or specific content of the decision as a source of affect, merely the level of conflict. Taken alone, their theory would suggest similar levels of aversive emotions in an emotionally neutral shopping task or a serious health decision, as long as the level of conflict between alternatives (difficulty) is similar. Carpenter and Niedenthal (2017) later addressed this conflict as the primary source of decision-related emotion in their review on health decision-making, although they did suggest that emotional regulation from coherence shifting can also reduce long-term negative emotions that are common in patients facing health decisions.

However, high conflict between outcomes is far from the only source of negative emotion in the health decision process. Luce (1998) defined decision "difficulty" as the extent to which the decision threatens higher-level goals, rather than differences in outcomes. In a consumer decision-making study, Luce presented participants with one of two numerically identical car buying tasks, changing the names of two attributes. These attributes had been pre-rated as being equally important to buying a car (e.g., Routine

handling vs. Occupant (crash) survival), but attributes in one condition threatened higher-level goals like safety or esteem. Luce found that decisions produced more general negative affect when attributes threatened higher-level goals, even if the attributes in question were rated of equal importance to the task. I am continuing Carpenter et al.'s research into coherence shifting and regulation of aversive feelings, but I expanded this research in Experiment 1 by using two decision scenarios that vary in the levels of goals threatened for participants (e.g., comfort in the job task vs. survival in the disease task).

In addition to threatening important goals, descriptions of health decisions often involve physical symptoms or side effects that are viscerally imaginable and produce an immediate physiological response, both of which can have direct impact on negative emotion (Carpenter and Niedenthal, 2017; Loewenstein et al, 2001). This is related to Damasio's somatic marker hypothesis, which states that emotions in decision-making arise from bioregulatory neural processes that either come from or mimic physiological states in the body (Bechara & Damasio, 2005). The somatic marker hypothesis, however, goes beyond theories of how difficult decisions produce general negative emotions, suggesting that negative and positive affect are also mechanisms driving choice. In Experiment 2, I expanded on research by Carpenter et al. (2016) by using two tasks that both threaten higher-level goals but vary in the unpleasantness of their outcomes (i.e., moderately unpleasant in mild disease treatments versus highly unpleasant in serious disease treatments.).

Emotions as Decision-Making

To explain the direct role of emotion in decision-making, Slovic, Finucane, Peters, and MacGregor (2004) proposed a process known as the affect heuristic. This had its basis in early research on public risk perception, which showed that people evaluate public health risks based primarily on emotional dread of consequences and unfamiliarity, rather than quantitative or expert information (Slovic, 1987). Slovic and colleagues proposed that features of an object, location, or situation are automatically compared to affect-based markers in memory, leading to an immediate emotional reaction. These somatic markers are aspects of past experiences that are associated with physiological emotion-states and are used to judge current situations or alternatives (Bechara and Damasio, 2005). In decision-making, this immediate emotional reaction can alter perceptions and evaluations of outcomes or alternatives or even be the only driver of the decision-making process (e.g., avoiding alternatives evoking negative feelings and seeking those evoking positive feelings). The affect heuristic has the benefit of being an experience-based and low-effort process, but the disadvantage of discounting quantitative and non-emotional evidence. Also, it does not function well if relevant experience is not available (Slovic et al., 2004). Even if a more complex decision-making strategy is used, short-term affect that arises from imagining future outcomes is often what participants weigh when making decisions (Lowenstein et al., 2001). Peters (2006) suggests that even in cognitively intensive decision making, people may be unable to place value or utility on outcomes or meaningfully compare outcomes on different attributes if those outcomes are not affectively evaluable, with affective reactions operating in place of utility ratings.

These affective reactions may not be stable over time or decisions, being constructed from the combination of somatic markers that happen to be activated by the specific presentation of the alternative as a whole rather than being a simple additive combination of affect towards those outcomes (Bechara & Damasio; Slovic et al.). These affective responses are fast, temporary, and *situation-specific* and can be influenced by factors such as the particular alternatives presented, the order of information perceived, and decision-irrelevant affect or arousal (Kuykendall & Keating, 1990; Loewenstein et al.; Slovic et al.) . Thus, the affect heuristic may be another type of preference construction. (See the discussion of stochastic models below.)

Some evidence of the direct predictive value of affect on decision-making has been found in the related field of probabilistic or risky decision-making. Charpentier et al. (2016) asked participants to rate their expected happiness or unhappiness upon hypothetically winning or losing various amounts of money, before exposing them to approximately 300 trials of a simulated decision between a sure option and a risky 50-50 gamble using the monetary amounts rated. Participant's choice of sure bets or gambles was best predicted by a model containing only emotions towards the monetary values in the outcomes. Adding additional terms to the model for probability of outcomes and framing effects (i.e. different weighting of losses versus gains, Kahneman & Tversky, 1979) did not improve this prediction. Note that Simon et al. (2004) asked participants to rate outcomes in the job offer task on a single-item scale of desirability, which is similar to this predicted happiness rating or a simple positive-negative (valence) affect scale (Lowenstein et al., 2001; Peters, 2006).

Schlösser et al. (2013) conducted a series of similar studies involving real monetary rewards to evaluate the predictive ability of both overall affect towards alternatives (similar to the affect heuristic) and affect towards specific outcomes. Their studies replicated well-known gambling tasks and were arranged such that the expected utility of each alternative was identical, similar to a conflicted trade-off multi-attribute decision. Participants were asked to rate their feelings toward each future outcome of a choice (won, lost, would've won, or would've lost) as well as overall alternative-related emotion (sure bet and gamble) using a more complex multi-dimensional measure of emotion (i.e., pleasure, arousal, and dominance) before selecting an alternative. They found that outcome-related and alternative-related affect each individually improved prediction of choice to a significant extent, but the model with only alternative-based emotions was the most predictive. Alternative-based emotions mediated the relationship between outcome-based emotions and choice, but also had independent effects beyond mediation.

This finding is conceptually similar to stochastic models of multi-attribute choice such the diffusion decision model (Ratcliff & McKoon, 2008) and the leaky competing accumulator model (Usher & McClelland, 2001). According to these models, as people randomly sample attributes when comparing two alternatives, positive affect towards the best outcome on each attribute will increase activation towards one alternative and inhibit activation towards the other alternative. A decision is made when one alternative reaches some threshold of sufficient activation, whether or not all outcomes have been considered. The diffusion decision model is capable of making specific predictions about

decision time that have been supported in repeated trials with simple choices (Ratcliff & McKoon).

However, the diffusion decision and leaky competing accumulator models only describe one direction of activation flow, where *outcomes* provide activation and lateral inhibition *towards alternatives*. Examining activation in the other direction (alternatives to outcomes/attributes) as well as lateral connections between alternatives and outcomes could explain positive and negative changes in preferences during coherence shifting. If we also treat outcomes and attributes as nodes with changing activation, then any activation towards an alternative may also increase activation towards outcomes and attributes that support that alternative (leading to higher importance and utility ratings) and laterally inhibit outcomes and attributes that do not (leading to lower importance and utility ratings). Simon et al. (2004) used a constraint satisfaction network to model preference shifting. In these networks, there are bidirectional connections between outcomes and alternatives as well as lateral connections between outcomes and attributes (Hunt, 2002). There is some evidence for this direction of activation in an analogous public risk perception context. Alhakami and Slovic (1994) found that positive or negative affect towards a wide variety of technologies (e.g., nuclear power) affected participants' subjective ratings of the risks and benefits (outcomes) of adopting those technologies.

In two experiments, I examined both directions of influence by adapting past methods used to study coherence shifting (*preference for an alternative* towards *outcomes and attributes*; Simon, 2004) and those used to predict choice (*outcomes,*

attributes, and affect towards *chosen alternative*; Charpentier et al., 2016; Schlösser et al., 2013). Using both types of analyses allowed me to fill an important gap in both research programs. In addition, using a predictive analysis allowed me to test multiple decision strategies in order to identify what decision information best predicts choice in simulated shared treatment decisions.

Summary of Decision Making

In summary, there are multiple models of decision making that could be used to explain how patients handle difficult trade-offs in a shared decision context. Weighted additive is thought to generally lead to decisions that maximize utility (von Neumann & Morgenstern, 1947) and deals with difficult trade-offs by expending considerable cognitive effort to consider preferences towards all available information. Attribute-based models (lexicographic, elimination by aspects) simplify the decision by discarding all information not related to one or a few critical attributes. These specific models reflect what experts instruct patients to do in shared decision-making. Weighted additive and attribute-based models rely on the idea that as an alternative becomes more preferred during the decision process, the increased preference for the alternative does not feed back to influence preferences for outcomes or attributes. This means that patients' preferences are stable over time and they will make decisions that optimally match their stable preferences in the long term. It is important to note that shared decision making advocates are not overly concerned with the assumption of stable preferences (Elwyn et al., 2012).

In contrast, emotion or affect-based models of decision making assume that preferences for outcomes, attributes, and alternatives are constructed in the context of a single decision and can change during the decision process. The affect heuristic suggests that choice is influenced by a rapid affective judgement of each alternative as a whole based on the particular combination of somatic markers activated by the outcomes of that alternative, but only within the specific framing of that decision and at that particular time (Slovic et al.2004; Kuykendall & Keating, 1990), . This construction of preferences is in the forward direction, with an affect-based preference for one alternative or another being constructed from past experience of outcomes, although not through a purely mathematical combination of stable preferences. This theory has some support from empirical findings (Charpentier et al., 2016; Schlösser et al., 2013) and a basis in choice construction models such as the diffusion decision model and leaky competing accumulator theory (Ratcliff & McKoon, 2008; Usher & McClelland, 2001), This system is low-effort and experience-based, and although prone to bias, often leads to effective decisions (Bechara & Damasio, 2005; Slovic et al., 2004). General (i.e., related to the decision as a whole) negative affect can arise from the level of conflict in a difficult trade-off (Carpenter et al., 2016), threats to closely held goals like survival (Luce, 1998), or viscerally unpleasant decision outcomes themselves (Slovic et al.). These aversive feelings can alter or disrupt the decision making process (Loewenstein et al., 2001; Luce, 1998).

Construction of preferences can also flow in the other direction. Under coherence shifting, after coming to some initial preference for one alternative, patients would

change their subjective feelings towards outcomes and the importance of attributes such that the initially preferred treatment becomes more attractive and other alternatives become less attractive (Simon et al., 2004). Patients who display high coherence shifting reduce some of the negative affect or arousal arising from a difficult trade-off decision. Patients who do less coherence shifting do not reduce negative affect and report higher subjective effort required to reach a decision (Carpenter et al., 2016).

Definitions of Affect in the Current Study

Affect is usually defined as a short-term emotional state or mood. Affective reactions refer to specific short-term emotions that are evoked by some stimuli in the environment. Medical treatment decisions may lead to longer-term emotional distress, which can represent more debilitating states (Carpenter & Niedenthal, 2017). Due to the short duration and simulated nature of the current tasks, however, such long-term emotions were not examined.

This leads to an important note about terminology. Throughout this project, separate terms are used to represent affect towards different aspects of a multi-attribute decision: *Affect* towards outcomes, *feelings* towards alternatives, and *aversive feelings* towards the decision as a whole. These are used for the purposes of clarity, to differentiate between the variables being measured. All of these descriptors are still meant to represent affect, or short-term emotional states or reactions. Another concern is the depth of affect measures. Schlösser et al. (2013) measured affect towards all outcomes and alternatives in their study on three dimensions: Positivity, Arousal, and Dominance. These dimensions have been shown in past research to account for a

substantial amount of variance (> 50%) in decision-making and preferences (Mehrabian, 1995). They have also been found through factor analysis to contribute the majority of the variance in other, more specific measures of emotion. Affect positivity, or valence, is a simple positive or negative reaction, and is the kind typically discussed in public risk perception (Lerner & Kentner, 2001) risky decision-making (Charpentier et al., 2016), and activation-based decision models (Ratcliff & McKoon, 2008). The arousal dimension represents a level of activity or alertness and can lead to either avoiding or seeking an action depending on the associated positivity. Many outcomes in health decision are unpleasant, so avoidance is more likely (Bechara, Tranel, Damasio, & Damasio., 1996). Arousal is also the only affect dimension that can be reliably associated with physiological arousal measures, such as skin conductance (Figner & Murphy, 2011). Dominance represents a spectrum of perceived control over a situation versus feeling controlled by external influences (Mehrabian). Dominance is strongly related to health decision-making, as perceived control or self-efficacy is considered essential for adopting and maintaining health-promoting behaviors (Strecher, McEvoy DeVellis, Becker, & Rosenstock, 1986; Weinstein & Sandman, 2002). Outside of these dimensions, Carpenter et al. (2016) included unpleasantness, stress, anxiety, and feeling conflicted to provide a broader measure of aversive feelings.

Ideally, affect towards both alternatives and outcomes would be measured with a multifaceted measure, as was done by Schlösser et al. (2013). Unfortunately, the larger number of outcomes used in these coherence shifting tasks (8 vs. their 4) and the repeated measures necessary to demonstrate coherence shifting (pre-, mid-, and post-choice)

makes this impractical. Charpentier et al. (2016) were able to find a reliable predictive relationship between affect and choice using one-dimensional positivity (happiness) measures. Thus, the following studies will include one-dimensional (positivity) affect ratings for constructs that must be measured repeatedly (outcomes), and multi-dimensional measures of affect for more complex mediating affective responses (feelings towards alternatives) and those that are compared to physiological arousal (aversive feelings towards the decision as a whole).

Purpose and Hypotheses

This research project will examine the role of emotion and direction of preference change in difficult multi-attribute health trade-off decisions. Decisions will involve choosing between two disease treatments that are conflicted, with each treatment (alternative) having outcomes that are superior on some attributes and inferior on other attributes. The most important questions addressed in this research focus on the role of affect in the decision making process. The relationship between affect and choice was examined in two directions. The first is predictive, where affective reactions predict choice:

> Hypothesis 1: In Experiments 1 and 2, choice will best be predicted by a model including affect towards outcomes outside of the decision context (pre-choice) and feelings towards alternatives within the context of a decision (mid-choice) such that lower ratings of negative affect towards outcomes related to an alternative and less negative feelings towards an alternative will increase its likelihood of being chosen. This model will be more predictive of choice than

models where choice is predicted only by affect towards outcomes, only by attribute importance (approximating Lexicographic/Elimination by Aspects), or by a linear combination of outcome affect and attribute importance (Weighted Additive).

According to evidence from a related risky decision paradigm, models using simple positive/negative affect towards outcomes (Charpentier et al., 2016) and multi-dimensional feelings about alternatives (Schlösser et al., 2013) are the most predictive of choice. These models are not improved by including additional rational factors like the probability of outcomes. This method, which is simpler than weighted additive, also aligns with the choice-construction models, where activation and inhibition towards alternatives come from the difference between outcomes rather than a focus on important attributes, e.g., Ratcliff & McKoon's (2008) diffusion decision model.

In contrast to the model above, if models relying on stable, revealed preferences are correct, then attribute importance ratings and affect towards outcomes measured independently of any specific alternatives or decision should be predictive of choice. For example, if affect is the mechanism underlying utility ratings, as suggested by Peters (2006), the weighted additive strategy suggests that choice will be predicted by a person's outcome affect ratings multiplied by attribute importance, whereas Lexicographic choice should only be predicted by attribute importance and a positive or negative sign based on which alternative is "winning" on that attribute.

These predictive analyses, which involve comparing affect-based models consistent with coherence shifting to more mathematically rational models, are important

to investigate because previous studies of coherence shifting either manipulated outcome utilities to make one choice more attractive (Simon et al., 2004), did not conduct predictive analysis (Carpenter et al., 2016), or did not compare models that follow from different theories (Simon & Spiller, 2016).

In addition to affect towards outcomes predicting choice, influence may occur in the opposite direction. Once an early preference for one specific treatment alternative emerges, preferences towards alternatives and outcomes should shift to support a final choice.

> Hypothesis 2: In Experiments 1 and 2, preferences towards decision information (outcome affect and attribute importance) will change during decision making based on time and final treatment choice in order to make the chosen alternative more attractive. I hypothesized a *strong* form of the coherence shifting, in which ratings of outcome affect for the *chosen* treatment will increase from before a decision to the middle of the decision, while ratings for the *non-chosen* treatment decrease over the same interval. Similarly, importance ratings for attributes favoring the chosen treatment will increase from before a decision to the middle of the decision, while attributes favoring the non-chosen treatment will decrease over the same interval.

This hypothesis relies on a strong interpretation of coherence shifting, where the slope of the line showing change in affect—or importance ratings—over time is positive or negative depending on whether the outcomes come from—or the attributes favor—the chosen treatment vs. the non-chosen treatment, respectively. In a more *general* form of

coherence shifting, the slope of the line showing change in outcome affect ratings over time for the chosen treatment will be greater than the corresponding slope for the non-chosen treatment. Similarly, the slope of the line showing change in importance ratings over time for attributes favoring the chosen treatment will be greater than the corresponding slope for attributes favoring the non-chosen treatment.

This hypothesis holds that coherence shifting, as demonstrated with a job offer task (Carpenter et al., 2016; Simon et al., 2004; Simon et al., 2008; Simon & Spiller, 2016), will also occur in health decision making tasks. To test whether this hypothesis fits the connectionist constraint-satisfaction model proposed by Simon et al., preferences must be measured at pre-, mid-, and post-decision times. This is important because if preferences were only shown to change between pre-decision and post-decision measures, that could be attributed to theoretical accounts of changes in preferences taking place *after* a decision, such as cognitive dissonance (Simon & Holyoak, 2002).

If preference changes do not occur until the post-choice stage, this would not be compatible with coherence shifting as an emotional regulation function that occurs during the decision process. This proposed function and findings by Carpenter et al. (2016) suggest the following hypothesis.

> Hypothesis 3: In both experiments, as overall coherence shifting from the pre-choice to mid-choice time increases, aversive feelings towards the decision as a whole will decrease. Also, as the extent of overall coherence shifting increases, the extent to which physiological arousal decreases will increase, particularly between mid- and post- choice times. In other words, coherence shifting and

arousal decrease will show a dose-response relationship in Experiment 2 (Abelson, 1995).

Carpenter et al. (2016) found evidence that a difficult trade-off decision with relatively emotionally-neutral outcomes and attributes led to aversive feelings during the decision process. In an experiment using physiological arousal as an index of aversive feelings, they found that the pattern of arousal over time was dependent on the extent of coherence shifting engaged in by participants. Those who shifted their preferences strongly over time showed a reduction in physiological arousal between the mid- and post-decision times, whereas those who did less shifting of preferences did not reduce their arousal.

Carpenter et al. found that aversive feelings arise from the level of conflict (distance between outcomes) within a decision but did not specify any difference based on the context or content (attributes, outcomes) of the decision. Given that health treatment decisions are likely to have higher aversive feelings from the decision context, this effect should be more pronounced in health decisions and should depend on the content of specific health decisions:

> Hypothesis 4: Tasks with attributes that threaten higher-level goals (Experiment 1) and tasks with more severe physical outcomes (Experiment 2) will lead to stronger aversive feelings towards the overall decision and higher physiological arousal compared to tasks with less severe outcomes and which threaten lower-level goals.

Difficult health decisions differ from relatively emotionally-neutral tasks such as job selection in a number of ways that may affect emotional salience. Aversive feelings might arise from attributes that affect higher-level goals like safety and survival (Luce, 1998), physical outcomes that are often visceral and easily imaginable, and outcomes that predict near-term physical discomfort or pain (Bechara & Damasio, 2005; Loewenstein et al., 2001). These negative emotions are likely to increase with the severity of physical outcomes, such that a relatively mild health decision will be less pleasant than a neutral task such as job selection, and a serious health decision with severe consequences will be less pleasant than that. This should lead to stronger aversive feelings towards the overall decision in these scenarios and higher physiological arousal when presented with more emotionally salient tasks. Experiment 1 will compare two tasks that threaten lower- or higher-level goals and Experiment 2 will compare two tasks threatening the same high-level goals but with different levels of outcome unpleasantness.

Exploratory Analysis. In addition to negative emotion, it important to know if coherence shifting becomes more or less common as a decision becomes more serious and if the magnitude of coherence shifting changes as outcome emotional salience and the level of goals threatened increase. Unfortunately, it is difficult to predict this based on available empirical evidence. Carpenter et al. (2016) found that people depleted of executive resources by a separate frustrating task showed lower coherence shifting, but these findings do not extend to the effect of feelings arising from the decision itself. Thus, how differences in disease seriousness influence coherence shifting was examined in an exploratory manner.

Carpenter et al. also found that coherence shifting correlated with existing measures of emotional regulation. This analysis was repeated in the interests of converging evidence. Additionally, increasing confidence that a person's final decision was correct may be a mechanism by which coherence shifting reduces negative emotion. Final decision confidence and changes in confidence were examined for a relationship with coherence shifting.

Overview of Experimental Studies. These hypotheses were investigated using two experiments. Experiment 1 expanded upon original research on coherence shifting by comparing an emotionally-neutral decision task (job offer selection) with an emotion-laden health trade-off decision (a serious disease treatment decision) that threatened lower- and higher-level goals respectively. Experiment 2 utilized objective, physiological measures to examine changes in emotion during coherence shifting with two difficult health trade-offs differing in severity (unpleasantness) of outcomes. Both of these experiments required the creation of balanced, difficult, emotionally-salient health trade-off stimuli, which were created using a pilot study.

PILOT STUDY: TRADE-OFF DEVELOPMENT

In order to simulate a difficult trade-off decision, it was necessary to create scenarios that included two alternatives with outcomes that were balanced on a number of attributes. Simon et al. (2004) accomplished this by creating two fictional companies, Bonnie's Best and Splendor, selecting four attributes common to job selection, and providing outcomes that were better or worse for either company. The outcomes chosen were often objectively better or worse than some moderate value (e.g., industry average salary.) Some decision making studies describe these outcomes as conflicted or negatively correlated: Every outcome that is positive for one job is negative for the other, as can be seen in Table 1.

Creating difficult health trade-off scenarios for the current project entailed additional challenges beyond those faced by Simon et al. (2004). This project examined bi-directional influences between preferences and choices, rather than solely focusing on choice affecting preferences. This means that participants must be free to choose either alternative using the same outcome information, but without overwhelmingly favoring one alternative. This required a more precise balancing of outcome favorability or utility between alternatives. In order to be a difficult trade-off decision, differences on attributes must be large enough to not be trivial (e.g., $5 vs. $7 co-pay would be trivial) while also avoiding individual outcomes that are so extreme that choice would be guided by them entirely (e.g., death as a side effect).

It was also necessary to balance outcomes across multiple attributes. Health treatments are complex and multi-faceted, and many of these attributes threaten very

closely-held goals and have outcomes that are not easily comparable. Thus, even if two health treatments are moderately different on price and moderately different in 5-year survival rate, the individual outcomes on those measures must also be somehow comparable or one attribute will become irrelevant. One potential solution to this problem discussed by Peters (2006) is affect as common currency, or the idea that people can compare many dissimilar possible outcomes using their affective reaction to each outcome rather than any kind of quantitative or reflective process. Thus, the purpose of the pilot study was to determine outcomes of real medical treatments that provided approximately equal affective reactions. In order to create a tradeoff that is predictably difficult across participants, affect towards these outcomes must also be generally shared across individuals.

In the pilot study, affect towards outcomes was assessed using a modified version of Thurstone's (1928) method of equal-appearing intervals. The original purpose of this method was to quantify attitudes. To do this, Thurstone would identify a large number of qualitative, textual statements about a topic (e.g., religion), and then ask judges to sort the statements into numbered piles ranging from least favorable to most favorable towards the concept. Statements that were placed under a single number reliably by several judges could be assigned that number as a value, and statements corresponding to consecutive values could be used to construct a ranked scale. A person could then give yes/no agreement to items on this scale, and the numerical values of their "Yes" items could be averaged to quantify their attitude.

Method

The goal of this pilot was to develop two difficult multi-attribute health treatment trade-off decisions, similar to the job offer task developed by Simon et al. (2004). These include treatments for a mild disease with visceral but moderate negative outcomes and treatments for a serious disease with severe negative outcomes. First, I identified real diseases that can lead to difficult trade-off decisions, including Type 1 diabetes, Crohn's disease, and malignant melanoma. Information about treatments and symptoms for these diseases was collected from reputable online sources and medical literature (Tiziani, 2017). Next, information about these treatments was separated from original sources into lists of outcomes for six attributes: side effects, administration method, out-of-pocket cost, duration of symptoms, efficacy, and mortality.

Participants. Participants included 62 Clemson University students (age $M = 19$; 87% female) recruited through Clemson's undergraduate psychology research system. Participants were treated in accordance with APA ethical guidelines.

Procedures. Participants completed six online card sort procedures following Thurstone's procedure with a separate card sort for each attribute. They sorted a list of outcomes based on either their immediate emotional reaction (for side effects) or how unpleasant it would be to experience them (for all other scales), placing all outcomes in categories ranging from 1 "Not at all unpleasant" to 7 "Most unpleasant." Participants then provided importance ratings for each of the six attributes towards any treatment decision on a scale 5-point scale from 1 ("Not at all important") to 5 ("Extremely important").

Results

Following guidelines provided by Thurstone (1928), outcomes were sorted according to their median affect value to find outcomes that represent a similar amount of unpleasantness. The other criterion for selecting outcomes was that the unpleasantness ratings showed relatively low variability, as outcomes with high variability in affect ratings would not generalize well to the experimental studies. Attributes were selected as follows. Attribute importance ratings on a 1 to 5 scale were highest for mortality (M = 4.71, SD = 0.64), followed by efficacy (M=4.29, SD=0.69), out-of-pocket costs (M=3.97, SD=0.94), side effects (M=3.65, SD=0.83), duration of symptoms (M=3.52, SD=0.84), and administration (M=2.94, SD=0.92). Given the higher importance of mortality rates versus other attributes and the subjective perceptions described below, mortality outcomes were discarded.

To construct a trade-off decision with sufficient differences between alternatives, outcomes with equivalent affect ratings were grouped, and then outcome groups 2 to 3 affect units apart were compared side-by-side. Based on this initial comparison, a serious disease treatment decision with serious outcomes was constructed using outcomes with low variability that had been rated 4 or 6 on unpleasantness (2 units apart). A mild disease treatment decision was constructed using outcomes rated 2 and 4 on unpleasantness (2 units). The tradeoffs for mild and serious disease are presented in Tables 2 (Mild) and 3 (Serious). These trade-off decisions were used in Experiments 1 and 2.

The scenarios presented to participants included a longer textual explanation of each outcome in addition to the summary table below.

Table 1*: Difficult job offer trade-off. Reproduced from Simon et al., 2004, with salary values adjusted from Simon & Spiller (2016). (+) and (-) indicate favorable and unfavorable outcomes. Participants did not see these symbols.*

	Splendor	Bonnie's Best
Salary (Industry average = $50,000)	$49,2500 (-)	$51,000 (+)
Office	Private office (+)	Noisy cubicle (-)
Vacation Package	2 weeks (-)	2 weeks plus retreat (+)
Commute	Short (18 min) (+)	Long (40 min) (-)

Table 2: *Treatment attributes and outcomes for a mild disease with moderate consequences. Median unpleasantness scores ("Unpleas." ,interquartile range in parentheses) are shown to the left (Treatment T) or right (Treatment N) of each outcome. Participants did not see these scores. Higher numbers are more unpleasant.*

	Unpleas. (IQR)	Treatment T	Treatment N	Unpleas. (IQR)
Administration method	2 (1)	One pill 3 times a day for 14 days	One intramuscular shot, then one pill 2 times a day for 7 days	4 (2)
Side effects	4(1)	Nausea	Chills	2 (0.25)
Efficacy (% of people cured of disease 1 week after treatment ends)	2 (1)	87%	63%	4 (1)
Duration of Symptoms (How long you'll feel disease symptoms.)	4 (1)	14 days	1 day	2 (0)

Table 3: *Treatment attributes and outcomes for a serious disease with severe consequences. Median unpleasantness scores (Unpleas., interquartile range in parentheses) are shown to the left (Treatment K) or right (Treatment M) of each outcome. Participants did not see these scores. Higher numbers are more unpleasant.*

	Unpleas. (IQR)	Treatment K	Treatment M	Unpleas. (IQR)
Administration method	4 (2)	One intramuscular shot, then one pill 2 times a day for 7 days	One injection into the spinal fluid	6 (2)
Side effects	6 (1)	Seizure (50% chance)	Nausea (50% chance)	4 (1)
Efficacy (% of people cured of disease 1 week after treatment ends)	4 (1.25)	67%	51%	6 (1)
Duration of Symptoms (How long you'll feel disease symptoms.)	4 (1.25)	7 days, in hospital	10 days (no hospitalization)	6 (1)

EXPERIMENT 1: EMOTIONALLY NEURTRAL VERSUS MEDICAL TREATMENT DECISIONS

Experiment 1 examined participant decision-making in both a relatively emotionally-neutral tradeoff decision (job offer selection) adapted from Simon et al. (2004) and a medical treatment tradeoff decision threatening higher-level goals and with more emotionally laden outcomes (treatments for a serious disease). These two tasks were selected to produce strong differences in negative affect from decision information (outcomes, attributes) between tasks. This experiment had three purposes. The first was to examine the predictive validity of affect on final choice in both tasks. The second was to determine if coherence shifting occurs in both a replication of the original task used for this research as well as a novel medical treatment decision. Third, the two tasks were compared on subjective measures of aversive feelings from the decision and for the magnitude of coherence shifting which had occurred. Participants completed both tasks to allow for individual (within-subjects) comparisons in decision-making.

Method

Participants. A statistical power analysis was performed for sample size estimation using data from Simon's et al. (2004) second job offer experiment, using the critical choice-by-time interaction (r = .404). With an alpha = .05 and power = 0.80, the projected sample size needed to reproduce this interaction effect size (GPower 3.1.9.2) was N = 83. For logistic regression, a minimum sample of N = 140 would be required for the most complex proposed model (weighted additive; Vittinghoff & MuCulloch, 2007, van Smeden et al., 2016).

Participants included 125 Clemson University psychology students (Age: M = 19.05, SD = 1.41, 60.8% female). Participants were treated according to APA ethical guidelines under the supervision of Clemson University's Institutional Review Board.

Design. The overall experiment followed a 2 (decision task: job offer, serious disease treatment) by 3 (measurement time: pre-choice, mid-choice, post-choice) within-subjects design, although analysis focused on pre- and mid-choice times. Final choice between alternatives for each decision task (Serious disease: Treatment K or M; Job offer: Bonnie's Best vs. Splendor) was dichotomous and was used as a between-subjects predictor variable or a criterion depending upon the analysis. Analysis for each decision task was conducted separately. The use of different attributes and outcomes between tasks and the number of dependent variables would make meaningful direct comparisons of shifting in preferences over time difficult.

Dependent variables measured at all three time points included affect towards outcomes (8 per task) and subjective attribute importance (4 per task). Other dependent variables included mid-choice feelings towards alternatives (valence, arousal, and dominance toward each alternative), aversive feelings from the decision as a whole (4 items), and final choice for each task. These were measured at (or shortly after, for final choice) the mid-choice time period. (A mid-choice initial leaning and confidence for both leaning and final choice were also recorded.)

Measures. Affect (valence) towards outcomes was rated on a 10-point scale of predicted happiness ranging from -5 ("Extremely unhappy") to +5 ("Extremely happy"), excluding 0. This replaced the desirability scale used in prior coherence shifting research

(Carpenter et al., 2016; Simon et al., 2004) with the direct measure of affect used by Charpentier et al. (2016). Desirability is conceptually strongly related to affect towards an object and can directly influence behavior in a similar manner (Peters, 2006). Subjective attribute importance was measured using a 9-point scale of importance ranging from 1 ("no weight") to 9 ("maximum weight") (Simon et al.).

Feelings towards alternatives were measured using a short, 3-item Self-Assessment Manikin developed by Lang (1980) and Bradley and Lang (1994). The Self-Assessment Manikin is a rapid self-report assessment of three dimensions of affect: positivity (valence), arousal, and dominance. Each dimension was rated on a 9-point pictorial semantic-differential scale, where participants selected one of the 5 images or the 4 midpoints between the images for an intermediate value (See Figure 1). Following Schlösser et al. (2013), I reversed the Positivity dimension images so that "desirable" ratings (positive, calm, and in control) for all three dimensions were on the right.

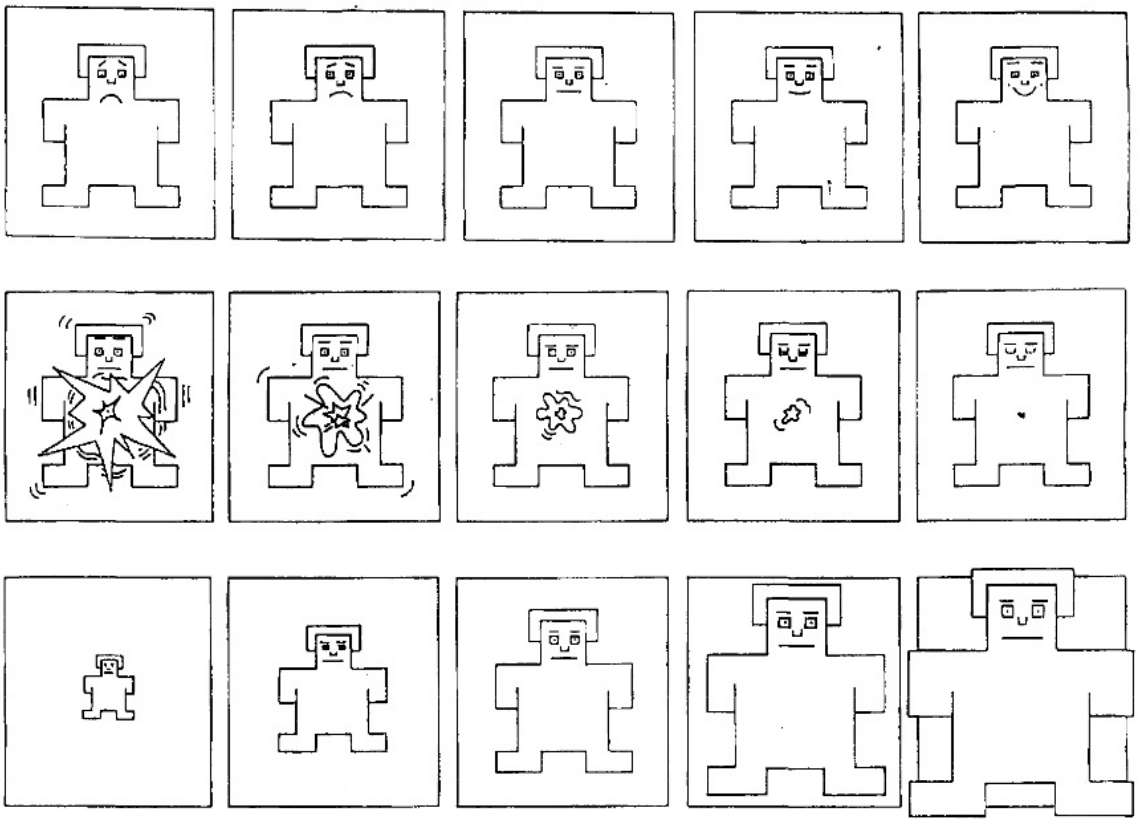

Figure 1. The Self-Assessment Manikin (Bradley & Lang, 1994). This is used to rate affective reaction dimensions of Positivity (top panel), Arousal (middle panel) and Dominance (bottom panel). The Positivity scale has been reversed from Bradley and Lang.

Aversive feelings towards the decision (affect) were measured with an averseness index utilized by Carpenter et al. (2016), including four 9-point scales using the terms Anxious, Stressed, Unpleasant, and Conflicted. These ranged from 1 ("Not at all anxious") to 5 ("Moderately Anxious") to 9 ("Extremely anxious"), except substituting the other emotions for "Anxious" in their scales. Final choice was a dichotomous measure of the alternative chosen.

Ratings of outcome affect ratings and attribute importance were collected at all three time points. Feelings towards alternatives and aversive feelings towards the decision were only collected at the mid-choice time. Final choice was collected between the mid-choice and post-choice ratings. For more details on timing, see the procedures.

Participants also completed several demographic and control questions that may impact emotions towards outcomes or alternatives. These included direct or close personal experience with the disease used in the serious disease task, as prior experience should increase the strength of affective response (Slovic et al., 2004; Bechara and Damasio, 2005). Fear of injection was assessed using a single-item measure adapted from the Marks & Matthews (1979), as the health decision task involved injections. Participants also reported their health insurance status, as Wong et al. (2013) found that cancer patients prioritize monetary cost over survival (and vice versa) depending on insurance availability and income. Participants completed two existing measures of emotional regulation strategy: the Emotion Regulation Questionnaire (ERQ; Gross & John 2003) and the Berkeley Expressivity Questionnaire (BEQ; Gross & John, 1998). Both of these scales use a 7-point Likert scale of agreement. The ERQ consists of 10 statements about how a participant controls their emotions and is divided into two subscales; Cognitive Reappraisal (e.g., "I control my emotions by changing the way I think about the situation I'm in.") and Expressive Suppression (e.g., "I control my emotions by not expressing them."). The BEQ consists of 20 items assessing the extent to which the participant expresses positive emotion (Positive Expressivity) and negative emotions (Negative Expressivity) in their daily life, and the strength of these expressions

(Impulse Strength). The BEQ was calculated as three separate subscales as well as an overall expressivity score.

Procedure. Experimental sessions took place a laboratory with 1 to 4 participants. All stimuli were presented through MediaLab (v2012). Participants completed both the job offer task and the serious disease task in a single session, and were randomly assigned to complete either the job offer or disease task first. Participants completed a practice Self-Assessment Manikin scale before the first task.

Job Task. The job offer task was adapted from Simon et al. (2004). During the first, pre-choice phase of the task, participants were presented with information entitled "Waiting for a Job Offer," where participants were asked to imagine that they are about to gradate and will be interviewing for a job. They were presented with 11 possible outcomes of job offers and rated each one in terms of their predicted happiness or unhappiness upon choosing it, i.e., outcome affect ratings. These included the 8 outcomes used in Table 1 plus 3 distractor outcomes, in list form. Participants were then presented with the four attributes available in Table 1 and asked to rate their importance in possible job offers, i.e., attribute importance ratings. These 12 measures were pre-choice ratings, outside of any decision context. No alternatives or decision matrix were present at this time.

During the second, mid-choice phase, participants were presented with job offers from two large fictional retail-store chains, Splendor and Bonnie's Best, using the decision information available in Table 1. These jobs were described as similar in all other attributes (e.g., promotion opportunities, size, and stability). This information was

presented in both paragraph and matrix form (as in Table 1), with explanations of each outcome in the paragraphs. After being told to consider all aspects of the job offers, participants were instructed that a third company is considering buying one company or the other, so they were not able to choose between jobs until later. Participants then rated their predicted affect towards the 8 outcomes and provided importance weights for the 4 attributes. For each job offer, participants were asked to imagine how they would feel if they selected that job and rate those feelings by completing a Self-Assessment Manikin. After this, they were be instructed to keep the two job offers in mind and complete the four-item index of aversive feelings towards the decision, as used in Carpenter et al., (2016).

Following this was the final choice and third, post-choice phase. Participants were instructed that the third company would not buy either of the companies in the decision, and that they should continue with the choice. Participants were shown the job offer information again and chose a job offer. Participants rated their affect towards the 8 outcomes and importance weights for the four attributes.

Serious Disease Task. Timing and measures for the serious disease task were the same as for the job offer task. The structure of disease task was adapted from the job offer task used by Simon et al. (2004). The pre-choice information for the serious disease task asked participants to imagine that they had been diagnosed with a strain of malaria after a mosquito bite and that a doctor would soon be with them to talk about treatment options. Malaria was selected from diseases rated in the pilot due to some participants not knowing of more common diseases (e.g., MRSA, tetanus). Participants then provided

ratings for the 8 outcomes and four attributes in Table 3, as well as three distractor outcomes selected from moderately unpleasant pilot items. Participants were presented with the treatments from Table 3 in both paragraph and matrix form, along with instructions that they were being asked to take part in the decision process (a shared decision) due to the trade-offs in costs and benefits of the two treatments. This included a delay instruction stating that the participants would have to wait to decide because the doctor was still waiting for laboratory bloodwork that could prevent them from taking one or both treatments. The participants then completed mid-choice measures. At the choice and post-choice times, participants were instructed that the blood tests were ready and that both treatments were still available, and then selected a treatment and provided and post-choice ratings.

Participants completed demographic and external influence measures after completing both tasks. Participants were fully debriefed including instructions that all treatments discussed were fictional before being released.

Past research (Simon et al., 2008; Carpenter et al., 2016) included unrelated reasoning tasks between each measurement time as distractors in order to reduce memory for previous ratings. Such distractors were omitted from Experiment 1 in both tasks due to a software error, but were implemented in Experiment 2.

Results

Statistics. Generalized eta-squared was used for MANOVA and ANOVA effect sizes, as partial eta-squared may overestimate effect sizes in repeated-measures designs (Bakeman, 2005). According to Cohen's (1988, p. 286) conventions for η_G^2, 0.26 and

above represents a large effect, 0.13 is a medium effect, 0.02 is a small effect, and less than 0.02 is negligible. Akaike's Information Criterion (AIC; Akaike, 1973) was used for predictive model fit as described below. Lower AIC indicates better fit. Using conventions from Burnham and Anderson (2004), models which have a fit within 2 AIC of the best-fitting model have substantial support, while models more than 10 AIC below the best-fitting model have essentially no support.

Predicting Choice: Serious Disease Task (Experiment 1). A separate predictive analysis was conducted on each task. Seven participants were eliminated from predictive analysis for the serious disease task as multivariate outliers on all predictor and outcome variables, resulting in a sample size for the disease task of $N = 119$. Treatment K was chosen by 77 participants (64.7%), and Treatment M was chosen by 42 participants (35.3%). Hypothesis 1 predicted that a model including pre-choice affect towards outcomes and mid-choice feelings towards alternatives would be the best predictor of choice (Model 2.1 in Table 4), when compared to other models based on outcome affect alone (1.0), attribute importance alone (3.0), or a linear combination of outcome affect and the related attribute importance (3.1, 4.0, 4.1). Multiple logistic regression analyses were conducted for several models to examine the predictive ability of different decision-making models for this task. See Appendix B for full model descriptions. Model omnibus statistical tests, effect sizes, and fit information for the serious disease tasks are available in Table 5. Models were evaluated individually rather than in a stepwise logistic regression because, due to the selection of predictors from multiple decision strategies, not all models were nested. Models that significantly predicted choice were Model 2.0

(Mid-choice feelings towards alternatives; χ^2 (6) = 90.39, $p < .001$), Model 2.1 (Mid-choice feelings and pre-choice outcome affect; χ^2 (14) = 107.41, $p < .001$), Model 3.1 (Lexicographic; χ^2 (2) = 19.65, $p < .001$), and Model 4.1 (Weighted additive with interaction; χ^2 (6) = 33.85, $p = .027$). No other models significantly predicted choice.

Models were compared for fit using AIC (Cohen et al., 2003). AIC approximates variance accounted for but favors more parsimonious models, with lower AIC values indicating better fit (Akaike, 1973). Model 2.1, including pre-choice affect towards outcomes and mid-choice feelings for alternatives, showed the best fit (AIC = 76.24). Models that are within 2 AIC of the best-fitting model are also considered to have substantial support. Model 2.0 (AIC = 77.26), containing only mid-choice feelings, is equally predictive of choice. No other models provide substantial fit. This only partially supported Hypothesis 1, which predicted that pre-choice affect with mid-choice feelings (Model 2.1) would out-predict mid-choice feelings alone (Model 2.0).

Regression coefficients and odds ratios for predictors in the serious disease task are presented in Appendix C (Tables C1.1 to C1.3). In both models 2.0 and 2.1, all six mid-choice ratings of feelings towards alternatives (positivity, arousal, and dominance) significantly predicted choice. Positivity most strongly predicted choice. In Model 2.1, participant became 7.53 times more likely to choose Treatment K for every 1 point increase in positivity towards K and became 4.55 (1/0.22) times more likely to choose Treatment M for every 1 point increase in positivity towards M. In model 2.1, pre-choice affect towards spinal injections (Treatment M administration) and 10 days without hospitalization (Treatment M duration) also significantly predicted choice. For every 1

point increase in affect towards spinal injections, participants became 1.77 times more likely to choose Treatment K. For every 1 point increase in affect towards 10 days without hospitalization, participants became 1.67 (1/0.60) times more likely to choose Treatment M.

Predicting Choice: Job Task. Thirteen participants were eliminated from predictive analysis for the job task as multivariate outliers, resulting in a sample size for the job task of $N = 112$. Only one participant was an outlier on both the serious disease and job tasks. Splendor was chosen by 70 participants (62.5%) and Bonnie's Best was chosen by 42 participants (37.5%). Model omnibus statistical tests, effect size, and fit information for the job task are available in Table 6. All models significantly predicted choice ($p < .01$) except for Models 3.0 (Attribute importance) and Model 3.1 (Lexicographic). Model 2.0, including only mid-choice feelings towards alternative, was the best-fitting model (AIC = 63.15). This only partially supports Hypothesis 1, which predicted that pre-choice affect ratings would improve prediction above mid-choice feelings, as in Model 2.1. Based on comparison of AIC, no other models provide substantial fit. Regression coefficients and odds ratios for predictors in the job-choice task are presented in Appendix C, Tables C2.1 to C2.3. In model 2.0, all six mid-choice ratings of feelings towards alternatives (positivity, arousal, and dominance) significantly predicted job choice. Positivity most strongly predicted choice. In this model, participants became 4.14 times more likely to choose Splendor for every 1 point increase in positivity towards Splendor. Participants became 2.63 (1/0.38) times more likely to choose Bonnie's Best for every 1 point increase in positivity towards Bonnie's Best.

Table 4: *Schematic versions of theory-based models and parameters. Hypothesis 2 states that Model 2.1 would be the most predictive of choice.*

Model	Predictors	Critical Parameters
1.0: Affect: Charpentier et al.	Affect for 8 outcomes (alternatives A and B each have 4 outcomes)	$\beta_{1\text{-}8}(\text{Affect}_{A1\text{-}4,\ B1\text{-}4})$
2.0: Mid-choice feelings only	Feelings for alternative A + feelings for alternative B	$\beta_1\text{Pos}_A + \beta_2\text{Arousal}_A + \beta_3\text{Dom}_A + \beta_4\text{Pos}_B + \beta_5\text{Arousal}_B + \beta 6\text{Dom}_B$
2.1: Affect: Schlösser et al.	Affect for 8 outcomes + feelings for alternative A + feelings for alternative B	$\beta_{1\text{-}8}(\text{Affect}_{A1\text{-}4,\ B1\text{-}4}) + \beta_9\text{Pos}_A + \beta_{10}\text{Arousal}_A + \beta_{11}\text{Dom}_A + \beta_{12}\text{Pos}_B + \beta_{13}\text{Arousal}_B + \beta_{14}\text{Dom}_B$
3.0: Attribute-based	Importance for 4 attributes	$\beta_{1\text{-}4}(\text{Imp}_{\text{Attribute1-4}})$
3.1: Lexicographic	Affect for 2 outcomes on high-importance attribute	$\beta_1(\text{Affect}_{\text{A,most imp.}}) + \beta_2(\text{Affect}_{\text{B,most imp.}})$
4.0: W.Add (Main effects)	Affect for 8 outcomes + Importance for 4 attributes	$\beta_{1\text{-}8}(\text{Affect}_{A1\text{-}4,\ B1\text{-}4}) + \beta_{9\text{-}12}(\text{Imp}_{\text{Attribute1-4}})$
4.1 : W.Add (Interaction)*	Affect for 8 outcomes + Importance for 4 attributes + Affect x Importance for A + Affect x Importance for B	$\beta_{1\text{-}8}(\text{Affect}_{A1\text{-}4,\ B1\text{-}4}) + \beta_{9\text{-}12}(\text{Imp}_{\text{Attribute1-4}}) + \beta_{13\text{-}16}((\text{Affect}_{A1\text{-}4}) \times (\text{Imp}_{\text{Attribute1-4}})) + \beta_{17\text{-}20}((\text{Affect}_{B1\text{-}4}) \times (\text{Imp}_{\text{Attribute1-4}}))$

See Appendix B for a detailed explanation of models. Pos = positivity, Dom = dominance, and Imp = importance. * Interaction is only between the outcome and its corresponding attribute (e.g., $\text{Import}_{\text{Attribute1}}$(Side effects) x Affect_{A1}(Seizure), $\text{Import}_{\text{Attribute2}}$ x Affect_{A2}, etc.). most imp. = most important attribute. W.Add = Weighted Additive.

Table 5: *Model statistical tests and fit for the Study 1 Serious Disease task.*

Model	*Log Likelihood*	*χ2 (df)*	*P*	*McFadden's Pseudo-R^2*	*McFadden's R^2 adjusted*	*AIC*
0 Null	-76.82	-	-	0	0	161.8
1.0 Affect (Charpentier)	-71.62	10.42 (8)	0.237	0.0678	-0.036	161.23
2.0 Affect (Mid-choice)	-31.63	**90.39** (6)**	**< .001**	0.588	0.510	77.26
2.1 Affect (Schlösser)	-23.12	**107.40 **(14)**	**< .001**	0.699	0.517	76.24
3.0 Attribute-based	-72.57	8.52 (4)	0.074	0.055	0.003	155.13
3.1 Lexicographic	-67.00	**19.65 **(2)**	**< .001**	0.128	0.102	140.00
4.0 W.Add (Main effects)	-68.30	17.05 (12)	0.148	0.111	-0.045	162.60
4.1 W.Add (Interaction)	-59.90	**33.85* (20)**	**0.027**	0.220	-0.040	161.80

N = 118. * $p < .05$, ** $p < .001$. AIC = Akaike's Information Criterion. W.Add = Weighted Additive.

Table 6: Predictive model statistical tests and fit for the Study 1 Job task.

Model	*Log Likelihood*	*χ2 (df)*	*P*	*McFadden's Pseudo-R^2*	*McFadden's R^2 adjusted*	*AIC*
0 Null	-74.10	-	-	0.000	0.000	150.19
1.0 Affect (Charpentier)	-61.58	**25.02* (8)**	0.002	0.169	0.061	141.17
2.0 Affect (Mid-choice)	-42.52	**63.15** (6)**	< .001	0.426	0.345	99.05
2.1 Affect (Schlösser)	-37.92	**72.36 **(14)**	< .001	0.488	0.299	105.83
3.0 Attribute-based	-72.84	2.51 (2)	0.285	0.017	-0.010	151.68
3.1 Lexicographic	-71.52	5.16 (4)	0.271	0.035	-0.019	153.03
4.0 W.Add (Main effects)	-58.84	**30.52* (12)**	0.002	0.206	0.044	143.67
4.1 W.Add (Interaction)	-51.83	**44.52* (20)**	0.001	0.300	0.031	145.67

N = 112. * $p < .05$, ** $p < .001$. AIC = Akaike's Information Criterion. W.Add = Weighted Additive.

Coherence Shifting: Serious Disease Task (Experiment 1).

Serious Disease: Importance. Mean comparisons for coherence shifting were also conducted separately for each task. The strong form of the coherence shifting Hypothesis (#2) predicted that attributes favoring the chosen treatment increase from before the decision to the middle of the decision, while attributes favoring the non-chosen treatment would decrease over the same interval. The general coherence shifting hypothesis only predicted a slope difference for the change in ratings over time rather than a positive vs. negative slope. These analyses did not include the post-decision data because changes between pre- and mid-choice times must be present if changes in scores are caused by coherence shifting, as opposed to cognitive dissonance.

In the context of the two treatments in the serious disease decision task, if the strong form of coherence shifting occurs, these changes should follow opposite patterns as follows. Participants who selected K should increase their importance ratings for attributes favoring K (Administration method, Efficacy) between the pre- and the mid-choice periods and decrease their importance ratings for attributes favoring M (Side effects, Duration) over the same interval. In contrast, participants who selected M should decrease their importance ratings for attributes favoring K between the pre- and the mid-choice periods and increase their importance ratings for attributes favoring M over the same interval.

To test these predictions, importance ratings were examined using a 2 (Treatment chosen: K or M; between subjects) by 2 (Alternative Favored: K or M; within subjects) by 2 (Time: pre- or mid-choice; within subjects) mixed-model MANOVA with

importance ratings paired as favoring K or M as the 2 dependent variables. A significant 3-way choice-by-favored-by-time interaction with importance scores shifting to make the chosen alternative more attractive would provide support for the general form of coherence shifting under Hypothesis 2. The strong form of coherence shifting would be supported if these changes follow the pattern described above.

Six participants were identified as multivariate outliers on serious disease importance scores and were removed from analysis, resulting in a sample of 119. Treatment K was chosen by 77 of these participants (64.7%) and M was chosen by 42 (35.3%). Full MANOVA results are available in Appendix D (Table D1.1). Figure 2 shows the data relevant to the coherence shifting hypothesis. M choosers appear to have shown the strong form of coherence shifting, while K choosers appear to show no coherence shifting, as the slopes of the time-change lines are parallel. In the omnibus multivariate test, the 3-way interaction of choice, time and whether attributes favor K or M was significant (*Pillai's* = .066, $F(2,116) = 4.11$, $p = .016$, $\eta_G^2 = 0.016$), with a negligible effect on importance. This supports the general form of the coherence shifting hypothesis. (Other significant main effects and interactions were found, as shown in Appendix D.)

Given the different patterns for choosers of K vs. M, separate within-subjects 2 by 2 (time by treatment favored) repeated-measures ANOVAs were conducted for these groups. A significant treatment-favored by time interaction supported the strong form of coherence shifting for M choosers ($F(1, 41) = 5.40$, $p = 0.025$, $\eta_G^2 = 0.062$), a small effect. This interaction was not significant for K choosers ($F(1, 76) = 2.12$, $p = 0.150$ η_G^2

=0.014), which provides no support for coherence shifting. See Table D1.2 for full ANOVA results. Thus, coherence shifting was supported for K choosers but not for M choosers.

In summary, the omnibus 3-way interaction was significant but showed a negligible effect size. Also, coherence shifting was found only for K choosers. Thus, the evidence for coherence shifting regarding importance scores is weak.

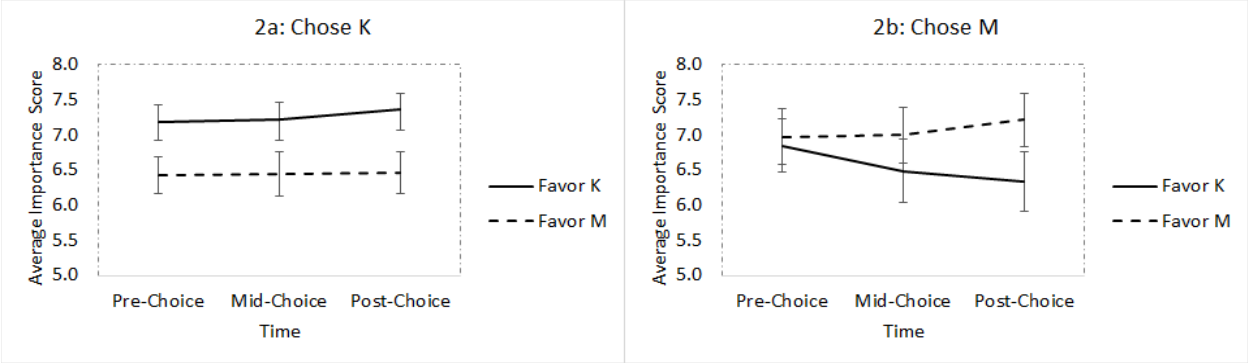

Figure 2.. Study 1 average serious disease attribute importance ratings favoring K or M by choice and time. M choosers (2b) altered their importance ratings between the pre- and mid-choice times showing the strong form of coherence shifting. K choosers (2a) did not shift scores. Error bars = 2SE.

Serious Disease: Affect. The strong form of coherence shifting predicted that affect towards outcomes of the chosen treatment would increase from before the decision to the middle of the decision, while affect towards outcomes of the non-chosen treatment would decrease over the same interval. Furthermore, for participants who chose K, affect towards treatment K outcomes should increase and affect towards treatment M outcomes should decrease between pre- and mid-choice ratings. In contrast, for those choosing M, affect towards M outcomes should increase and affect for K outcomes should decrease in

the same interval. To test for coherence shifting, affect ratings were examined using a 2 (Treatment chosen: K or M; between subjects) by 2 (Outcome Treatment: K or M; within subjects) by 2 (Time: pre- or mid-choice; within subjects) mixed-model MANOVA with outcome ratings on each attribute as the 4 dependent variables. A significant choice-by-treatment-by-time interaction in a direction that made the chosen alternative more attractive would provide support for coherence shifting.

One participant was identified as a multivariate outlier in all affect dependent variables, resulting in a sample of 124. Of these, 81 chose Treatment K (65.3%) and 42 chose M (34.7%). Figure 3 shows the data relevant to the coherence shifting hypothesis. Although the data do not support the strong form of coherence shifting, they seem consistent with the general form in which affect ratings for outcomes of the chosen treatment increase from pre- to mid-choice more than ratings for outcomes of the non-chosen treatment. The predicted choice by outcome treatment by time interaction was significant. (*Pillai's* = .089, $F(4,119) = 2.89$, $p = 0.023$, $\eta_G^2 = 0.023$), showing a small effect on affect ratings. Thus, the outcome affect ratings supported the general form of coherence shifting.

The MANOVA also showed that outcome affect ratings increased significantly from before ($M = -1.86$, $SD = 1.17$) to in the middle of the decision ($M = -1.30$, $SD = 1.08$), (*Pillai's* = 0.819, $F(4,119) = 134.81$, $p = <.001$, $\eta_G^2 = .0.513$), with a very large effect size. Figure 3 shows that all outcomes were rated negatively, which makes sense given that these outcomes were selected to be highly unpleasant. This main effect means that participants optimistically rated outcomes as less unpleasant in the decision context

than in the de-contextualized situation before the decision. This could explain why the data supported the general but not the strong form of coherence shifting. Other significant main effects and interactions were found, as shown in Appendix D (Tables D2.1 and D2.2).

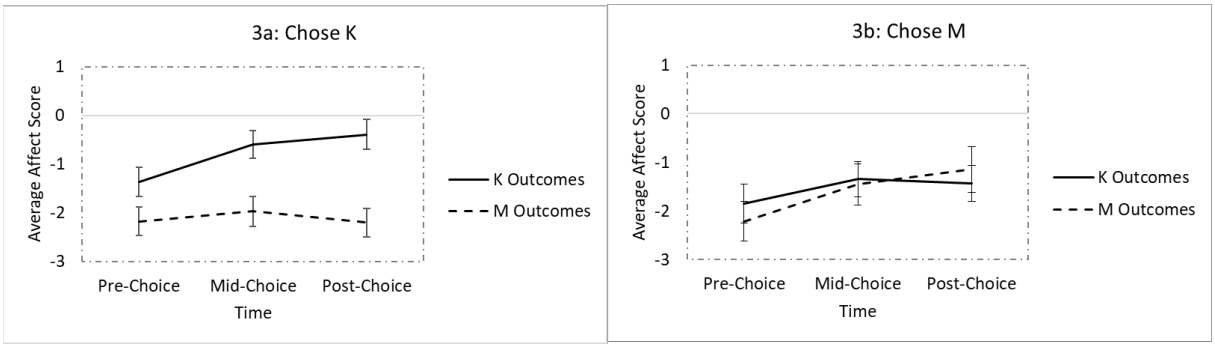

Figure 3.. Study 1 average serious disease outcome affect ratings by choice and time.Both types of choosers generally increased their affect scores between the pre- and mid-choice times, but sharper increases occurred for outcomes of the chosen treatment. All sub-figures are in the same scale. Error bars = 2SE.

Four univariate ANOVAs with the same independent variables showed that the predicted choice by outcome treatment by time interaction was only significant for Duration of symptoms ($F(1,122) = 9.20$, $p = 0.003$, $\eta_G^2 = 0.015$), a negligible effect size. Duration followed the pattern expected with general coherence shifting, with scores increasing over time but with steeper increases for the chosen treatment.

In summary, the general form of coherence shifting was supported for affect towards outcomes and showed a small effect size. The strong form of the prediction in Hypothesis 2 was not supported.

Coherence Shifting: Job Task.

Job task: Importance. Similar to the serious disease task, the strong version of coherence shifting would be supported in the Job Task if those choosing Splendor rated

attributes favoring Splendor (Office, Commute) as more important and attributes favoring Bonnie's Best (Salary, Vacation) as less important between the pre- and mid-choice times. Bonnie's Best choosers should show the opposite importance changes. Importance ratings were examined using a 2 by 2 by 2 (final choice by job favored by time) mixed-model MANOVA with the importance ratings grouped by job favored as the 2 dependent variables. Seven participants were identified as multivariate outliers on importance scores, resulting in a sample of 118. Of these, 75 chose Splendor (65.2%) and 43 chose Bonnie's Best (34.8%). Figure 5 shows importance scores relevant to the coherence shifting hypothesis. The predicted choice by time by job favored interaction was significant, (*Pillai's* = 0.222, $F(2, 115) = 16.36$, $p < .001$, $\eta_G^2 = 0.064$), showing a small effect on importance ratings. As shown in Figure 4, importance ratings shifted as predicted by the strong form of the coherence shifting hypothesis. Other significant main effects and interactions were found. See Appendix D (Table D3) for full statistical results.

Job task: Affect. Coherence shifting of affect ratings towards job outcomes was examined using a 2 by 2 by 2 (final choice by job outcome by time) mixed-model MANOVA with the outcome affect ratings paired across the 4 attributes as dependent variables. Seven participants were identified as multivariate outliers on affect scores, resulting in a sample of 118. Of these, 75 chose Splendor (65.2%) and 43 chose Bonnie's Best (34.8%). Figure 5 shows a summary of changes in affect scores by choice, outcome job, and time. The predicted 3-way interaction was not significant, (*Pillai's* = 0.073., $F(4, 113) = 2.22$, $p = 0.072$, $\eta_G^2 = 0.017$), with a negligible effect size. Thus, coherence shifting

in job outcome affect scores was not supported. Other significant main effects and interactions were found (see Appendix D, Tables D4.1 and D4.2). (In univariate ANOVAs with the same independent variables, the 3-way interaction testing coherence shifting was only significant for Office type ($F(1.\ 116) = 5.53$, $p = 0.020$, $\eta_G^2 = 0.006$), but this effect was negligible in size.)

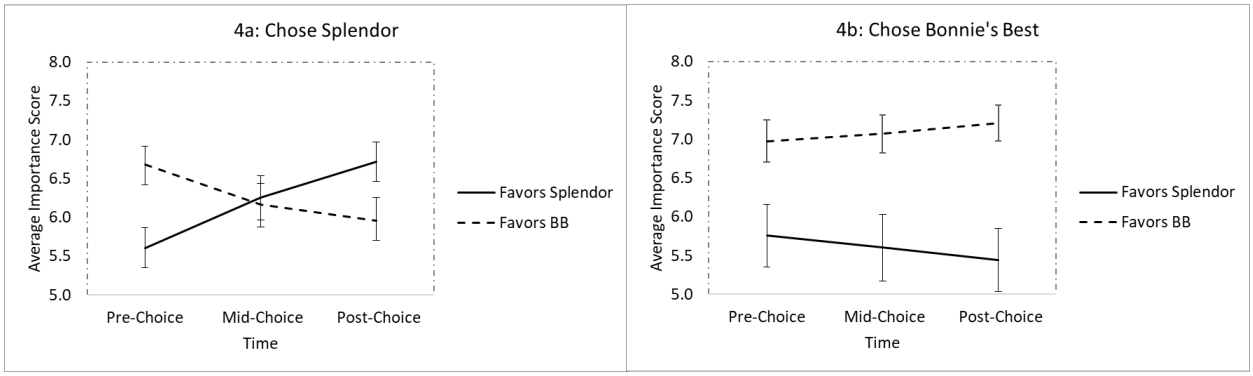

Figure 4. Study 1 average job task attribute importance ratings favoring Splendor or BB by choice and time.Choosers of each treatment changed scores from the pre- to mid-time in opposite directions in accordance with the strong form of coherence shifting under Hypothesis 2. BB = Bonnie's Best. Error bars = 2 SE.

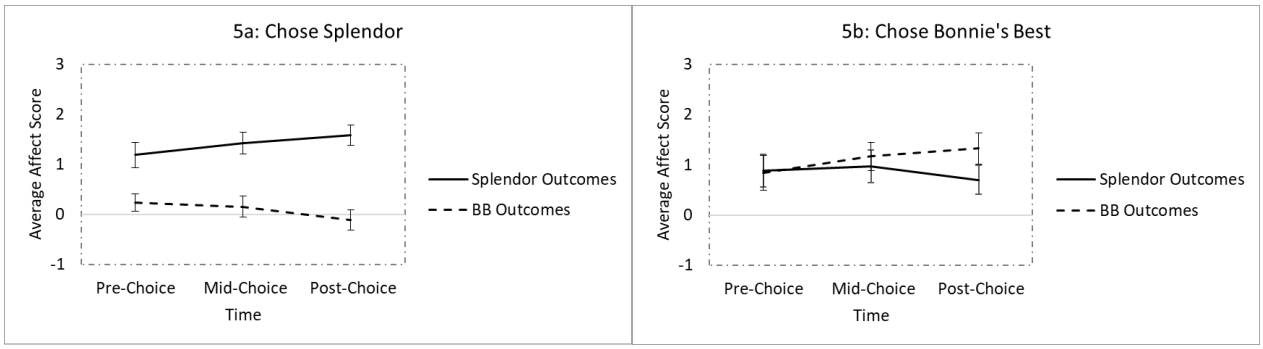

Figure 5. Study 1 average job task affect outcome ratings by choice and time. The choice by treatment by time interaction for changes from pre- to mid-choice ratings was significant. BB = Bonnie's Best. Error bars = 2 SE.

To summarize the results for the job task, coherence shifting of importance scores was significant and showed a small effect size. Coherence shifting was not supported for affect ratings.

Hypothesis 3 predicted that higher coherence shifting would lead to a decrease in aversive feelings towards the decision and in arousal after the mid-choice time. Hypothesis 3 was not examined in Experiment 1.

Between-Task Comparisons and Exploratory Analysis.

Participants who were identified as multivariate outliers on any previous analysis were removed from the following analyses, resulting in a sample size of $N = 106$.

Between-Task Aversive Feelings: Hypothesis 4 suggested that aversive feelings towards a decision would be higher in a serious disease treatment task than the job selection task, which threatened lower-level goals. To assess differences in aversive feelings related to the decision task as a whole, participant scores on the 4 items on the aversive feelings scale were compared using 4 paired-samples t-tests. Descriptive statistics and statistical tests are presented in Appendix D (Table D5). Participants reported significantly higher aversive feelings during the serious disease task than during the job task for all scales. Aversive feelings were rated 1.32 points higher on an 8-point scale, on average. Reliability for the aversive feelings scale was high in both the serious disease ($\alpha = 0.872$) and job ($\alpha = 0.875$) tasks.

Between-Task Coherence Shifting. One important contribution of this research program is a comparison of coherence shifting between tasks that involve different decision contexts and outcomes on substantially different attributes. Past research has

identified differences in decision strategy based on attributes threatening higher- vs. lower-level goals, but not differences in the degree of coherence shifting (Luce, 1998; Payne & Bettman, 2004). In Experiment 1, job task attributes threatened relatively lower-level goals such as comfort (e.g., office size, commute), whereas the serious disease task threatened higher level goals such as health and survival. (Serious disease outcomes were also manipulated to be highly unpleasant, creating a likely confound with differences in emotional salience. Differences in emotional salience will be examined independently in Experiment 2.) Given the lack of empirical evidence to form a hypothesis, I have presented between-task comparisons in an exploratory manner.

To compare the levels of coherence shifting individuals engaged in between tasks, attribute importance and outcome affect ratings were combined into aggregate variables using a technique used by Simon et al. (2004) and Carpenter et al. (2016). Importance and affect ratings at each time were scaled and then combined into measures where +1 would indicate complete favorability towards one alternative (Treatment K or Splendor job) and -1 would indicate complete favorability towards the other (M or Bonnie's Best). This is mathematically defined in Appendix G. Aggregated scores at the pre-choice time are subtracted from the mid-choice time, indicating the degree of change in favorability towards attributes or outcomes is consistent with coherence shifting. Absolute values (A) of these scores represent overall strength of coherence shifting (CS) for affect (ACS_{Aff}) and importance (ACS_{Imp}).

Paired-samples t-tests were used to compare absolute pre- to mid-choice coherence shifting between the job and serious disease tasks (See Appendix D, Table

D5). Overall shifting of *importance* scores was not significantly different between the disease task ($\underline{M}$ = 0.052, *SD* = 0.048) and the job task (*M*= 0.066, *SD* = 0.069), *t* (105) = -1.91, *p* = 0.058. In contrast, overall shifting in *affect* scores was significantly higher in the disease task (*M* = 0.110, *SD* =0.092) than in the job task (*M*=0.076, *SD*=0.056). *t* (105) = 3.20, *p* = 0.002.

Emotional Regulation and Confidence. Following Carpenter et al. (2016), correlations were calculated in both tasks between absolute scores of coherence shifting and self-report measures of emotional regulation strategies. Descriptive statistics and statistical tests are reported in Appendix D (Table D5). Correlations are reported in Appendix D (Table D6). Pre- to mid-choice shifting of affect scores (ACS_{Aff}) in the serious disease task were weakly but significantly correlated with Cognitive Reappraisal, *r* (106) = 0.227. *p* = 0.019, r^2 = 0.051, i.e., stronger coherence shifting was associated with a higher tendency to change one's thinking in response to an emotional stimulus. No other significant correlations between coherence shifting and emotion regulation strategies were found. This pattern differs from findings by Carpenter et al. (2016) using the same job task as in the current study. They found that coherence shifting of combined desirability (similar to affect) and importance scores correlated positively with the expression suppression scale of emotional regulation, but not with cognitive reappraisal.

Another exploratory analysis focused on the idea that coherence shifting may serve to increase participants' confidence that they have made the correct decision, which should reduce their negative emotions related to the decision. Ratings of confidence in the final decisions in the serious disease and job tasks were examined for correlations

with overall coherence shifting. See Appendix D (Table D6) for full correlations. Stronger shifting of affect scores in the serious disease task was correlated with greater decision confidence, r (106) = 0.304, p = 0.002, r^2 = 0.092. Stronger shifting of job importance scores correlated with greater confidence in the final job chosen, r (106) = 0.208, p = 0.033, r^2 = 0.043. Coherence shifting was not significantly correlated with any *changes* in confidence across the course of the decision, as calculated by subtracting confidence in the mid-choice leaning from final choice confidence ratings.

Prior Experience and Outside Influences. In order to determine if any factors outside of the decision context affected treatment choice (such as those reported in Wong et al, 2013), participants completed single-item measures of health insurance coverage (yes or no), experience with malaria in themselves or someone close to them (yes or no), and the extent to which they avoid medical procedures due to a fear of injections or needles. Only 3 participants (2.8%) indicated previous experience with malaria and only 4 participants (3.8%) reported that they did not have health insurance. Given the small number of participants in these categories, Fisher's exact tests were used for these comparisons. An independent-samples t-test was used to examine fear of needles and treatment choice. Those with malaria experience chose treatment K about as often (*1 out of 3*, 66.7%) as those without experience (67 of 103, 65.0%), *Fisher's Exact p* = 0.722. Those with health insurance chose treatment K about as often (66 of 102, 64.7%) as those without insurance (3 out of 4, 75% chose K), *Fisher's Exact p* = 0.564. Fear of needles was not significantly different between those who chose K (n = 69, M = 7.49, SD = 2.45) and those who chose M (n=37, M = 6.78, SD = 1.72), t (96.74) = 1.87, p = 0.086.

Experiment 1 Summary

The hypothesis regarding the predictive analyses (#1) was partially supported in both tasks. A model including pre-choice outcome affect and mid-choice feelings (2.1) was either similarly predictive (serious disease task) or less predictive (job task) of choice than a model only including mid-choice feelings towards alternatives (Model 2.0). Thus, including pre-choice affect ratings did not substantially improve prediction of choice. These results support the theory that choice is constructed from the affective evaluation of alternatives as a whole, since adding in pre-existing preferences towards individual outcomes either did not improve prediction much or made it worse. However, they do not support the hypothesis that pre-choice affect towards outcomes would improve the ability to predict choice. The expectation that the above affect-based models would predict choice better than attribute-based models like weighted additive and lexicographic was supported.

The coherence-shifting hypothesis (#2), which predicted that affect and importance ratings would shift prior to a final decision to be more favorable toward the alternative chosen, was supported for some importance ratings in both tasks, and (in a more general form) for affect ratings only in the disease task. These findings support the theory of coherence shifting in a health context, where emotional reactions are adjusted before choice as part of the decision-making process. However, the size of the coherence shifting effects was low; ranging from negligible ($\eta_G^2 = 0.016$) to small ($\eta_G^2 = 0.064$). Although shifting of affect scores in the job task did not replicate similar findings by

Simon et al. (2004) in outcome desirability ratings, shifting in importance scores was replicated.

Participants reported stronger task-related aversive feelings during the serious disease task, supporting Hypothesis 4. Participants showed stronger aversive feelings in a task with highly unpleasant outcomes that threaten higher-level goals than in a task with relatively emotionally neutral outcomes threatening lower-level goals. Participants also shifted affect scores more strongly between the pre- and mid-choice times in the serious disease task, but did not show differences in importance score shifting. A significant positive correlation was found between an emotion regulation strategy (cognitive reappraisal) and coherence shifting, but only in the serious disease task and only for affect scores. Taken together, these findings suggest that participants changed their affect towards outcomes to reduce overall aversive feelings towards a decision, but only when the task produced sufficiently negative emotions. Participants showed coherence shifting for importance scores in both tasks, but these shifts were not correlated with broad emotion regulation strategies. Experiment 2 examined the use of coherence shifting as emotion regulation more directly, using an objective physiological affect measure.

EXPERIMENT 2: PHYSIOLOGICAL AROUSAL AND MULTIPLE DISEASE TREATMENT DECISIONS

Experiment 2 expanded upon Experiment 1 in two ways. First, rather than using an emotionally neutral job selection task, participants in this experiment completed treatment decision tasks for a relatively mild disease with moderately negative physical outcomes and a serious disease with severe physical outcomes. Unlike the two dissimilar tasks in Experiment 1, these two tasks used the same attributes and used outcomes selected from the pilot to provide similar levels of conflict between alternatives. This was intended to reduce between-task differences in attribute-based threats to important goals (Luce, 1998). In addition, selecting outcomes that were an equal distance apart on pilot affect ratings (ratings of 2 & 4 vs. ratings of 4 & 6) provided a level of control over decision conflict as defined by Carpenter and Niedenthal (2017).

Second, in addition to self-report measures, this study used an objective measure of physiological arousal as an index of aversive emotions. If coherence shifting serves an emotional regulation role, physiological arousal should decrease for people who have shifted their preferences, as found by Carpenter et al. (2016).

Skin Conductance for Physiological Arousal

In addition to all measures used in Experiment 1, skin conductance was used as an objective measure of physiological arousal. Skin conductance is a measure of the electrical conductivity of the skin based on eccrine sweating, which is directly related to the autonomic nervous system. Skin conductance has seen extensive use as a measure of affective response in decision-making research (Figner & Murphy, 2011) and of general

stress or arousal in experimental settings (Boucsein, 2012). With roughly a 1 to 5 s delay before response to a specific stimulus, skin conductance is a slow or time-lagged measure compared to physiological measures such as event-related potentials. It is also sensitive to artifacts related to movement, respiration, and speaking. Even if it were a perfectly reliable measure of autonomic nervous system activity or physiological arousal, physiological arousal is simply an activation of the sympathetic nervous system, and relies further on cognitive evaluation of the context of the situation to be interpreted as a positive or negative emotional state (Loewenstein et al. 2001). Arousal is affected by many other physiological and mental processes, making it difficult to claim that any particular skin conductance response is definitely related to affect. Within affect, skin conductance only reliably aligns with the arousal dimension. Additional subjective measures, such as those included in this experiment, are usually required to assess positivity, dominance, or other dimensions (Figner & Murphy). In spite of these limitations, skin conductance response is sufficiently sensitive to detect immediate affective responses to anticipated rewards and punishments in decision making, and even to detect anticipatory affect that predicts decision-making in repeated-decision tasks such as the Iowa Gambling Task in line with the somatic marker hypothesis (Bechara, Tranel, & Damasio, 2000; Bechara et al., 1996).

One important issue in selecting a measure of skin conductance is the difference between phasic and tonic skin conductance measures, and their use for short-term or persistent affective responses. Tonic changes in skin conductance *level* (SCL) are gradual increases or decreases in skin conductance. Phasic changes or skin conductance

responses (SCR) are rapid, short-term elevations in conductance followed by a delayed decline. To examine the role of coherence shifting in regulating negative affect, Carpenter et al. (2016) measured SCRs occurring between 1 and 3 s after each screen where participants completed importance and desirability ratings. They defined these SCRs as negative emotion caused by decision conflict and proposed that lower SCRs in the post-decision period represented a regulation of emotion. However, they selected this measure and time period due to its use in past decision making research, where it was used to index short-term affective responses to the outcomes of decision-making (Bechara et al, 2000; Bechara et al., 1996). These are known as specific responses, or SCRs that are related to a specific stimulus. Carpenter et al. (2016) intended to measure medium-term (e.g., minutes) changes in aversive feelings towards a decision as a whole, but used a measurement window traditionally used to identify rapid (e.g., seconds) affective reactions to a specific stimulus (e.g., reading one decision outcome). It is not possible to distinguish between these sources of affect using SCR, a potential confound.

Carpenter et al. (2106) proposed that aversive feelings arise from the difficulty of a decision itself, with coherence shifting serving to reduce these feelings. If these feelings are persistent until a regulatory process occurs, their role should be more similar to an ongoing stressor than to a rapid affective reaction. Measures of SCL, or tonic skin conductance level, are more commonly used and reliable for measuring the impact of laboratory stressors (Boucsein, 2012). These includes threats of physical pain or watching disturbing video clips, but some are even accurate for detecting weaker, instruction-based threats (Boucsein; Kilpatrick, 1972). These measures use various methods of eliminating

specific responses, such as measuring only in windows devoid of responses or statistically removing spikes characteristic of a response. It is still not possible to distinguish between sources of negative affect or the nature of the decision-making process through this measure, but an SCL measure should be able to distinguish between changes in persistent aversive feelings on the one hand and simple, short-term affective responses to decision information on the other. Equipment used to record skin conductance necessarily records both SCR and SCL, so both were analyzed.

Method

Participants. An additional power analysis was conducted for the effect of time and psychological threat on skin conductance using a similar interaction taken from Kilpatrick (1972), $r = .331$. Using a multilevel model power analysis method provided by Bickel (2007) with an alpha = .05, power = 0.80, the projected sample size needed to reproduce this interaction effect size was $N = 46$. Thus, power analyses from Experiment 1 still provide the conservative estimate.

Participants included 95 Clemson University psychology students (Age: M=19.87, SD=2.61, 70.2% female). Participants were treated according to APA ethical guidelines under the supervision of Clemson University's Institutional Review Board. Recruitment in Study 2 was lower than in Study 1 because measuring skin conductance required testing participants individually and there were two sessions (two weeks apart) compared to Study 1's single session. This resulted in a lower power of analyses in Study 2. Participants were compensated at a rate of $10 for an initial session and $20 for a second session. This incentive was used to reduce participant dropout between sessions.

Design. The overall experiment followed a 2 (decision task: mild vs. serious disease treatment) by 3 (measurement time: pre-, mid-, post-choice) within-subjects design. Analysis focused only on the pre- and mid-choice times. Final choice between alternatives for each decision task (Mild disease treatment: Treatment T or N; Serious disease treatment: Treatment K or M) was dichotomous and was used as a between-subjects predictor variable or a criterion depending upon the analysis. An aggregate measure of preference shifting based on outcome affect and attribute importance ratings was used to predict changes in physiological arousal. Analyses for each decision task were conducted separately.

All dependent variables measured in Experiment 1 were included in Experiment 2. In addition, SCR and SCL were measured at a baseline and at pre-choice, mid-choice, and post-choice rating times.

Measures. All dependent variables and demographics measures used in Experiment 1 were also collected in Experiment 2. Skin conductance was recorded using a Biopac GSR100 skin conductance module, MP150 base module, and STP100C digital interface. AcqKnowledge software was used to record skin conductance data. Based on recommendations by Figner and Murphy (2011), disposable electrodes were placed on the distal (first) phalanges of the index and middle finger of the participant's non-dominant hand. Skin conductance sampling acquisition was set to 500 Hz (samples/sec). Hardware was set to record using DC, with amplification set to 5 μSiemens/V and a low-pass filter set to 1 Hz. Experimenters measured ambient temperature within the laboratory at the beginning and end of every session. Coded markers were automatically

placed in AcqKnowledge software through MediaLab software and were verified in real time by experimenters. Experimenters also recorded the time and nature of any disturbances that may have led to recording artifacts.

Procedure. Participants attended 2 sessions, a minimum of 2 weeks apart. Experimental sessions were conducted with one participant and one experimenter. Each session included either the serious disease treatment decision (identical to Experiment 1, see Table 3) or the mild disease treatment decision shown in Table 2. Participants were randomly assigned to complete the mild or serious disease task first. The materials and measures for the mild disease treatment decision were the same as those for the serious disease decision, except that the instructions asked participants to imagine that they had been diagnosed with a strain of influenza and that the outcomes listed were those in Table 2 rather than those in Table 3. Attributes were the same between tasks.

Stimuli were presented in MediaLab (v2012) software. The general timing of materials and measures were the same as in Experiment 1, except that the experimenter controlled when the participants proceeded between rating times. Skin conductance was only recorded during the baseline measure and at the pre-, mid-, and post-choice rating times. For each session, after providing consent, participants first had electrodes placed on their non-dominant hand. After at least five minutes, a baseline measure of skin conductance was recorded. This included an active baseline (participants taking a deep breath) followed by a two-minute passive baseline (during which they were instructed not to move or talk.). The process was repeated with fresh electrodes for null responses. Participants were asked to avoid body movement, any motion in their non-dominant

hand, or speech during tasks as to avoid recording artifacts. Each measurement period took no longer than 5 minutes and participants were informed when measurement was occurring and when they were free to move or talk.

After these instructions, the experimenter began recording skin conductance and the participants were shown the pre-choice instructions and ratings as in Experiment 1. Markers were automatically placed at each screen including experimental stimuli and at each rating. Skin conductance was recorded during both outcome and attribute ratings (and, for later tasks, choice). After this, participants completed either a spatial reasoning task (Ekstrom, French, & Harman, 1976) or items from a personal interests survey (Goldberg, 2010) as a distractor task. Distractors used different items between tasks and were presented in a cross-balanced order with instructions that they were not intelligence tests to avoid external stress.

After the first distractor, the participant was shown mid-choice materials including the instruction that the treatment decision must be delayed for laboratory bloodwork. This delay instruction was present in both disease tasks. The experimenter resumed recording and the participants completed mid-choice ratings. Participants then completed the second distractor. The experimenter then resumed recording and the participant made a final decision and provided post-choice ratings as in Experiment 1.

Participants then completed demographic and external influence measures, which were recorded in the first session. The experimenter then removed the disposable electrodes. The participant was then compensated with $10 in cash and asked to schedule a second session at least two weeks later. This was partially to allow time for preferences

to return to a pre-decision baseline. Simon et al. (2008) found that preferences revert very quickly over time, but only using relatively emotionally neutral tasks. More importantly, this time lapse was meant to reduce order effects and prevent affect from the prior decision from influencing arousal in the later decision.

In the second session, each participant returned and completed the second task. This followed the same pattern and timeline as the first session, only with the other disease treatment decision task and with alternate distractor task content. Participants were compensated with $20, fully debriefed, and dismissed after the second session.

Results

Predicting Choice: Mild Disease Task. A separate predictive analysis was conducted for each task. One participant was identified as a multivariate outlier on all predictor and outcome variables and eliminated from predictive analysis for the mild disease task. An additional 4 participants were excluded due to computer errors resulting in incomplete data. This resulted in a sample size for the mild disease task of $N = 90$. Treatment T was chosen by 24 participants (26.7%), and Treatment N was chosen by 66 participants (73.3%). Model multiple regression omnibus statistical tests, effect size, and fit information for the mild disease task are available in Table 7. All models significantly predicted choice, $p < .05$.

Table 7: *Model statistical tests and fit for the Study 2 Mild Disease task*

Model	*Log Likelihood*	*χ2 (df)*	*P*	*McFadden's Pseudo-R^2*	*McFadden's R^2 adjusted*	*AIC*
0 Null	-52.19	-	-	0.000	0.000	106.39
1.0 Affect (Charpentier)	-43.02	**18.36* (8)**	**0.019**	0.176	0.022	104.03
2.0 Affect (Mid-choice)	-27.19	**50.00** (6)**	**< .001**	0.479	0.364	68.39
2.1 Affect (Schlösser)	-19.23	**65.94** (14)**	**< .001**	0.632	0.363	68.45
3.0 Attribute-based	-46.01	**12.38* (4)**	**0.015**	0.118	0.042	102.01
3.1 Lexicographic	-48.54	**7.30* (2)**	**0.026**	0.070	0.032	103.09
4.0 W.Add (Main effects)	-38.7	**29.98* (12)**	**0.008**	0.258	0.029	103.4
4.1 W.Add (Interaction)	-35.6	**33.19* (20)**	**0.032**	0.318	-0.065	113.2

$N = 90$. * $p < .05$, ** $p < .001$. AIC = Akaike's Information Criterion. W.Add = Weighted Additive.

Hypothesis 1 predicted that Model 2.1, which contained affect towards outcomes outside of the decision context (pre-choice) and feelings towards alternatives during decision making (mid-choice), would best predict choice. Model 2.0, which included only mid-choice feelings for alternatives, showed the best fit (AIC = 68.39). Model 2.1 (AIC = 68.45) also strongly predicted choice (as it was within 2 AIC of Model 2.1). No other models provide substantial fit.

Regression coefficients and odds ratios for predictors in the mild disease task are presented in Appendix E (Tables E1.1 through E1.3). In both models 2.0 and 2.1, Positivity towards both alternatives significantly predicted choice. In Model 2.0, Dominance towards N also significantly predicted choice. Positivity most strongly predicted choice. In Model 2.0, participants became 2.35 times more likely to choose

Treatment T for every 1 point increase in positivity towards T. Participants became 2.85 (1/0.35) times more likely to choose Treatment N for every 1 point increase in positivity towards N.

The similar fits of Model 2.1 and 2.0 and the fact that none of the specific pre-choice affect scores in Model 2.1 significantly predicted treatment choice suggests that affect towards alternatives during decision making is the strongest predictor of choice and affect towards outcomes outside of the decision context was not predictive. This only partially supported Hypothesis 1.

Predicting Choice: Serious Disease Task (Experiment 2). Five participants were eliminated from predictive analysis for the serious disease task (Experiment 2) as multivariate outliers and an additional 2 participants were excluded due to experimental errors, resulting in a sample size of $N = 88$. Treatment K was chosen by 41 participants (46.5%) and Treatment M was chosen by 47 participants (53.4%). Model omnibus statistical tests, effect sizes, and fit information for the serious disease task are available in Table 8.

Findings were similar to the mild disease task. Only Models 2.0 and 2.1 significantly predicted choice. Model 2.0 (only mid-choice feelings towards alternatives) was the best-fitting model (AIC = 64.9). No other models provided substantial fit, including the hypothesized best-fitting model, 2.1. This only partially supported Hypothesis 1. In Model 2.0, positivity towards both alternatives significantly predicted job choice and most strongly predicted choice. In Model 2.0, participants became 1.97 times more likely to choose Treatment K for every 1 point increase in positivity towards

K. Participants became 2.63 (1/0.38) times more likely to choose Treatment M for every 1 point increase in positivity towards M (see Appendix E, Tables E2.1 and E2.2).

The findings from these predictive analyses were similar to those from Study 1, which also found that Model 2.0 was always the most predictive of choice or equally as predictive as Model 2.1 and that ratings of positivity towards alternatives were the strongest individual predictors of choice. Across both studies, Hypothesis 1 was partially supported—as affect towards alternatives during decision making predicted choice—and partially disconfirmed—as affect towards outcomes outside of the decision context did not predict choice.

Table 8: *Model statistical tests and fit for the Study 2 Serious Disease task*

Model	*Log Likelihood*	*χ2 (df)*	*P*	*McFadden's Pseudo-R^2*	*McFadden's R^2 adjusted*	*AIC*
0 Null	-60.79	-	-	0.000	0.000	123.58
1.0 Affect (Charpentier)	-57.72	6.15 (8)	0.631	0.051	-0.081	133.44
2.0 Affect (Mid-choice)	-25.46	**70.66** (6)**	< .001	0.581	0.482	64.92
2.1 Affect (Schlösser)	-20.15	**81.29** (14)**	< .001	0.669	0.471	70.30
3.0 Attribute-based	-57.77	6.04 (4)	0.196	0.050	-0.016	125.55
3.1 Lexicographic	-58.79	4.00 (2)	0.135	0.033	0.000	123.59
4.0 W.Add (Main effects)	-54.72	12.15 (12)	0.434	0.100	-0.098	140.36
4.1 W.Add (Interaction)	-48.57	24.44 (20)	0.224	0.201	0.037	139.15

$N = 88$. * $p < .05$, ** $p < .001$. AIC = Akaike's Information Criterion. W.Add = Weighted Additive.

Coherence Shifting: Mild Disease Task.

Mild Disease: Importance. The coherence shifting hypothesis (#2) stated that when treatment T was chosen, importance ratings for attributes favoring T would increase from pre- to mid-choice and attributes favoring N would decrease; with the opposite pattern expected when N was chosen. Four participants were excluded due to experimental errors, resulting in a sample of 91. Of these, 25 chose Treatment T (26.4%) and 67 chose Treatment N (73.6%). Figure 6 shows importance scores averaged across favored treatment by choice and time. Participants who chose N seemed to show the strong form of coherence shifting, with importance ratings increasing or decreasing over time depending on whether attributes favored their choice or not, respectively. In contrast, T choosers seemed to show little coherence shifting (parallel lines for change over time). In a 2 (choice) by 2 (treatment favored) by 2 (time) MANOVA, the predicted choice by treatment-favored by time interaction for importance ratings was significant, (*Pillai's* = .075, $F(2, 88) = 3.55$, $p = 0.033$, $\eta_G^2 = 0.018$), but showed a negligible effect size. Other significant main effects and interactions were found, see Appendix F (See Tables F1.1 and F1.2).

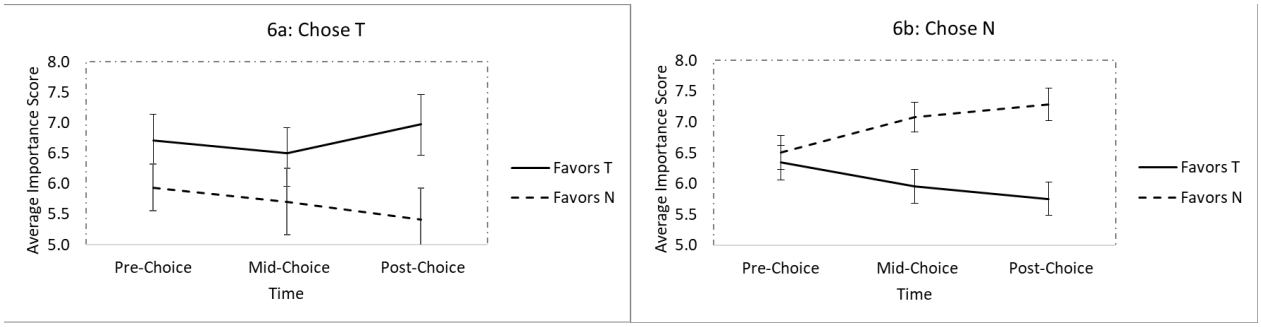

Figure 6. Study 2 average mild disease attribute importance ratings favoring T or N by choice and time. N choosers shifted importance scores between the pre- and mid-choice times in accordance with strong-form coherence shifting, whereas T choosers generally decreased all importance ratings but showed high individual variability. N = 91. Error bars = 2SE.

Separate within-subjects 2 by 2 (time by treatment favored) repeated-measures ANOVAs were conducted for T and N choosers. A significant treatment-favored by time interaction supported the strong form of coherence shifting for N choosers (F (1, 66) = 20.37, $p < .001$, η_G^2=0.105), a small effect. This interaction was not significant for T choosers (F (1, 23) < 1, p = .946 η_G^2=0.000), which provides no support for coherence shifting. See Appendix F (Table F1.2) for full ANOVA results. Thus, for the mild disease task, coherence shifting of importance ratings was supported for N choosers but not for T choosers.

Mild disease task: Affect. Four participants were excluded due to experimental errors and 1 was identified as a multivariate outlier on affect scores, resulting in a sample of 90. Of these, 24 chose Treatment T (26.7%) and 67 chose Treatment N (73.3%). Figure 7 shows outcome ratings changing in the direction predicted by the strong version of coherence shifting for both T and N choosers. In a 2 (choice) by 2 (outcome treatment) by 2 (time) MANOVA, the predicted choice by time interaction in overall outcome affect ratings was significant, (*Pillai's* = 0.226, F(2, 88) = 6.29, p = <.001, η_G^2=0.060),

representing a small effect. Appendix F (Tables F2.1 & F2.2) includes complete statistical analysis. Other significant main effects and interactions were found.

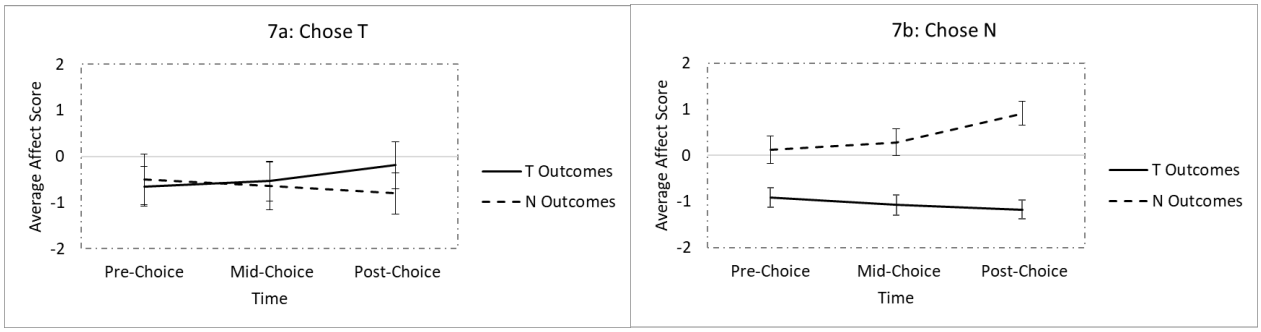

Figure 7. Study 2 average mild disease affect outcome ratings by choice and time. The choice by time interaction for changes from pre- to mid-choice ratings is visible. $N = 90$. Error bars = 2SE.

To summarize, for the mild disease task, coherence shifting of importance scores was supported for participants who made one choice, but not for those who made the other choice. Coherence shifting of affect scores was supported, with an effect size of $\eta_G^2 = 0.08$.

Coherence Shifting: Serious Disease Task (Experiment 2)

Serious Disease: Importance. Three participants were excluded as multivariate outliers and another 2 were excluded due to experimental error, resulting in a sample of 90. Of these, 41 chose Treatment K (45.6%) and 49 chose Treatment M (54.4%). Figure 8 shows changes in importance scores consistent with the general form of coherence shifting for K choosers and the strong form for M choosers. From the choice by treatment favored by time MANOVA, the 3-way interaction for importance predicted by the coherence shifting hypothesis scores was significant, (*Pillai's* = 0.098, $F(2, 87) = 4.72$, p

= .011, η_G^2 =0.025), and showed a small effect size. Other significant main effects and interactions were found, see Appendix F (Tables F2.1 and F2.2).

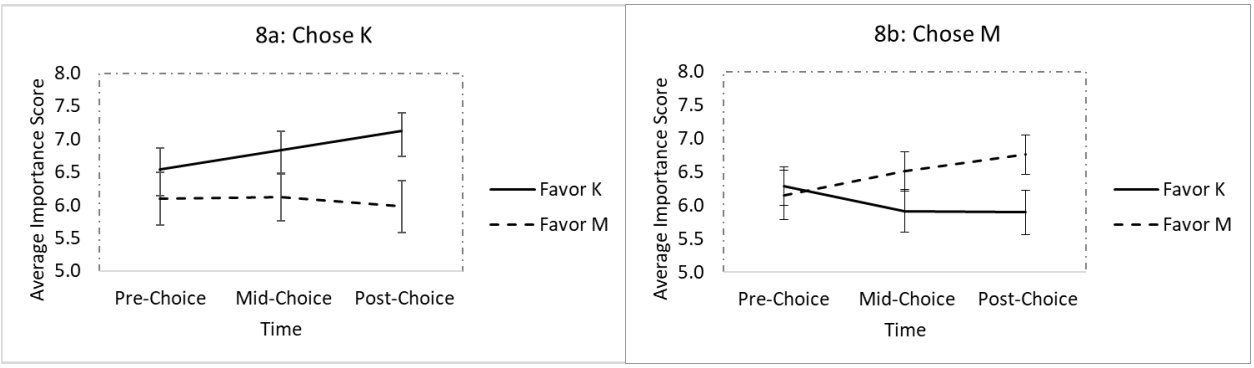

Figure 8. Study 2 average serious disease attribute importance ratings favoring K or M by choice and time. The predicted choice by time interaction was significant. $N = 90$. Error bars = 2SE.

.

Serious disease task: Affect. Two participants were excluded as multivariate outliers on affect scores and 2 were excluded due to experimental errors, resulting in a sample of 91. Of these, 43 chose Treatment K (47.3%) and 48 chose Treatment M (52.7%). The direction of changes shown in Figure 9 do not match the predictions of strong coherence shifting but do match the pattern of general coherence shifting seen in the serious disease task in Experiment 1. In the 2 by 2 by 2 choice by time by outcome-treatment MANOVA, the predicted choice by time interaction in overall outcome affect ratings was significant, (*Pillai's* = 0.368., $F(4, 86) = 12.52$, $p < .001$, $\eta_G^2 = 0.117$), and had a small effect size. However, this effect size was more than five times as large as the same interaction for this task in Experiment 1. The general form of coherence shifting (but not the strong form predicted in Hypothesis 2) was supported for coherence shifting in serious disease treatment outcome affect scores. Other significant main effects and

interactions were found, Appendix F (Tables F4.1 and F4.2) for full statistical results.

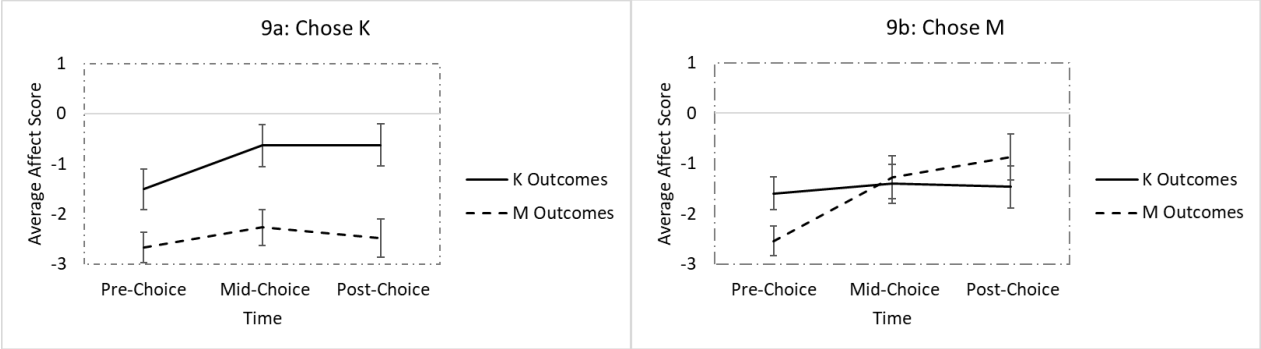

Figure 9.: Study 2 average serious disease outcome affect ratings by choice and time. Similar to the same task in Experiment 1, participants generally increased affect ratings between the pre- and mid-choice times, but showed steeper increases for their chosen treatment. $N = 90$. Error bars = 2SE.

Experiment 2 Coherence Shifting Summary

Findings in coherence shifting were similar between Experiments 1 and 2. Overall coherence shifting was supported for importance scores in all four tasks. Affect coherence shifting was not supported for the Experiment 1 job task, but was supported in all treatment decision tasks across the two experiments. More support was found for the general form of coherence shifting than for the strong form predicted in Hypothesis 2. The general form appeared in affect ratings for the serious disease task in both experiments and in Experiment 2 serious disease importance ratings. Effect sizes for the coherence shifting interaction were much larger in Experiment 2 than Experiment 1, especially for the serious disease task used in both experiments. This is likely an effect of the inclusion of a distractor task between the pre- and mid-choice ratings, which was omitted in the first Experiment.

Experiment 2 Coherence Shifting, Aversive Feelings and Physiological Arousal.

Hypothesis 3 predicted that higher coherence shifting would lead to a decrease in aversive feelings and physiological arousal after the mid-choice time, as shifting has been proposed as a mechanism for reducing decision-related negative affect (Carpenter et al. 2016; Carpenter & Niedenthal, 2017).

Physiological data were analyzed for a small pilot sample from Experiment 2 participants (n = 39). (Full analyses will be presented in an upcoming publication.) SCR and SCL values were calculated for the full period during which participants rated outcome affect, with responses to known artifacts (e.g., participant movement) subtracted. Tonic skin conductance (SCR) was calculated as the average amplitude of responses across this period in μS. Phasic skin conductance (SCL) was calculated as the minimum amplitude of skin conductance during this period. SCR or SCL scores from the resting baseline were subtracted from the respective measure. Participants who failed to display any skin conductance responses during at least one full recording period were eliminated from analysis for possible recording errors, resulting in sample sizes of n = 16 (Mild disease task) and n =14 (Serious disease task).

For each task, two multiple linear regression models were conducted with a composite measure of coherence shifting (zACSoverall, see Appendix G), linear and quadratic terms for recording time (pre-, mid-, and post-choice), and interactions between coherence shifting, linear time, and quadratic time as predictors and skin conductance (SCL or SCR) at the three recording times as the criterion. Phasic skin conductance

(SCL) is the preferred measure for long-term changes in arousal, such as overall arousal due to the decision as a whole (Boucsein, 2012). Hypothesis 3 would be supported if a coherence shifting x time (quadratic) interaction was significant and showed a sharper decrease in SCL between the mid- and post-choice rating times for people who showed stronger coherence shifting.

Neither time nor coherence shifting significantly predicted SCR in either task (*p*'s > .05), so no further data is presented. Figure 10 shows simple quadratic slopes of SCL for high, average, and low coherence shifting over time. Regression coefficients predicting SCL are shown in Appendix F (Table F5). Coherence shifting and the critical coherence shifting x time (quadratic) interaction were not significant predictors of SCL. Thus, coherence shifting did not lead to slope differences in SCL over time. Hypothesis 3 was not supported in this pilot sample. Strength of coherence shifting did not predict changes in arousal, but participants did show similar patterns of change in arousal across decision times. SCL was significantly predicted by both linear and quadratic terms for time in both the mild and serious disease tasks. In both tasks, SCL generally increased between pre- and mid-choice times and then decreased between mid- and post-choice times.

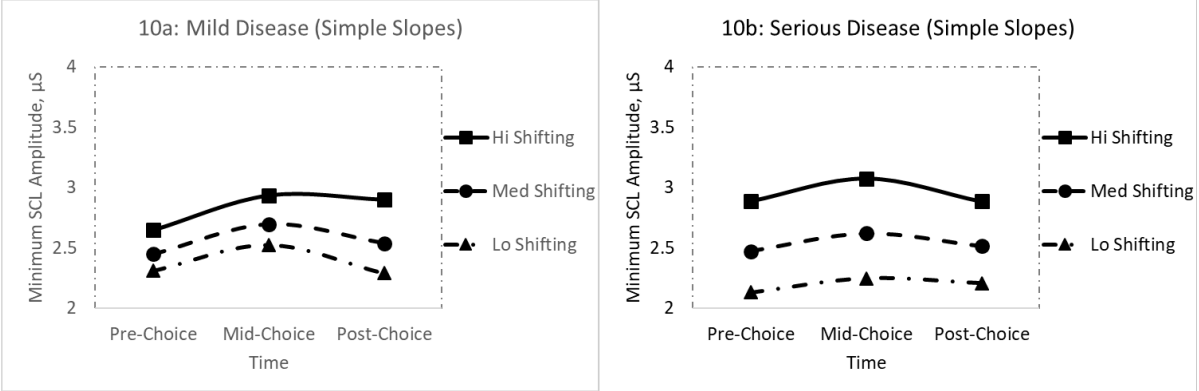

Figure 10. Study 2 simple quadratic slopes of SCL over time by level of coherence shifting. SCL changed significantly over time, but coherence shifting did not significantly predict slope differences. Hi = maximum non-outlier participant coherence shifting score (< 2SE). Med = average coherence shifting. Lo = minimum non-outlier participant coherence shifting. Mild disease $N = 16$, Serious disease $N = 14$).

Experiment 2 Between-Task Comparisons and Exploratory Analysis.

Participants who had been identified as multivariate outliers on any previous analysis were removed from the following analyses, resulting in a sample size of $N = 84$. Group descriptive statistics and statistical tests are presented in Appendix F (Table F6), while correlational results and overall descriptive statistics are available in Table F7.

Between-Task Aversive Feelings: Hypothesis 4 predicted that aversive feelings would be higher in tasks that have less pleasant physical outcomes. Both the mild and serious disease tasks threaten goals important to health and survival (e.g., efficacy of a treatment, intrusive side effects and administration methods), but serious disease outcomes were selected to produce higher negative affect than mild disease outcomes. Accordingly, aversive feelings towards a decision should be higher in a serious disease treatment than in the mild disease task. To assess differences in aversive feelings caused by decision tasks, participants' scores on the 4 items on the aversive feelings scale (feeling anxious, stressed, unpleasant, and conflicted) were compared using 4 paired-

samples t-tests. Participants reported significantly higher aversive feelings during the serious disease task than during the mild task for all scales, rating aversive feelings 1.17 points higher on an 8-point scale on average. Reliability for the aversive feelings scale was high in both the mild disease ($\alpha = 0.863$) and serious disease ($\alpha = 0.916$) tasks.

Between-Task Coherence Shifting. To compare the degree of coherence shifting individuals engaged in between tasks, attribute importance and outcome affect were combined into aggregate variables using the technique described in Experiment 1. Absolute values of these scores represent overall strength of coherence shifting for affect (ACS_{Aff}) and importance (ACS_{Imp}). Paired-samples t-tests were used to compare absolute pre- to mid-choice coherence shifting between the mild and serious disease tasks. See Appendix F for full statistics (Table F6). Overall shifting of importance scores was not significantly different between the mild disease task ($M = 0.084$, $SD = 0.075$) and the serious disease task ($M = 0.083$, $SD = 0.069$), t (83) = -0.08, $p = 0.934$. Overall shifting in affect scores was also not significantly different between the mild disease task ($M = 0.101$, $SD = 0.076$) than in the serious disease task ($M = 0.113$, $SD = 0.082$). t (83) = 0.94, p =0.359.

Emotional Regulation and Confidence. To further investigate Carpenter et al.'s (2016) findings concerning coherence shifting and emotional regulation, correlations were calculated between absolute coherence-shifting scores for both tasks and self-report measures of emotional regulation strategies from the BEQ and ERQ. Descriptive statistics and correlations are reported in Appendix F (Table F7). Pre- to mid-choice shifting of affect scores in the mild disease task had small positive correlations with

Negative Emotionality (BEQ; $r(82) = 0.293$, $p = 0.007$, $r^2 = 0.085$), Impulse Strength (BEQ; $r(82) = 0.273$, $p = 0.012$, $r^2 = 0.074$), Overall Emotional Expressivity (BEQ; $r(82) = 0.288$, $p = 0.008$, $r^2 = 0.083$), but was not significantly correlated with BEQ Positive Emotionality, ERQ expressive suppression, or ERQ cognitive reappraisal (p's > 0.05). Together, these findings indicate that people who are more likely to express their negative feelings strongly in their daily lives engaged in more coherence shifting regarding affect in the mild disease task. These results are in direct contradiction with results reported by Carpenter et al., who found that coherence shifting was significantly negatively correlated with the exact same subscales of the BEQ in a relatively emotionally neutral job selection task, in the opposite direction of correlation. No significant correlations were found between strategy measures and mild disease importance shifting or either type of serious disease score shifting.

Coherence shifting was also not found to correlate with final decision confidence or changes in confidence within each task.

Prior Experience and Outside Influences. In order to determine if any outside factors affected treatment choice, choice in both treatments was compared based on health insurance status, needle fear, and malaria experience (serious disease only). Two additional measures of experience with influenza were added for the mild disease task. One question asked if they or anyone close to them had ever been diagnosed with influenza or "flu" (yes or no), and the second question asked if they had personally been diagnosed with influenza within the past 6 months (yes or no). Only 1 participant indicated previous experience with malaria, so the effect of that experience on choice was

not compared. General flu experience was high (54 of 84, 64%), but only 3 participants reported that they had been personally diagnosed with influenza recently and only 5 participants reported that they did not have health insurance. Fisher's exact tests were used for these dichotomous comparisons. Participants with general influenza experience chose treatment T significantly less often (10 out of 54, 18.5% chose T over N) than those who did not report any lifetime experience with influenza (12 out of 30, 40.0% chose T), *Fisher's Exact p* = 0.031. It is possible that those without experience were less threatened by Treatment T's longer duration of flu symptoms, but without further unplanned statistical tests it is unclear which aspects of Treatment T might have been more attractive to those without flu experience (See Table 2 for outcomes). Those with a recent personal influenza diagnosis chose treatment T about as often (1 out of 3 chose T, 33.3%) as those without experience (21 out of 81 chose T, 29.5%), *Fisher's Exact p* = 0.603.

In the mild disease task, those with health insurance chose treatment T about as often (*N* = 21, 26.9% chose T) as those without insurance (*N*=1, 16.7% chose T), *Fisher's Exact p* = 0.501. For all participants, needle fear was not significantly different between those who chose Treatment T (*n*=22, *M*=3.55, *SD*=2.69) as those who chose N (*n*=62,*M*=2.81, *SD*=1.89), *t* (28.71) =1.19, *p* = 0.244. Similarly, in the serious disease task, those with health insurance chose Treatment K about as often (*n*=36, 46.2% chose K) as those without health insurance (*n*=2, 33,3% chose K). Needle fear was not significantly different between those who chose K (n=38, M=2.97, SD=1.87) and those who chose N (*n* = 46, *M* = 3.50, *SD* = 2.23), *t* (82) = -1.16, p=0.250.

Experiment 2 Discussion and Comparison with Experiment 1

Experiment 2 replicated the predictive results of Experiment 1, showing that a model with only mid-choice feelings towards alternatives was either the most predictive or equally predictive to a less parsimonious model also including pre-choice outcome affect ratings.

Coherence shifting received stronger support in Experiment 2, with some form of significant coherence shifting appearing in importance and affect ratings for both tasks. Experiment 2 serious disease affect scores replicated the general-form pattern of coherence shifting seen in the same task in Experiment 1. Effect sizes for coherence shifting were much larger in Experiment 2 in general than in Experiment 1. This is likely due to the influence of the unrelated reasoning task used as a distractor between ratings. However, it is notable that coherence shifting was still found in Experiment 1 without the distractor.

The finding that coherence shifting depended on which treatment participants chose in Experiment 1 serious disease importance ratings was also found in Experiment 2 mild disease importance ratings (See Figures 2 and 6). Participants who chose Treatment K in the Experiment 1 serious disease task and those who chose Treatment T in the Experiment 2 mild disease task generally did not significantly change their importance scores between the pre- and mid-choice times. Participants choosing M or N behaved in line with the strong form of coherence shifting. This finding is unexpected, especially given that participants facing the identical serious disease task in Experiment 2 showed coherence shifting for choosers of both treatments. The simple-effects ANOVAs for these

two tasks showed a significant main effect of treatment favored for those who did not shift scores (i.e., Treatment K and T choosers), and showed a large difference between attributes favoring each treatment at the pre-choice time. Alternatively, participants who showed coherence shifting had similar importance ratings for both types of attribute at the pre-choice time. This suggests that participants who were initially less conflicted, showing large differences in importance between attributes favoring each alternative, were more likely to choose one treatment (K or T) and did not need to shift their scores for that alternative to seem dominant. Participants who began the task feeling more conflicted were more likely to choose the other treatment (M or N) and shifted their scores to make that treatment more attractive.

Between-task comparisons in Experiment 2 replicated Experiment 1 in that aversive feelings towards the decision as a whole were stronger in the more emotionally-salient task, with scores being higher for the serious disease. This difference was found despite tasks in Experiment 2 threatening the same higher-level goals, supporting the prediction that increasing outcome unpleasantness also increases aversive feelings towards the task as a whole. More correlations were found between coherence shifting and emotion regulation strategies in Experiment 2 than in Experiment 1, but findings that coherence shifting was related to post-choice confidence were not replicated.

Significant positive correlations between mild disease affect shifting and expressivity, which is a component of emotion regulation, run counter to the finding that higher coherence shifting did not significantly predict changes in physiological arousal. This is likely due to the higher reliability of self-report measures in this study (affect

ratings and the expressivity questionnaire) when compared to the physiological arousal measure. Skin conductance is highly sensitive and often reflects artifacts such as bodily movement and task-irrelevant arousal (Boucsein, 2012), and many participants were excluded due to potential recording errors. In addition, the analyses of physiological data were statistically underpowered when compared to the self-report measures, as the pilot physiological data sample included less than $1/5^{th}$ the participants in correlational analyses.

GENERAL DISCUSSION

The purpose of this dissertation was four-fold. The first goal was to examine what decision information best predicts choice in a difficult health tradeoff decision that would also elicit coherence shifting. This represents a new contribution to research on coherence shifting, as past studies have focused on the impact of early preference for one alternative or another on decision information ratings without examining the source of the initial leaning. This project also integrated years of research on the extent to which affective reactions predict choice. The second goal was to determine if coherence shifting occurs in disease-treatment shared-decision tasks. These tasks, which represent a recent and important field of decision-making, threaten higher-level goals and include more emotionally salient outcomes than the decision tasks used in prior research (e.g., the job task). In both of our studies, participants completed two decision tasks that differentially threatened higher-level goals (Experiment 1) or had higher outcome unpleasantness (Experiment 2). This led to a third goal—investigating whether the degree of coherence shifting changed as higher-level goals are more threatened or as outcomes become less pleasant. The fourth goal was to determine if coherence shifting served as a strategy to reduce task-related negative affect and if it corresponds with previously-identified emotion regulation strategies. These three research questions were investigated using three tasks, including a relatively emotionally-neutral job selection task used in prior research, a mild disease treatment choice that presented moderately unpleasant outcomes, and a serious disease treatment choice that presented highly unpleasant outcomes. The

disease tasks also threaten higher-level goals (e.g., health, survival) as opposed to relatively lower-level goals in the job task (e.g., comfort, convenience).

Predicting Choice

In Hypothesis 1, I predicted that final choice would best be predicted by a model including pre-choice affect towards outcomes and mid-choice feelings towards alternatives. (See Table 9 for a summary of results regarding the predictive analyses and coherence shifting within and across tasks.) Hypothesis 1 was only partially supported. In every decision scenario, a model only including mid-choice feelings towards alternatives either showed better fit or substantially equivalent fit to a model that included both pre-choice affect ratings and mid-choice feeling towards alternatives. There was no case in which adding pre-choice affect ratings substantially improved model fit. This finding strongly supports the preference construction view of decision making and provides only weak support for the view of revealed preferences. Choice was best predicted by feelings towards alternatives after outcomes were presented *in the decision context* (e.g., arranged into two treatments). Under the revealed preferences conception of choice, participants should have based their decisions on their general predicted affect towards outcomes regardless of the decision context, combining or comparing their pre-existing feelings and reaching a decision that will match their overall preferences and long-term goals. The finding that preferences measured outside the decision context did not improve the prediction of participants' choices is important for doctors and patients facing a shared treatment decision because it indicates that patients are strongly influenced by the presentation of decision information in the specific context of the alternatives given.

Table 9: *Summary of Experiment 1 and 2 Results.*

	Study 1 Job Task		Study 1 Serious Disease Task	Study 2 Mild Disease Task		Study 2 Serious Disease Task
Prediction						
Best-Fit Model	Model 2.0		Models 2.1 & 2.0	Models 2.0 & 2.1		Model 2.0
AIC	**99.05****		**76.24** / 77.26****	**68.39** / 68.45****		**64.92****
Coherence Shifting	η_G^2		η_G^2	η_G^2		η_G^2
Importance Shift	**0.064****		**0.016***	**0.018***		**0.025***
Affect Shift	0.017		**0.023***	**0.060***		**0.118****
Between-Task Analyses	M (SD)	*t*	M (SD)	M (SD)	*t*	M (SD)
Aversive Feelings	**4.58 (1.86)**	**<***	**5.90 (2.19)**	**4.51 (1.81)**	**<***	**5.68 (1.95)**
Importance ACS	0.066 (0.069(	~	0.052 (0.48)	0.084 (0.075)	~	0.083 (0.069)
Affect ACS	**0.076 (0.056)**	**<***	**0.110 (0.076)**	0.101 (0.076)	~	0.113 (0.082)

* $p < .05$, ** $p < .001$. ~ = Not significantly different. AIC = Akaike's Information Criterion. ACS = Absolute Coherence Shifting.

Based on model fit, there was essentially no support for any model including pre-choice attribute importance, such as lexicographic and weighted additive (AIC differences from best-fitting model < 10; Burnham & Anderson, 2004). This suggests that even for decisions that strongly impact a person's future (e.g., career or health), people may use fast, affect-based strategies over slower, more considered strategies that weigh the importance of various attributes. However, participants received no coaching in decision strategies. Doctors in real shared treatment decisions usually coach their patients to use an attribute-based strategy (lexicographic decision making; Elwyn et al., 2012), and actual cancer patients have been found to prioritize different attributes based

on external life circumstances and goals (Wong et al., 2013). Also, the Wong et al. finding suggest an alternative interpretation of the predictive-analysis findings—that participants used fast heuristics because they were making a hypothetical decision. Research using decisions with real consequences for participants is needed to clarify this question.

In summation, predictive analysis of all 4 tasks provided support for a preference construction view of decision-making and the use of affect as the primary decision-making strategy over conceptions of revealed preferences and more cognitively-effortful importance-based strategies.

Coherence Shifting

Predictive analysis showed the impact of affect about decision alternatives on choice. Coherence shifting represents influence in the opposing direction, where an early preference towards one alternative will lead to changes in ratings of affect and importance for decision information. In Hypothesis 2, I predicted that participants would change ratings of affect towards outcomes and importance of attributes from the pre-choice time (outside the decision context) to the mid-choice time (in the decision context) in a manner that makes the chosen alternative more attractive and the non-chosen alternative less attractive. Changes must exist between the pre- and mid-choice times in order for coherence shifting to be a mechanism for choice rather than post-decision shifting of scores (e.g., cognitive dissonance.)

Importance Scores. There was general support for coherence shifting in importance scores across the four tasks in Experiments 1 and 2. Statistically significant

coherence shifting in some form was found in all four tasks. The evidence for coherence shifting of importance ratings was strongest for the job task and Study-2 serious disease task (see Table 9). In these tasks, participants increased sores for attributes that favored their choice and decreased scores (or increased them less strongly) for attributes that did not favor their choice. For the other two tasks, overall effect sizes were negligible and coherence shifting was only shown for participants who made one of the two possible choices.

Affect Scores. The evidence for coherence shifting of outcome affect ratings was stronger in Experiment 2 than Experiment 1. In Experiment 2, significant affect shifting was found for both tasks and effect sizes ranged from 0.06% to 0.12% of variance explained (see Table 9). In Study 1, effect sizes were lower and coherence shifting was not significant for the job task. This was probably due to the distractor tasks between measurement periods being inadvertently omitted in Experiment 1 but included in Experiment 2. The lack of distractor tasks in Experiment 1 may have allowed participants ratings at the mid- and post-choice measurement periods to be influenced by their memory of previous ratings.

In the serious disease tasks for both studies, participants increased all affect ratings between the pre- and mid-choice times (main effect of time), but increased outcome ratings more sharply for their chosen treatment than for their non-chosen treatment. This represents the general form of coherence shifting, where affect ratings for outcomes of the non-chosen alternative do not decrease over time but increase less than outcomes of the chosen alternative. This pattern makes sense for the serious disease

tasks, where all outcomes were highly unpleasant. The main effect of time may represent a strategy of optimistically increasing ratings for all of these unpleasant outcomes when they are experienced in the context of a decision relative to the decontextualized non-decision situation. An injection into the spinal fluid may seem less frightening when it is seen as part of an effective treatment for a serious disease than when it is contemplated alone. Participants did increase scores for their chosen treatment and decrease scores for their non-chosen treatment in the mild disease task, which showed higher outcome ratings in general than in the serious disease tasks.

Hypothesis 3, concerning the relationship between the strength of coherence shifting and the corresponding sharpness of decrease in physiological arousal post-shifting (aka a dose-response relationship), was not significantly supported in a pilot sample. This will be examined further in an upcoming publication.

Between-Task and Exploratory Analysis

In Hypothesis 4, I predicted that tasks that had more unpleasant outcomes and threatened higher level goals would lead to stronger aversive feelings. Carpenter and colleagues (Carpenter et al., 2016; Carpenter and Niedenthal, 2017) proposed that coherence shifting may serve as a strategy to reduce these aversive feelings rather than being purely a strategy to make fast or accurate decisions. This hypothesis was supported by comparisons between tasks in Experiments 1 and 2, which found that the task with more emotionally salient outcomes (serious disease in both experiments) and threatened higher-level goals (serious disease over job selection in Study 1) produced significantly

higher aversive feelings towards the decision as a whole than the less emotionally-salient tasks (job selection, mild disease).

If coherence shifting serves to reduce aversive feelings, shifting of affect scores is likely a direct way to reduce negative affect and this could lead to different patterns of affect coherence shifting between tasks. Accordingly, I conducted an exploratory analysis comparing coherence shifting between tasks. A significant difference in overall affect shifting was found in in Experiment 1, where tasks differed in context, outcome salience, and level of goals threatened. No differences were found between Experiment 2 tasks, which shared a context and only differed in the unpleasantness of outcomes and the disease treated.

In exploratory correlational analyses of Experiment 1, job task importance shifting was positively correlated with cognitive reappraisal (ERQ), which is a more cognitively effortful emotional regulation strategy then expressive suppression (Gross & John, 2003). Shifting importance scores may represent a more abstract or explicit cognitive strategy to regulate emotion than affect score shifting, making this strategy similar to cognitive reappraisal. However, in the serious disease treatment decision in Experiment 2, affect coherence shifting was correlated positively with multiple measures of emotional expressivity from the Berkeley Expressivity Questionnaire, such that participants who express negative emotions more openly and strongly engaged in stronger coherence shifting regarding affect towards outcomes. Shifting their affect ratings to reduce aversive feelings may be a more intuitive and effective emotion regulation method for people who feel and express their emotions more strongly. The

expressivity correlations found in the serious disease task did not replicate findings by Carpenter et al. (2016) in the job task, indicating that further replication may be necessary.

Is Coherence Shifting a Bias?

These findings fit together in a clear pattern. Participants in the current study shifted their ratings of affect towards outcomes and importance of attributes during decision making to align with their eventual final choice. Their final choice was predicted well by feelings towards the two choices during decision making and poorly by affect towards outcomes and importance ratings of attributes prior to decision making. By down-weighting pre-choice affect and importance ratings and engaging in coherence shifting, participants reduced aversive feelings towards the choices they made.

Although coherence shifting may help participants feel better about their choices, it is not clear whether down-weighting initial preferences in order to construct new ones during decision making is adaptive. Coherence shifting can be characterized as people relying on initial hunches about their preferred choice rather than the actual strength of evidence. Thus, it could lead to decisions that do not align with a person's long-term goals, especially given past research that has found that shifted preferences quickly revert to baseline preferences after a simulated decision (Simon et al., 2008; Simon & Spiller, 2016). Imagine a person who is deciding between insulin injections and an oral medication for Type-2 diabetes and who, outside of the context of a treatment decision, strongly dislikes needles and is strongly motivated to reduce diabetes complications by reducing blood sugar. Because insulin is more effective at reducing blood sugar, this

person might develop an early preference for insulin during decision making, use coherence shifting to perceive injections as less negative and eventually choose insulin. This person might later find that in practice, his or her fear of needles is so strong that he or she does not comply regularly with the regular injections.

Simon (2004) points out that coherence shifting can occur when people are judging whether a factual claim is true based on evidence as well as during preference decision making. Using the example of a jury making a judgment about the guilt of a defendant, he suggests that the strict reasonable doubt standard in criminal cases may encourage jurors to engage in coherence shifting when there is conflicting evidence regarding guilt. In this context, Simon characterizes coherence shifting as a bias.

Cognitive Mechanisms of Coherence Shifting as Emotional Regulation

These studies provided mixed correlational evidence that coherence shifting may serve as a strategy for emotion regulation. If shifting serves as a method to reduce negative emotions, then coherence shifting would be evidence that people engage in biased reasoning because they are trying to satisfy multiple goals: Reaching a decision with the highest future utility while also minimizing present, short-term negative affect. Marr (1982) would characterize this level of reasoning at the abstract *computational* level, representing the goals a person is trying to accomplish by engaging in coherence shifting.

However, these results do not directly address the specific cognitive mechanisms by which coherence shifting reduces negative emotion, at Marr's *algorithmic* level (1982). Some evidence was found for a positive relationship between strength of shifting

and decision confidence in Experiment 1. If participants were using effortful cognitive processes to increase confidence or justify their decision, then it is likely that they would have engaged in more shifting of relatively abstract attribute importance ratings than fast, affective outcome ratings. However, effect sizes were stronger for affect shifting in almost all tasks. In their third study, Carpenter et al. (2016) found less coherence shifting when people were depleted of regulatory resources, disrupting emotional processing. Future research could use secondary tasks that specifically tax cognitive resources to determine if systematic or affective processes underlie coherence shifting.

Alternatively, participants may attempt to reduce negative emotion by reducing perceived risk. Risk is usually not considered in models of multi-attribute decision-making, which generally treat outcomes as deterministic (Hastie & Dawes, 2010). However, both the simulated disease tasks in this experiment and real health decisions involve some level of uncertainty. Future experiments could measure perceived risk at multiple rating times or manipulate the probability of outcomes to examine well-studied biases in risky (probabilistic) decision-making (e.g., Kahneman & Tversky, 1979).

Impact and Contributions

The experiments reported above provide multiple new contributions to the study of coherence shifting and preference construction, along with an avenue towards important applications. Coherence shifting was proposed by Simon et al. (Simon, 2004) as an example of a bi-directional interaction between choice and preferences, where an initial leaning towards a certain alternative in a difficult trade-off decision will lead to changes in perceived preferences to make that alternative more dominant. This is in the

opposite direction of many revealed-preference models, which suggest that stable long-term preferences or revealed preferences for decision information are constructed unidirectionally into choice. In these experiments, I have examined changes in both of these directions, incorporating prior research on the construction of preferences from both prior and mid-decision affect. The predictive findings in both experiments provided support for preference construction as equal or superior to revealed or stable preferences for predicting choice, and that affect-based models have much stronger fit in predicting decisions than several more cognitively-effortful attribute-based decision strategies. These findings suggest that the initial leaning which participants will shift scores to support is based on affect towards alternatives as a whole, only assessed after that have seen outcomes within a decision context.

These experiments have also replicated coherence shifting in shared health treatment decisions for a mild and a serious disease. These tasks are more emotionally salient than the job task (e.g., Simon et al., 2004; Carpenter et al., 2016) and mock-jury civil court tasks (Simon, 2004) utilized in prior coherence shifting experiments, and were designed to be specifically balanced on outcome unpleasantness rather than abstract task information. They also personally threaten higher-level goals such as health and survival, as opposed to a criminal jury who may be contemplating another person's future. Coherence shifting appears to follow different patterns depending on the emotional salience of the outcomes and the level of goals threatened by the task (e.g., comfort vs. health and survival), which supports past findings in preference construction (Luce, 1998). These findings were also replicated despite using more conservative and accurate

analyses and measures of effect size such as omnibus MANOVAS and generalized η^2 (Bakeman, 2005; Tabachnick & Fidel, 2018).

In addition, these simulated disease tasks represent a first step towards applying coherence shifting findings to an important real-world decision context. These tasks were designed to share aspects with difficult health treatment tradeoff decisions using outcomes identified from real health treatments and diseases. These decisions are increasingly shared between patients and doctors (Elwyn et al., 2010; Makoul & Clayman, 2006), so understanding the impact of processes like coherence shifting is important to understanding how patients make these decisions.

Given the suggestion earlier that coherence shifting is a cognitive bias, shared decision-making templates and tools could be designed to reduce coherence shifting engaged in by patients while still providing them with assurances that will reduce their aversive feelings and increase their confidence in the effectiveness of their selected treatment. For example, based on the correlation between affect-shifting and final decision confidence and the low tendency to utilize more effortful, attribute-based strategies, doctors could ask patients what attributes of a decision are more important to them before presenting them with the treatments in a decision context. In addition, decision tools could be designed that allow participants to engage in coherence shifting consciously and openly, indicating their initial preferences outside of a decision context and then actively and visibly changing those scores once the decision context has been presented. Such a decision tool may help patients and doctors avoid unconscious bias and

provide feedback during shifting in order to avoid decisions that do not align with long-term goals.

APPENDICES

Appendix A: Instructions and Affect Scores for Tradeoff Development Pilot

General Instructions

In the following survey, you will be asked to place items into categories. This will include one long scale and several shorter scales.

These scales are RELATIVE. This means that items only need to be higher or lower on the scale than items in the categories next to them.

The directions WILL BE DIFFERENT for each question. Please read them each time.

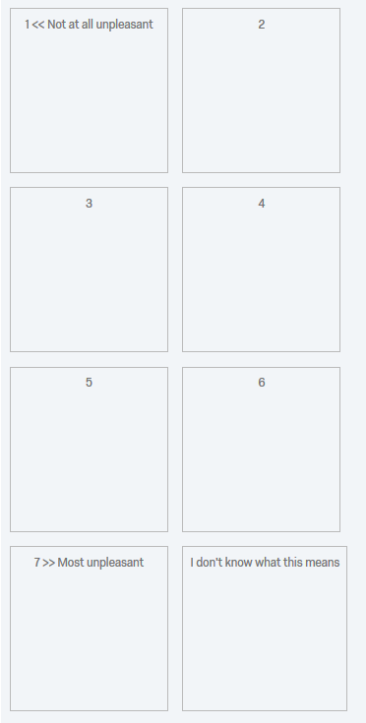

Figure A1: *Pilot rating scales. "I don't know what this means" was not present in all scales.*

Scale 1: Side Effects

Instructions: “The following items represent possible side effects for treatments for a disease. Please drag these side effects into one of 7 categories based on your immediate emotional reaction to them. Scores closer to 7 mean more unpleasant. If you don't know the meaning of an item, please place it in the "I don't know what this means" group.”

Original Scale Options: No side effects, Cough, Runny nose, Chills, Bruising, Congestion, Itching, Loss of appetite, Sore throat, Skin blemishes, Sweats, Indigestion, Blisters, Heartburn, Muscle cramping, Muscle spasms, Rash, Dizziness, Headache, Lack of coordination, Dehydration, Puffy face, Scarring, Bronchitis, Nausea, Tremors, Blood in stool, Cyst, High fever, Paranoia, Weakness in limbs, Prolonged double vision, Temporary hearing loss, Bone marrow loss, Tuberculosis

Scale 2: Administration of Treatment

Instructions: “The following items represent the administration methods for a variety of treatments for a disease. Please drag these items into 7 categories based on how unpleasant they would be to experience. Scores closer to 7 mean more unpleasant. If you don't know the meaning of an item, please place it in the "I don't know what this means" group.”

Original Scale Options: One pill only, One pill per day for 10 days, One pill immediately then one pill each at 6 24 and 48 hours later, One pill 3 times a day for 14 days, Ointment in nostrils 2 times per day for 5 days, 2 inhalations 2 times a day for 10 days, One pill 2 times a day for 5 days, One pill per day for 30 days, Oral rinse once per hour for 3 hours, One injection per week for 3 weeks, One intramuscular shot, then one pill 2 times a day for 7 days, One intramuscular injection, One injection in the buttocks, IV drip 4 hours a day for 5 days, One injection into the spinal fluid, Continuous IV drip for 7 days, Immediate hospitalization

Scale 3: Out-Of-Pocket Cost

Instructions: "The following items represent the out-of-pocket costs of treatments for a disease. This is the total amount you would personally need to pay, even if you have insurance. Please drag these items into 7 categories based on how you would feel if you had to pay that amount of money. Scores closer to 7 mean more unpleasant."

Original Scale Options: No cost, $7.00, $11.00, $20.00, $34.00, $60.00, $104.00, $180.00, $320.00, $550.00, $960.00, $2,900.00, $1,700.00, $8,900.00, $15,500.00, $5,100.00

Scale 4: Duration of Symptoms

Instructions: "The following items represent the time that you will still feel disease symptoms over the course of treatment. Please drag these items into 7 categories

based on how unpleasant they would be to experience. Scores closer to 7 mean more unpleasant."

Original Scale Options: 12 hours, 24 hours, 48 hours, 5 days, 7 days, 14 days, 10 days, 1 day of hospitalization, 3 weeks, 1 month, 3 days of hospitalization, 7 days of hospitalization, 2 months, 1 month of hospitalization

Scale 5: Efficacy

Instructions: "The following items represent the efficacy of a treatment, or chance you will be cured of a disease after 7 days. Please drag these items into 7 categories based on how you would feel if you were offered a treatment with this efficacy. Scores closer to 7 mean more unpleasant."

Original Scale Options: 99%, 95%, 91%, 83%, 87%, 71%, 79%, 75%, 63%, 67%, 55%, 59%, 47%, 51%, 43% (same as without treatment)

Scale 6: Mortality

Instructions: "The following items represent the mortality rate, or that people will die from a disease even if they have received a treatment.

Please drag these items into 7 categories based on how you would feel if you were offered a treatment with this mortality rate. Scores closer to 7 mean more unpleasant."

Original Scale Items: 0.00%, 0.50%, 2.00%, 3.50%, 5.00%, 6.50%, 8.00%,

9.50%, 11.00%, 14.00%, 12.50%, 15.50%, 17.00%, 18.50%, 20.0% (without treatment)

Appendix B: Predictive Models of Choice

The following regression models examine predictors of choice that are relevant to specific proposed conscious or unconscious strategies of decision making. These models are all based on simplifying assumptions and do not represent the full process of decision making proposed by each theory. This is necessary due to several statistical and methodological concerns. First, choice is only recorded at the person level. It is possible to distinguish between predictors in any give category (outcomes, attributes) at a between-subjects level, but the effect of these predictors on choice cannot be distinguished at a within-subjects level. For example, I might find that the importance of side effects is more predictive of choice across all participants than the importance of efficacy, but I cannot say that side effect importance was more predictive than efficacy for one person's final choice. This prevents modeling any processes where participants do not use all decision information, such as the lexicographic decision-making (Hastie & Dawes, 2010) or stochastic threshold models such as the diffusion decision model (Ratcliff & McKoon, 2008). Second, multiple self-report measures of affect and importance are likely too slow and explicit to capture rapid attentional or neurobiological processes like the diffusion decision model and somatic marker hypothesis (Bechara and Damasio, 2005).

This appendix includes the four primary models of interest, necessary statistical variations, and an explanation of simplifying assumptions for each model. The following models use notation for a decision between 2 alternatives (A & B) described through 8

outcomes within 4 attributes. See Table B1 below for this decision in a matrix form along with a summary of measures.

Table B1: *Decision Matrix.*

	(Measure)	Alternative A	*(Measure)*	Alternative B	*(Measure)*
Attribute 1	*(Importance$_1$)*	Outcome A1	*(Affect$_{A1}$)*	Outcome B1	*(Affect$_{B1}$)*
Attribute 2	*(Importance$_2$)*	Outcome A2	*(Affect$_{A2}$)*	Outcome B2	*(Affect$_{B2}$)*
Attribute 3	*(Importance$_3$)*	Outcome A3	*(Affect$_{A3}$)*	Outcome B3	*(Affect$_{B3}$)*
Attribute 4	*(Importance$_4$)*	Outcome A4	*(Affect$_{A4}$)*	Outcome B4	*(Affect$_{B4}$)*
	Feelings	Positivity A	*(Pos$_A$)*	Positivity B	*(Pos$_B$)*
	towards	Arousal A	*(Aro$_A$)*	Arousal B	*(Aro$_B$)*
	Alternatives	Dominance A	*(Dom$_A$)*	Dominance B	*(Dom$_B$)*

Model 1: Affect Heuristic (Outcome Based)

This model is based on risky decision research by Charpentier et al. (2016), who found that predicted happiness towards monetary outcomes rated prior to any decision predicted choice more reliably than the values themselves or values transformed according to known influences on decision-making (e.g., framing effects). In the context of this project, this corresponds to affect towards the 8 outcomes rated at the pre-choice time, producing the following model. (An aggregated version of this model, Model 1.1, will be examined under Model 2.)

Model 1.0: Only Outcomes

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 Affect_{A1} + \beta_2 Affect_{A2} + \beta_3 Affect_{A3} + \beta_4 Affect_{A4} + \beta_5 Affect_{B1} + \beta_6 Affect_{B2} + \beta_7 Affect_{B3} + \beta_8 Affect_{B4}$$

where P_A is the probability of choosing Alternative A, ($1 - P_A$) is the probability of choosing Alternative B, and each *Affect* term is affect towards an outcome in Table B1.

Model 2: Affect Heuristic (Outcomes and Alternatives)

Schlösser et al. (2013) measured positivity, arousal, and dominance towards both outcomes and alternatives. They found that a model using affect towards alternatives was strongly predictive of choice and was not significantly improved by the addition of outcomes as predictors. Additionally, they found that the relationship between outcome affect and choice was mediated by affect towards alternatives. The first two comparisons are relatively easy to replicate. First, Model 1 will be compared with a model containing only feelings towards alternatives A and B:

Model 2.0: Only Alternatives

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 Pos_A + \beta_2 Aro_A + \beta_3 Dom_A + \beta_4 Pos_B + \beta_5 Aro_B + \beta_6 Dom_B$$

where *Pos, Aro, Dom* are Positivity, Arousal, and Dominance ratings for Alternatives A or B.

Model 2.0 could be compared with a model including these 6 alternative ratings and the 8 outcome affect ratings:

Model 2.1: Outcomes and Alternatives

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_{1-4}(Affect_{A1-4}) + \beta_5 Pos_A + \beta_6 Aro_A + \beta_7 Dom_A + \beta_{8-11}(Affect_{B1-4}) + \beta_{12} Pos_B + \beta_{13} Aro_B + \beta_{14} Dom_B$$

Mediation analysis is more difficult. Including all possible outcome-by-alternative affect mediation terms would result in a model that is difficult to interpret and inappropriate given the lack of attribute- or outcome-level within-person variance in final choice. Also, positivity, arousal, and dominance are sufficiently different constructs that they are not easy to aggregate. Schlösser et al. (2013) simplified their model by conducting cluster analysis on alternative ratings and using cluster membership as predictors. To reduce the complexity of the model and still differentially examine the impact of positivity, arousal, and dominance, affect towards outcomes will be aggregated into predictors for each alternative using the following formulas:

$$AFF_A = \sum_{i=1}^{4} Affect_{Ai}$$

$$AFF_B = \sum_{i=1}^{4} Affect_{Bi}$$

A model using only these two aggregate terms as predictors will be compared with model 1 to determine if using these aggregate terms results in a loss of fit or explained variance.

Model 1.1: Only Aggregated Outcomes

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 AFF_A + \beta_2 AFF_B$$

For all tasks in both experiments, Model 1.1 showed substantially worse fit than model 1.0, making it inappropriate to examine interactions using aggregated terms. Two additional models are proposed to examine the interaction between pre-choice alternatives, but were not utilized for either experiment:

Model 2.2: Aggregated Outcomes + Alternatives

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 AFF_A + \beta_2 Pos_A + \beta_3 Aro_A + \beta_4 Dom_A + \beta_5 AFF_B + \beta_6 Pos_B + \beta_7 Aro_B + \beta_8 Dom_B$$

These models can be compared with a model aggregated outcome terms, feelings towards alternatives, and six alternative by outcome interaction terms:

Model 2.3: Outcome x Alternative Interaction

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 AFF_A + \beta_2 Pos_A + \beta_3 Aro_A + \beta_4 Dom_A + \beta_5 (Pos_A * AFF_A) + \beta_6 (Aro_A * AFF_A) + \beta_7 (Dom_A * AFF_A) + \beta_8 AFF_B + \beta_9 Pos_B + \beta_{10} Aro_B + \beta_{11} Dom_B + \beta_{12} (Pos_B * AFF_B) + \beta_{13} (Aro_B * AFF_B) + \beta_{14} (Dom_B * AFF_B)$$

According to the findings of Schlösser et al. (2013), with some theoretical support from stochastic and affective choice construction models, Model 2.3 should show the best performance and interaction terms should be significant predictors of choice.

Model 3: Attribute-Based Decision Strategies

Lexicographic decision making and elimination by aspects (see Hastie & Dawes, 2010) both rely on selecting the most important attribute before examining outcomes to make a decision. This fits with the shared decision making instructions from

Elwyn et al (2012), who instruct patients to think about what is most important to them. These decision-making strategies include a temporal sequence for prioritizing information in decision-making. A person first examines the most important attribute and then either selects the alternative with the best outcome (lexicographic) or rejects alternatives with unacceptable outcomes (elimination by aspects). Thresholds for "winning," "unacceptable," or "acceptable" are subjective and not defined here. If a choice cannot be reached through one alternative with these methods, the person then makes comparisons on the second most important alternative. First, a simplified model using only the 4 importance weights as predictors will be used to evaluate the impact of attribute importance in general.

Model 3.0: Only Importance

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 Importance_1 + \beta_2 Importance_2 + \beta_3 Importance_3 + \beta_4 Importance_4$$

where each *Importance* term represents an Attribute in Table B1.

In order to examine lexicographic decision making, a model using only the two affect outcome ratings for the attribute rated most important for each participant. If more than one attribute receives the highest importance rating, the outcome ratings for those attributes will be averaged.

Model 3.1 Lexicographic

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_1 Affect_{A:Most\ Important} + \beta_2 Affect_{B:Most\ Important}$$

Model 4: Weighted Additive Strategy

The weighted additive strategy instructs that importance weights for attributes should be multiplied by utility ratings for outcomes, and these products should be summed for alternatives to produce overall alternative scores (Hastie & Dawes, 2010), but without being instructed participants cannot be assumed to use this exact process. Instead, a model representing weighted additive will include 8 outcomes ratings, 4 attribute importance ratings, and 8 outcome-by-attribute interaction terms. (It is necessary to keep the 12 main effect terms in the model to assess the interactions.) If using importance to weight outcome affect ratings is the most common decision-making method, this model should explain the most variance. In order to examine the interaction, we must first examine a model only including main effects. This is a combination of Models 1.0 and 3.0:

Model 4.0: Importance and Outcomes

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_{1-4}(Importance_{1-4}) + \beta_{5-8}(Affect_{A1-4}) + \beta_{9-12}(Affect_{B1-4})$$

Adding the interaction terms provides the full weighted additive model (next page):

Model 4.1: Importance x Outcomes Interaction

$$\ln\left(\frac{P_A}{1-P_A}\right) = \beta_0 + \beta_{1-4}(Importance_{1-4}) + \beta_{5-8}(Affect_{A1-4})$$
$$+ \beta_{9-12}(Affect_{B1-4}) + \beta_{13}(Importance_1 * Affect_{A1})$$
$$+ \beta_{14}(Importance_1 * Affect_{B1}) + \beta_{15}(Importance_2 * Affect_{A2})$$
$$+ \beta_{16}(Importance_2 * Affect_{B2}) + \beta_{17}(Importance_3 * Affect_{A3})$$
$$+ \beta_{18}(Importance_3 * Affect_{B3}) + \beta_{19}(Importance_4 * Affect_{A4})$$
$$+ \beta_{20}(Importance_4 * Affect_{B4})$$

Appendix C: Experiment 1 Predictive Analysis Regression Coefficients

Table C1.1: *Logistic regression coefficients for Study 1 Serious disease pre- and mid-choice affect models.*

	Model 1: Affect (Charpentier)				Model 2.0: Affect (Mid-choice)				Model 2.1: Affect (Schlösser)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-0.09	0.09	0.92	[0.77,1.08]	-	-	-	-	-0.13	0.20	0.88	[0.57,1.31]
K Side Effects	-0.08	0.23	0.92	[0.58,1.43]	-	-	-	-	0.37	0.51	1.45	[0.52,4.09]
K Efficacy	-0.21	0.13	0.81	[0.62,1.03]	-	-	-	-	-0.57	0.31	0.57	[0.29,0.98]
K Duration	-0.05	0.12	0.95	[0.75,1.20]	-	-	-	-	0.33	0.26	1.39	[0.85,2.47]
M Admin	**0.19***	0.09	1.20	[1.01,1.46]	-	-	-	-	**0.57***	0.24	1.77	[1.17,3.05]
M Side Effects	0.05	0.14	1.05	[0.78,1.39]	-	-	-	-	-0.15	0.39	0.86	[0.39,1.88]
M Efficacy	0.15	0.14	1.16	[0.89,1.54]	-	-	-	-	0.54	0.32	1.71	[0.95,3.41]
M Duration	-0.09	0.08	0.92	[0.77,1.08]	-	-	-	-	**-0.51***	0.21	0.60	[0.38,0.86]
K Positivity	-	-	-	-	**1.42****	0.30	4.14	[2.45,8.18]	**2.02****	0.49	7.53	[3.36,24.02]
K Arousal	-	-	-	-	**0.66***	0.27	1.93	[1.17,3.50]	**0.77***	0.38	2.16	[1.09,5.01]
K Dominance	-	-	-	-	**0.73****	0.23	2.08	[1.37,3.38]	**0.96****	0.32	2.61	[1.49,5.39]
M Positivity	-	-	-	-	**-0.97****	0.32	0.38	[0.19,0.66]	**-1.51****	0.50	0.22	[0.07,0.50]
M Arousal	-	-	-	-	**-0.87****	0.32	0.42	[0.21,0.75]	**-0.86***	0.40	0.42	[0.17,0.87]
M Dominance	-	-	-	-	**-0.56***	0.25	0.57	[0.34,0.92]	**-0.92***	0.37	0.40	[0.17,0.78]
Constant	-0.32	0.92	0.73	[0.11,4.32]	-2.79	1.75	0.06	[0.00,1.68]	0.62	2.68	1.87	[0.01,544.14]

$N = 118$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table C1.2: *Logistic regression coefficients for Study 1 Serious disease pre-choice attribute-based models.*

	Model 3.0: Attribute-based				Model 3.1: Lexicographic			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-	-	-	-	-	-	-	-
K Side Effects	-	-	-	-	-	-	-	-
K Efficacy	-	-	-	-	-	-	-	-
K Duration	-	-	-	-	-	-	-	-
M Admin	-	-	-	-	-	-	-	-
M Side Effects	-	-	-	-	-	-	-	-
M Efficacy	-	-	-	-	-	-	-	-
M Duration	-	-	-	-	-	-	-	-
Admin Importance	-0.13	0.12	0.88	[0.69,1.10]	-	-	-	-
Side Effect Importance	0.31	0.17	1.36	[0.99,1.94]	-	-	-	-
Efficacy Importance	-0.34	0.18	0.71	[0.49,1.02]	-	-	-	-
Duration Importance	0.17	0.14	1.18	[0.91,1.55]	-	-	-	-
K Important Outcome	-	-	-	-	**-0.44****	-0.44	0.64	[0.50,0.79]
M Important Outcome	-	-	-	-	**0.46****	0.46	1.58	[1.19,2.18]
Constant	-0.32	1.61	0.72	[0.03,17.01]	-0.02	-0.02	0.98	[0.53,1.82]

$N = 118$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table C1.3: *Logistic regression coefficients for Study 1 Serious disease pre-choice weighted additive models.*

	Model 4.0: W.Add (main effects)				Model 4.1: W.Add (interaction)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-0.05	0.09	0.95	[0.80,1.14]	-0.28	0.31	0.76	[0.40,1.38]
K Side Effects	-0.19	0.25	0.82	[0.50,1.32]	-2.74	1.88	0.06	[0.00,1.78]
K Efficacy	**-0.26***	0.13	0.77	[0.59,0.99]	0.67	1.00	1.95	[0.28,15.97]
K Duration	0.00	0.12	1.00	[0.78,1.28]	**1.24***	0.62	3.46	[1.08,12.73]
M Admin	0.16	0.10	1.18	[0.97,1.44]	-0.72	0.41	0.49	[0.20,1.03]
M Side Effects	0.07	0.15	1.07	[0.78,1.43]	0.54	0.85	1.72	[0.31,9.13]
M Efficacy	0.19	0.15	1.21	[0.91,1.63]	-0.41	1.12	0.66	[0.06,5.79]
M Duration	-0.06	0.09	0.94	[0.78,1.13]	-0.49	0.42	0.61	[0.26,1.36]
Admin Importance	-0.12	0.13	0.89	[0.68,1.15]	0.25	0.24	1.29	[0.83,2.15]
Side Effect Importance	0.29	0.18	1.34	[0.96,1.92]	1.79	1.15	6.00	[0.85,72.26]
Efficacy Importance	-0.34	0.20	0.71	[0.48,1.04]	-0.20	0.39	0.82	[0.38,1.79]
Duration Importance	0.16	0.16	1.17	[0.86,1.62]	-0.53	0.40	0.59	[0.25,1.25]
K Admin x Importance	-	-	-	-	1.21	2.36	3.35	[0.03,386.12]
K Side Effects x Importance	-	-	-	-	15.39	11.97	$4.83*10^{6}$	$[0.00,3.53*10^{17}]$
K Efficacy x Importance	-	-	-	-	-4.79	5.30	0.01	[0.00,227.80]
K Duration x Importance	-	-	-	-	**-10.20***	4.74	0.00	[0.00,0.25]
M Admin x Importance	-	-	-	-	**-6.89***	3.09	0.00	[0.00,0.28]
M Side Effects x Importance	-	-	-	-	2.92	5.58	18.54	[0.00,1,187,352.30]
M Efficacy x Importance	-	-	-	-	-3.01	6.13	0.05	[0.00,6,579.19]
M Duration x Importance	-	-	-	-	-4.08	3.15	0.02	[0.00,6.91]
Constant	-0.15	1.82	0.86	[0.02,31.43]	-10.77	9.33	0.00	[0.00,267.10]

$N = 118$. * $p < .05$, ** $p < .001$. OR = Odds Ratio. W.Add = Weighted Additive.

Table C2.1: *Logistic regression coefficients for Study 1 Job task pre- and mid-choice affect models.*

	Model 1: Affect (Charpentier)				Model 2.0: Affect (Mid-choice)				Model 2.1: Affect (Schlösser)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
Splendor Salary	**-0.30***	0.14	0.74	[0.56,0.96]	-	-	-	-	-0.13	0.20	0.88	[0.57,1.31]
Splendor Office	-0.29	0.17	0.75	[0.52,1.01]	-	-	-	-	0.37	0.51	1.45	[0.52,4.09]
Splendor Vacation	**-0.22***	0.11	0.80	[0.64,0.99]	-	-	-	-	-0.57	0.31	0.57	[0.29,0.98]
Splendor Commute	-0.08	0.15	0.92	[0.68,1.24]	-	-	-	-	0.33	0.26	1.39	[0.85,2.47]
BB Salary	**0.52****	0.17	1.69	[1.23,2.41]	-	-	-	-	**0.57***	0.24	1.77	[1.17,3.05]
BB Office	0.23	0.13	1.26	[0.98,1.66]	-	-	-	-	-0.15	0.39	0.86	[0.39,1.88]
BB Vacation	0.13	0.15	1.14	[0.86,1.56]	-	-	-	-	0.54	0.32	1.71	[0.95,3.41]
BB Commute	0.25	0.13	1.28	[1.00,1.68]	-	-	-	-	**-0.51***	0.21	0.60	[0.38,0.86]
K Positivity	-	-	-	-	**1.42****	0.30	4.14	[2.45,8.18]	-	-	-	-
K Arousal	-	-	-	-	**0.66***	0.27	1.93	[1.17,3.50]	-	-	-	-
K Dominance	-	-	-	-	**0.73****	0.23	2.08	[1.37,3.38]	**2.02****	0.49	7.53	[3.36,24.02]
M Positivity	-	-	-	-	**-0.97****	0.32	0.38	[0.19,0.66]	**0.77***	0.38	2.16	[1.09,5.01]
M Arousal	-	-	-	-	**-0.87****	0.32	0.42	[0.21,0.75]	**0.96****	0.32	2.61	[1.49,5.39]
M Dominance	-	-	-	-	**-0.56***	0.25	0.57	[0.34,0.92]	**-1.51****	0.50	0.22	[0.07,0.50]
Constant	-0.76	0.82	0.47	[0.09,2.31]	-2.79	1.75	0.06	[0.00,1.68]	**-0.86***	0.40	0.42	[0.17,0.87]

$N = 118$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table C2.2: *Logistic regression coefficients for Study 1 Job task pre-choice attribute-based models.*

	Model 3.0: Attribute-based				Model 3.1: Lexicographic			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
Splendor Salary	-	-	-	-	-	-	-	-
Splendor Office	-	-	-	-	-	-	-	-
Splendor Vacation	-	-	-	-	-	-	-	-
Splendor Commute	-	-	-	-	-	-	-	-
BB Salary	-	-	-	-	-	-	-	-
BB Office	-	-	-	-	-	-	-	-
BB Vacation	-	-	-	-	-	-	-	-
BB Commute	-	-	-	-	-	-	-	-
Salary Importance	-0.03	0.20	0.97	[0.66,1.44]	-	-	-	-
Office Importance	-0.12	0.15	0.89	[0.65,1.20]	-	-	-	-
Vacation Importance	**0.31***	0.15	1.37	[1.02,1.87]	-	-	-	-
Commute Importance	0.05	0.13	1.05	[0.83,1.35]	-	-	-	-
Splendor Important Outcome	-	-	-	-	-0.10	-0.10	0.90	[0.74,1.10]
BB Important Outcome	-	-	-	-	0.07	0.07	1.07	[0.89,1.30]
Constant	-1.79	1.77	0.17	[0.00,5.18]	**-0.71****	-0.71	0.49	[0.28,0.82]

$N = 112$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table C2.3: *Logistic regression coefficients for Study 1 Job task pre-choice weighted additive models.*

	Model 4.0: W.Add (main effects)				Model 4.1: W.Add (interaction)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
Splendor Salary	**-0.35***	0.15	0.70	[0.51,0.93]	-2.15	2.09	0.12	[0.00,6.58]
Splendor Office	-0.28	0.17	0.76	[0.52,1.05]	-2.15	2.09	0.12	[0.00,6.58]
Splendor Vacation	-0.20	0.11	0.82	[0.65,1.01]	-2.15	2.09	0.12	[0.00,6.58]
Splendor Commute	-0.14	0.16	0.87	[0.64,1.19]	-2.15	2.09	0.12	[0.00,6.58]
BB Salary	**0.60****	0.19	1.82	[1.29,2.73]	-2.15	2.09	0.12	[0.00,6.58]
BB Office	0.20	0.14	1.22	[0.94,1.62]	-2.15	2.09	0.12	[0.00,6.58]
BB Vacation	0.07	0.16	1.08	[0.79,1.49]	-2.15	2.09	0.12	[0.00,6.58]
BB Commute	**0.34***	0.17	1.40	[1.02,2.00]	-2.15	2.09	0.12	[0.00,6.58]
Salary Importance	-0.11	0.24	0.89	[0.55,1.44]	-2.15	2.09	0.12	[0.00,6.58]
Office Importance	-0.09	0.18	0.92	[0.64,1.32]	-2.15	2.09	0.12	[0.00,6.58]
Vacation Importance	0.18	0.18	1.20	[0.84,1.73]	-2.15	2.09	0.12	[0.00,6.58]
Commute Importance	0.30	0.18	1.35	[0.96,1.97]	-2.15	2.09	0.12	[0.00,6.58]
Splendor Salary x Importance	-	-	-	-	1.69	7.15	5.41	[0.00,4.83x10^{6}]
Splendor Office x Importance	-	-	-	-	8.35	5.58	4,249.12	[0.10,4.62x10^{8}]
Splendor Vacation x Importance	-	-	-	-	2.62	3.55	13.78	[0.01,15,120.54]
Splendor Commute x Importance	-	-	-	-	**-13.25****	5.12	0.00	[0.00,0.02]
BB Salary x Importance	-	-	-	-	-3.10	7.89	0.04	[0.00,3.61x10^{5}]
BB Office x Importance	-	-	-	-	0.86	4.13	2.37	[0.00,9,129.14]
BB Vacation x Importance	-	-	-	-	-4.83	3.83	0.01	[0.00,13.23]
BB Commute x Importance	-	-	-	-	1.03	3.65	2.81	[0.00,5,338.29]
Constant	-2.15	2.09	0.12	[0.00,6.58]	-0.05	6.37	0.95	[0.00,2.23x10^{5}]

N = 112. * $p < .05$, ** $p < .001$. OR = Odds Ratio. W.Add = Weighted Additive.

Appendix D: Experiment 1 (Serious Disease and Job) Coherence Shifting and Exploratory Statistics.

Table D1.1: *Coherence shifting of serious disease importance scores by choice, time, and favored treatment: MANOVA results.*

Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
3-way (Choice by Time by Favored)	0.066	**4.11***	2, 116	0.019	0.016
2-way (Choice by Time)	0.154	**10.52****	2, 116	<.001	0.042
2-way (Choice by Favored)	0.043	2.59	2, 116	0.079	0.010
2-way (Time by Favored)	0.060	**3.73***	2, 116	0.027	0.015
Choice (Main Effect)	0.001	0.04	2, 116	0.962	0.000
Time (Main Effect)	0.572	**77.43****	2, 116	<.001	0.237
Favored Treatment (Main Effect)	0.572	**77.43****	2, 116	<.001	0.026

$N = 119$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table D1.2: *Coherence shifting of serious disease importance scores: Simple-effects univariate tests.*

		2-way (Time by Favored)				Time				Treatment Favored (K vs M)			
Univariate	*N*	*F*	*Df*	*p*	η_G^2	*F*	*Df*	*p*	η_G^2	*F*	*df*	*p*	η_G^2
Chose K	77	2.12	1, 76	0.150	0.014	0.035	1, 76	0.852	0.000	**28.87****	1, 76	<.001	0.160
Chose M	42	**5.40***	1, 41	0.025	0.062	**4.60***	1, 41	0.038	0.053	2.88	1, 41	0.097	0.034

$N = 119$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table D2.1: *Coherence shifting of serious disease affect scores by choice, time, and treatment: Interactions.*

	3-way (Choice by Time by Treatment)					2-way (Choice by Time)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
Affect	0.089	**2.89***	4, 119	0.023	0.023	0.135	**4.65***	4, 119	0.002	0.036
Univariate		*F*	*Df*	*p*	η_G^2		*F*	*Df*	*p*	η_G^2
Admin		1.43	1, 122	0.233	1.43	0.002	**10.76***	1, 122	0.001	0.042
Side Effects		1.22	1, 122	0.272	1.22	0.002	0.07	1, 122	0.786	0.000
Efficacy		0.12	1, 122	0.731	0.12	0.000	**5.33***	1, 122	0.023	0.011
Duration		**9.20***	1, 122	0.003	9.20	0.015	3.20	1, 122	0.076	0.011

	2-way (Choice by Treatment)					2-way (Time by Treatment)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*		*Pillai's Trace*	*F*	*Df*	*P*	η_G^2
Affect	0.039	1.21	4, 119	0.312	0.009	0.454	**24.74****	4, 119	<.001	0.162
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Admin		0.00	1, 122	0.967	0.000		**5.10***	1, 122	0.026	0.006
Side Effects		0.86	1, 122	0.356	0.001		**22.06****	1, 122	<.001	0.033
Efficacy		2.22	1, 122	0.139	0.005		**90.22****	1, 122	<.001	0.068
Duration		1.90	1, 122	0.171	0.003		0.40	1, 122	0.526	0.001

$N = 124$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared, K = Treatment K, M = Treatment M. See table 3 for specific outcomes.

Table D2.2: *Coherence shifting of serious disease affect scores by choice, time, and treatment: Main effects.*

	Choice					Time				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*P*	η_G^2
Affect	0.004	0.45	4, 119	0.770	0.004	0.819	**134.81****	4, 119	<.001	0.513
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Admin		0.00	1, 122	0.990	0.000		**73.75****	1, 122	<.001	0.224
Side Effects		1.11	1, 122	0.294	0.002		**161.86****	1, 122	<.001	0.366
Efficacy		0.75	1, 122	0.389	0.002		**362.78****	1, 122	<.001	0.427
Duration		0.55	1, 122	0.461	0.001		**23.69****	1, 122	<.001	0.078

Multivariate	Treatment (K vs M)				
Affect	*Pillai's Trace*	*F*	*df*	*p*	η_G^2
Univariate	0.353	**16.26****	4, 119	<.001	0.113
Admin		*F*	*Df*	*p*	η_G^2
Side Effects		**5.64***	1, 122	0.019	0.008
Efficacy		**25.86****	1, 122	<.001	0.042
Duration		**6.59***	1, 122	0.011	0.016
Multivariate		**41.28***	1, 122	<.001	0.062

$N = 124$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared, K = Treatment K, M = Treatment M. See Table 3 for specific outcomes.

Table D3: *Coherence shifting of job task importance scores by choice, time, and job favored: MANOVA results.*

Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
3-way (Choice by Time by Favored)	0.222	**16.36****	2, 115	<.001	0.064
2-way (Choice by Time)	0.110	**7.13***	2, 115	0.001	0.028
2-way (Choice by Favored)	0.017	0.99	2, 115	0.374	0.004
2-way (Time by Favored)	0.213	**15.57****	2, 115	<.001	0.057
Choice (Main Effect)	0.032	1.93	2, 115	0.150	0.008
Time (Main Effect)	0.612	**90.70****	2, 115	<.001	0.261
Job Favored (Main Effect)	0.009	0.52	2, 115	0.596	0.002

$N = 118$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table D4.1: *Coherence shifting of job task affect scores by choice, time, and job: Interactions.*

	3-way (Choice by Time by Job)					2-way (Choice by Time)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*P*	η_G^2
Affect	0.073	2.22	4, 113	0.072	0.017	0.298	**12.00****	4, 113	<.001	0.092
Univariate		*F*	*Df*	*p*	η_G^2		*F*	*Df*	*p*	η_G^2
Salary		0.37	1, 116	0.545	0.001		**14.90****	1, 116	<.001	0.058
Office		**5.53***	1, 116	0.020	0.006		**14.24****	1, 116	<.001	0.071
Vacation		1.21	1, 116	0.273	0.001		**4.85***	1, 116	0.030	0.023
Commute		2.54	1, 116	0.114	0.003		**6.05***	1, 116	0.015	0.026

	2-way (Choice by Job)					2-way (Time by Job)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*P*	η_G^2
Affect	0.057	1.69	4, 113	0.156	0.013	0.299	**12.06****	4, 113	<.001	0.084
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Salary		0.67	1, 116	0.414	0.001		**14.90****	1, 116	<.001	0.020
Office		3.74	1, 116	0.056	0.004		**5.85***	1, 116	0.017	0.007
Vacation		0.58	1, 116	0.446	0.001		**4.72***	1, 116	0.032	0.004
Commute		0.60	1, 116	0.440	0.001		**22.98****	1, 116	<.001	0.026

$N = 118$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared. Spl. = Splendor, BB = Bonnie's Best

Table D4.2: *Coherence shifting of job task affect scores by choice, time, and job: Main effects.*

	Choice					Time				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
Affect	0.068	2.06	4, 113	0.091	0.016	0.938	**427.01****	4, 113	<.001	0.765
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*p*	η_G^2
Admin		0.08	1, 116	0.783	0.000		**873.50****	1, 116	<.001	0.773
Side Effects		2.23	1, 116	0.138	0.002		**627.56****	1, 116	<.001	0.757
Efficacy		0.01	1, 116	0.910	0.000		**183.18****	1, 116	<.001	0.468
Duration		**5.74***	1, 116	0.018	0.010		**873.50****	1, 116	<.001	0.716

	Job (Spl vs BB)				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2
Affect	0.176	**6.04****	4, 113	<.001	0.044
Univariate		*F*	*Df*	*p*	η_G^2
Admin		0.05	1, 116	0.829	0.000
Side Effects		**16.39****	1, 116	<.001	0.015
Efficacy		0.18	1, 116	0.675	0.000
Duration		**4.36***	1, 116	0.039	0.005

$N = 118$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared. Spl. = Splendor, BB = Bonnie's Best

Table D5: *Aversive feelings and overall strength of Coherence Shifting in the Serious Disease Task* and the Job Task.

	Serious Disease Task	Job Task				
Feeling	*M (SD)*	*M (SD)*	*r (Between-Task)*	*Paired-samples t*	*df*	*P*
Anxious	5.90 (2.00)	4.67 (1.84)	**0.371****	**5.85****	105	< .001
Stressed	5.98 (2.25)	4.81 (1.90)	**0.476****	**5.62****	105	< .001
Unpleasant	6.12 (2.10)	3.92 (1.80)	**0.338****	**10.07****	105	< .001
Conflicted	5.59 (2.40)	4.91 (1.88)	0.165	**2.54***	105	0.012
Coherence Shifting						
Importance	0.052 (0.048)	0.066 (0.069)	0.117	-1.91	105	0.058
Affect	0.110 (0.092)	0.076 (0.056)	-0.064	**3.20***	105	0.002

$N = 106$. * $p < .05$, ** $p < .001$.

Table D6: Descriptive statistics and correlations for Experiment 1 between-task and exploratory analysis.

	M (SD)	1	2	3	4	5	6	7	8	9
1 Serious Imp. ACS	0.05 (0.05)									
2 Serious Affect ACS	0.11 (0.09)	.006								
3 Job Imp. ACS	0.07 (0.07)	.117	.104							
4 Job Affect ACS	0.08 (0.06)	**-.238***	-.064	.112						
5 BEQ Negative Emotionality	22.27 (6.06)	-.019	-.049	.071	.100					
6 BEQ Positive Emotionality	21.15 (3.72)	.123	.190	.009	.095	**.584****				
7 BEQ Impulse Strength	26.75 (8.08)	.060	-.058	-.063	-.083	**.468****	.339**			
8 BEQ Overall	70.17 (14.43)	.057	-.004	-.003	.020	**.832****	.692**	**.844****		
9 ERQ Cognitive Reappraisal	29.45 (5.12)	.007	**.227***	.119	-.135	.020	.346**	-.128	.026	
10 ERQ Expressive Suppression	14.12 (4.6)	.049	-.082	-.063	-.117	**-.696****	-.678**	**-.311****	**-.641****	**-.213***
11 Serious Choice Confidence	3.67 (0.81)	-.009	**.304****	.150	-.009	-.088	-.031	-.061	-.079	.039
12 Job Choice Confidence	4.15 (1.04)	-.080	.155	**.208***	.022	-.008	-.092	-.108	-.087	.017
13 Serious Confidence Change	0.41 (0.63)	-.118	.104	.135	.043	.078	.181	-.081	.034	.064
14 Job Confidence Change	0.56 (0.81)	-.095	.105	.175	.125	.173	.150	-.012	.105	.153

	10	11	12	13
11 Serious Choice Confidence	.021			
12 Job Choice Confidence	.082	.071		
13 Serious Confidence Change	-.109	**.302****	.095	
14 Job Confidence Change	-.157	-.008	**.615****	.170

N = 106. * $p < .05$, ** $p < .001$, ACS = Absolute Coherence Shifting, Imp. = Importance, BEQ = Berkeley Expressivity Questionnaire, ERQ = Emotion Regulation Questionnaire.

Appendix E: Experiment 2 Predictive Analysis Regression Coefficients

Table E1.1: *Logistic regression coefficients for Study 2 Mild disease pre- and mid-choice affect models.*

	Model 1: Affect (Charpentier)				Model 2.0: Affect (Mid-choice)				Model 2.1: Affect (Schlösser)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
T Admin	-0.15	0.15	0.86	[0.64,1.15]	-	-	-	-	-0.12	0.26	0.89	[0.51,1.46]
T Side Effects	-0.32	0.26	0.73	[0.42,1.16]	-	-	-	-	0.30	0.48	1.36	[0.50,3.51]
T Efficacy	0.03	0.25	1.03	[0.63,1.73]	-	-	-	-	0.40	0.45	1.49	[0.54,3.74]
T Duration	-0.55	0.29	0.58	[0.32,1.01]	-	-	-	-	-1.20	0.65	0.30	[0.07,0.95]
N Admin	0.13	0.12	1.14	[0.90,1.47]	-	-	-	-	0.45	0.29	1.56	[0.94,3.10]
N Side Effects	0.19	0.21	1.21	[0.82,1.84]	-	-	-	-	-0.06	0.37	0.94	[0.44,1.96]
N Efficacy	0.26	0.19	1.30	[0.90,1.92]	-	-	-	-	0.62	0.35	1.86	[1.00,4.10]
N Duration	0.10	0.12	1.10	[0.87,1.39]	-	-	-	-	0.24	0.23	1.27	[0.82,2.08]
T Positivity	-	-	-	-	**0.85****	0.30	2.35	[1.37,4.55]	**1.29****	0.47	3.64	[1.65,11.04]
T Arousal	-	-	-	-	0.34	0.33	1.40	[0.76,2.78]	0.18	0.37	1.20	[0.60,2.65]
T Dominance	-	-	-	-	0.32	0.24	1.37	[0.87,2.26]	0.72*	0.36	2.05	[1.05,4.48]
N Positivity	-	-	-	-	**-1.05****	0.34	0.35	[0.17,0.63]	**-1.06***	0.43	0.35	[0.13,0.72]
N Arousal	-	-	-	-	0.23	0.31	1.26	[0.69,2.34]	0.55	0.45	1.74	[0.72,4.61]
N Dominance	-	-	-	-	-0.31	0.23	0.73	[0.46,1.13]	**-0.94***	0.47	0.39	[0.13,0.87]
Constant	-1.92	1.22	0.15	[0.01,1.52]	-0.96	1.97	0.38	[0.01,18.02]	-7.71*	3.89	0.00	[0.00,0.41]

$N = 90$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table E1.2: *Logistic regression coefficients for Study 2 Mild disease pre-choice attribute-based models.*

	Model 3.0: Attribute-based				Model 3.1: Lexicographic			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
T Admin	-	-	-	-	-	-	-	-
T Side Effects	-	-	-	-	-	-	-	-
T Efficacy	-	-	-	-	-	-	-	-
T Duration	-	-	-	-	-	-	-	-
N Admin	-	-	-	-	-	-	-	-
N Side Effects	-	-	-	-	-	-	-	-
N Efficacy	-	-	-	-	-	-	-	-
N Duration	-	-	-	-	-	-	-	-
Admin Importance	-0.22	0.15	0.81	[0.59,1.08]	-	-	-	-
Side Effect Importance	-0.13	0.21	0.88	[0.57,1.33]	-	-	-	-
Efficacy Importance	-0.13	0.19	0.87	[0.59,1.27]	-	-	-	-
Duration Importance	**0.63****	0.21	1.87	[1.26,2.95]	-	-	-	-
T Important Outcome	-	-	-	-	**-0.18***	-0.18	0.83	[0.69,0.97]
N Important Outcome	-	-	-	-	0.20	0.20	1.22	[0.99,1.54]
Constant	0.09	2.34	1.10	[0.01,120.49]	**1.20****	1.20	3.31	[1.98,5.98]

$N = 90$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table E1.3: *Logistic regression coefficients for Study 2 Mild disease pre-choice Weighted Additive (W.Add) models.*

	Model 4.0: W.Add (main effects)				Model 4.1: W.Add (interaction)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
T Admin	-0.16	0.16	0.85	[0.61,1.16]	0.55	0.57	1.73	[0.58,5.76]
T Side Effects	-0.38	0.30	0.68	[0.36,1.15]	0.60	1.25	1.82	[0.15,23.11]
T Efficacy	0.11	0.28	1.12	[0.64,1.95]	1.64	1.53	5.16	[0.32,154.42]
T Duration	-0.35	0.31	0.70	[0.38,1.28]	1.03	1.78	2.81	[0.08,98.78]
N Admin	0.14	0.14	1.15	[0.88,1.54]	0.15	0.39	1.16	[0.56,2.71]
N Side Effects	0.20	0.23	1.23	[0.79,1.96]	0.06	1.12	1.06	[0.13,10.28]
N Efficacy	0.19	0.20	1.21	[0.82,1.82]	0.39	0.97	1.48	[0.22,11.19]
N Duration	0.15	0.13	1.17	[0.91,1.52]	0.29	0.61	1.34	[0.42,4.97]
Admin Importance	-0.16	0.18	0.86	[0.58,1.22]	-0.13	0.25	0.88	[0.51,1.42]
Side Effect Importance	-0.25	0.26	0.78	[0.46,1.27]	-0.83	0.78	0.44	[0.08,1.94]
Efficacy Importance	-0.18	0.22	0.84	[0.53,1.29]	0.45	0.71	1.57	[0.44,7.89]
Duration Importance	**0.66***	0.26	1.93	[1.19,3.37]	-0.23	1.34	0.79	[0.06,12.69]
T Admin x Importance	-	-	-	-	-5.85	4.59	0.00	[0.00,13.08]
T Side Effects x Importance	-	-	-	-	-8.01	10.43	0.00	[0.00,1.12*e^5]
T Efficacy x Importance	-	-	-	-	-8.91	9.36	0.00	[0.00,2.13*e^3]
T Duration x Importance	-	-	-	-	-10.31	13.38	0.00	[0.00,9.69*e6]
N Admin x Importance	-	-	-	-	0.25	3.38	1.29	[0.00,1.85*e^3]
N Side Effects x Importance	-	-	-	-	-1.57	9.06	0.21	[0.00,8.62*e^6]
N Efficacy x Importance	-	-	-	-	1.18	5.77	3.24	[0.00,4.81*e^5]
N Duration x Importance	-	-	-	-	1.32	4.84	3.73	[0.00,1.22*e^5]
Constant	-1.92	2.92	0.15	[0.00,46.84]	2.08	10.45	8.01	[0.00,8.52*e^9]

$N = 90$. * $p < .05$, ** $p < .001$. OR = Odds Ratio. W.Add = Weighted Additive. Extreme values are in scientific notation.

Table E2.1: *Logistic regression coefficients for Study 2 Serious disease pre- and mid-choice affect models.*

	Model 1: Affect (Charpentier)				Model 2.0: Affect (Mid-choice)				Model 2.1: Affect (Schlösser)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-0.06	0.10	0.94	[0.76,1.15]	-	-	-	-	-0.27	0.26	0.76	[0.43,1.28]
K Side Effects	-0.35	0.32	0.71	[0.37,1.30]	-	-	-	-	1.14	0.65	3.13	[0.96,13.82]
K Efficacy	0.06	0.15	1.06	[0.79,1.41]	-	-	-	-	-0.14	0.38	0.87	[0.38,1.76]
K Duration	-0.21	0.13	0.81	[0.62,1.03]	-	-	-	-	0.31	0.28	1.37	[0.79,2.50]
M Admin	0.26	0.16	1.29	[0.96,1.80]	-	-	-	-	-0.15	0.30	0.86	[0.45,1.58]
M Side Effects	0.19	0.19	1.22	[0.84,1.79]	-	-	-	-	**-1.27***	0.63	0.28	[0.06,0.80]
M Efficacy	0.00	0.14	1.00	[0.76,1.30]	-	-	-	-	0.52	0.44	1.67	[0.76,4.64]
M Duration	0.10	0.17	1.11	[0.80,1.55]	-	-	-	-	0.04	0.35	1.05	[0.52,2.15]
K Positivity	-	-	-	-	**1.40****	0.37	4.07	[2.15,9.53]	**2.19****	0.65	8.92	[3.20,47.54]
K Arousal	-	-	-	-	0.38	0.35	1.46	[0.74,3.05]	0.80	0.58	2.22	[0.76,8.26]
K Dominance	-	-	-	-	0.13	0.26	1.14	[0.68,1.92]	0.45	0.36	1.57	[0.82,3.57]
M Positivity	-	-	-	-	**-1.01****	0.30	0.37	[0.19,0.62]	**-1.91****	0.62	0.15	[0.03,0.40]
M Arousal	-	-	-	-	-0.24	0.35	0.79	[0.38,1.56]	-0.48	0.47	0.62	[0.22,1.53]
M Dominance	-	-	-	-	**-0.76***	0.30	0.47	[0.24,0.80]	**-1.34****	0.51	0.26	[0.07,0.60]
Constant	-0.36	1.24	0.70	[0.06,8.22]	0.76	2.05	2.14	[0.04,146.29]	4.25	3.44	70.25	[0.12,1.55•e^5]

N = 88. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table E2.2: *Logistic regression coefficients for Study 2 Serious disease pre-choice attribute-based models.*

	Model 3.0: Attribute-based				Model 3.1: Lexicographic			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-	-	-	-	-	-	-	-
K Side Effects	-	-	-	-	-	-	-	-
K Efficacy	-	-	-	-	-	-	-	-
K Duration	-	-	-	-	-	-	-	-
M Admin	-	-	-	-	-	-	-	-
M Side Effects	-	-	-	-	-	-	-	-
M Efficacy	-	-	-	-	-	-	-	-
M Duration	-	-	-	-	-	-	-	-
Admin Importance	-0.08	0.15	0.92	[0.68,1.24]	-	-	-	-
Side Effect Importance	-0.11	0.17	0.90	[0.63,1.26]	-	-	-	-
Efficacy Importance	**-0.40***	0.18	0.67	[0.46,0.94]	-	-	-	-
Duration Importance	0.12	0.15	1.13	[0.84,1.54]	-	-	-	-
K Important Outcome	-	-	-	-	-0.09	-0.09	0.91	[0.76,1.10]
M Important Outcome	-	-	-	-	-0.10	-0.10	0.91	[0.71,1.15]
Constant	3.40	1.90	29.89	[0.80,1,437.88]	-0.16	-0.16	0.85	[0.45,1.57]

$N = 88$. * $p < .05$, ** $p < .001$. OR = Odds Ratio.

Table E2.3: *Logistic regression coefficients for Study 2 Serious disease pre-choice Weighted Additive (W.Add) models.*

	Model 4.0: W.Add (main effects)				Model 4.1: W.Add (interaction)			
Predictors	*B*	*SE B*	*OR*	*95%CI OR*	*B*	*SE B*	*OR*	*95%CI OR*
K Admin	-0.05	0.11	0.95	[0.76,1.18]	0.30	0.46	1.35	[0.55,3.42]
K Side Effects	-0.51	0.35	0.60	[0.29,1.18]	-0.11	1.75	0.89	[0.03,28.54]
K Efficacy	0.16	0.16	1.17	[0.85,1.63]	-0.62	0.86	0.54	[0.09,2.83]
K Duration	-0.16	0.14	0.86	[0.64,1.11]	-1.28	0.67	0.28	[0.06,0.86]
M Admin	0.23	0.16	1.26	[0.93,1.77]	0.21	0.43	1.23	[0.52,2.95]
M Side Effects	0.16	0.22	1.17	[0.77,1.82]	-0.80	0.93	0.45	[0.07,2.91]
M Efficacy	-0.06	0.15	0.94	[0.69,1.26]	1.16	0.75	3.18	[0.78,16.34]
M Duration	0.19	0.20	1.21	[0.83,1.81]	0.50	0.71	1.65	[0.42,6.97]
Admin Importance	-0.18	0.20	0.84	[0.56,1.23]	-0.26	0.35	0.77	[0.38,1.56]
Side Effect Importance	-0.04	0.19	0.96	[0.65,1.40]	0.05	1.32	1.05	[0.08,15.06]
Efficacy Importance	**-0.44***	0.20	0.64	[0.43,0.93]	**-0.99****	0.37	0.37	[0.17,0.72]
Duration Importance	0.06	0.17	1.07	[0.76,1.51]	0.86	0.62	2.37	[0.73,8.65]
K Admin x Importance	-	-	-	-	-3.12	3.58	0.04	[0.00,44.69]
K Side Effects x Importance	-	-	-	-	-4.94	14.51	0.01	[0.00,2.48•e^{10}]
K Efficacy x Importance	-	-	-	-	5.10	5.30	163.84	[0.01,1.01•e^{7}]
K Duration x Importance	-	-	-	-	9.23	4.89	10,208.62	[1.78,4.24•e^{8}]
M Admin x Importance	-	-	-	-	0.26	3.37	1.29	[0.00,1.07•e^{3}]
M Side Effects x Importance	-	-	-	-	-8.54	7.39	0.00	[0.00,408.32]
M Efficacy x Importance	-	-	-	-	7.79	4.66	2,416.59	[0.41,6.36•e^{7}]
M Duration x Importance	-	-	-	-	2.32	4.92	10.17	[0.00,2.09•e^{5}]
Constant	3.14	2.27	23.09	[0.30,2,383.15]	1.62	9.40	5.04	[0.00,7.71•e^{8}]

N = 88. * $p < .05$, ** $p < .001$. OR = Odds Ratio. W.Add = Weighted Additive. Extreme values are in scientific notation.

Appendix F: Experiment 2 (Mild and Serious Disease) Coherence Shifting, Skin Conductance, and Exploratory Statistics

Table F1.1: *Coherence shifting of Study 2 mild disease importance scores by choice, time, and favored treatment: MANOVA results*

Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
3-way (Choice by Time by Favored)	0.075	**3.55***	2, 88	0.033	0.018
2-way (Choice by Time)	0.269	**16.18****	2, 88	<.001	0.081
2-way (Choice by Favored)	0.025	1.11	2, 88	0.335	0.006
2-way (Time by Favored)	0.086	**4.14***	2, 88	0.019	0.020
Choice (Main Effect)	0.066	3.10	2, 88	0.050	0.016
Time (Main Effect)	0.226	**12.88****	2, 88	<.001	0.061
Treatment Favored (Main Effect)	0.078	**3.71***	2, 88	0.028	0.018

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F1.2: *Coherence shifting of mild disease importance scores: Simple-effects univariate tests.*

		2-way (Time by Favored)				Time				Treatment Favored (K vs M)			
Univariate	*N*	*F*	*Df*	*p*	η_G^2	*F*	*Df*	*p*	η_G^2	*F*	*df*	*p*	η_G^2
Chose T	24	0.01	1, 23	0.946	0.000	2.42	1, 23	0.133	0.050	**8.95***	1, 23	0.007	0.163
Chose N	67	**15.52****	1, 66	<.001	0.105	0.66	1, 66	0.419	0.005	**23.51****	1, 66	<.001	0.151

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F2.1: *Coherence shifting of Study 2 mild disease affect scores by choice, time, and treatment: Interactions.*

	3-way (Choice by Time by Treatment)					2-way (Choice by Time)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
Affect	0.226	**6.29****	4, 86	<.001	0.060	0.317	**9.97****	4, 86	<.001	0.095
Univariate		*F*	*Df*	*p*	η_G^2		*F*	*Df*	*p*	η_G^2
Admin		3.55	1, 89	0.063	0.005		**8.36***	1, 89	0.005	0.049
Side Effects		1.23	1, 89	0.270	0.002		3.31	1, 89	0.072	0.018
Efficacy		**6.41***	1, 89	0.013	0.010		**4.19***	1, 89	0.044	0.016
Duration		**6.28***	1, 89	0.014	0.013		**15.29****	1, 89	<.001	0.069

	2-way (Choice by Treatment)					2-way (Time by Treatment)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	P	η_G^2
Affect	0.022	0.49	4, 86	0.743	0.005	0.022	**9.52****	4, 86	<.001	0.083
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Admin		0.37	1, 89	0.544	0.001		**4.52***	1, 89	0.036	0.007
Side Effects		0.19	1, 89	0.662	0.000		3.37	1, 89	0.070	0.006
Efficacy		1.24	1, 89	0.269	0.002		**19.65****	1, 89	<.001	0.031
Duration		0.21	1, 89	0.648	0.001		**17.18****	1, 89	<.001	0.033

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F2.2: *Coherence shifting of Study 2 mild disease affect scores by choice, time, and treatment: Main effects.*

	Choice					Time				
Multivariate	*Pillai's Trace*	*F*	*df*	p	η_G^2	*Pillai's Trace*	*F*	*Df*	p	η_G^2
Affect	0.096	2.28	4, 86	0.067	0.022	0.907	**209.44****	4, 86	<.001	0.666
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Admin		0.00	1, 89	0.957	0.000		**15.93****	1, 89	<.001	0.085
Side Effects		0.98	1, 89	0.324	0.002		**40.5****	1, 89	<.001	0.177
Efficacy		**7.78***	1, 89	0.006	0.026		**273.25****	1, 89	<.001	0.505
Duration		0.16	1, 89	0.693	0.000		**725.23****	1, 89	<.001	0.767

	Treatment (T vs N)				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2
Affect	0.307	**9.52****	4, 86	<.001	0.083
Univariate		*F*	*Df*	*p*	η_G^2
Admin		1.06	1, 89	0.307	0.002
Side Effects		0.57	1, 89	0.451	0.001
Efficacy		2.77	1, 89	0.100	0.005
Duration		**33.44****	1, 89	<.001	0.078

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F3: *Coherence shifting of Study 2 serious disease importance scores by choice, time, and favored treatment: MANOVA results.*

Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
3-way (Choice by Time by Favored)	0.098	**4.72***	2, 87	0.011	0.025
2-way (Choice by Time)	0.139	**7.03***	2, 87	0.002	0.037
2-way (Choice by Favored)	0.042	1.89	2, 87	0.157	0.046
2-way (Time by Favored)	0.054	2.48	2, 87	0.090	0.010
Choice (Main Effect)	0.054	2.49	2, 87	0.089	0.013
Time (Main Effect)	0.054	2.49	2, 87	0.089	0.013
Treatment Favored (Main Effect)	0.174	**9.14****	2, 87	<.001	0.046

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F4.1: *Coherence shifting of Study 2 serious disease affect scores by choice, time, and treatment: Interactions.*

	3-way (Choice by Time by Treatment)					2-way (Choice by Time)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
Affect	0.368	**12.52****	4, 86	<.001	0.118	0.234	**6.58****	4, 86	<.001	0.062
Univariate		*F*	*Df*	*p*	η_G^2		*F*	*Df*	*p*	η_G^2
Admin		**14.17****	1, 89	<.001	0.024		**7.26***	1, 89	0.008	0.035
Side Effects		**8.51***	1, 89	0.004	0.016		**10.15***	1, 89	0.002	0.045
Efficacy		**14.66****	1, 89	<.001	0.019		1.20	1, 89	0.276	0.004
Duration		**5.47***	1, 89	0.022	0.010		**5.15***	1, 89	0.026	0.025

	2-way (Choice by Treatment)					2-way (Time by Treatment)				
Multivariate	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*P*	η_G^2
Affect	0.052	1.19	4, 86	0.321	0.011	0.353	11.73	4, 86	<.001	0.099
Univariate		*F*	*df*	*p*	η_G^2		*F*	*Df*	*P*	η_G^2
Admin		1.53	1, 89	0.219	0.003		0.08	1, 89	0.780	0.000
Side Effects		1.20	1, 89	0.277	0.002		**9.33***	1, 89	0.003	0.017
Efficacy		0.96	1, 89	0.329	0.003		**22.28****	1, 89	<.001	0.028
Duration		0.28	1, 89	0.601	0.001		**26.14****	1, 89	<.001	0.045

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F4.2: *Coherence shifting of Study 2 serious disease affect scores by choice, time, and treatment: Main effects.*

	Choice					Time				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2	*Pillai's Trace*	*F*	*Df*	*p*	η_G^2
Affect	0.012	0.26	4, 86	0.902	0.002	0.817	**95.73****	4, 86	<.001	0.473
Univariate		*F*	*df*	*P*	η_G^2		*F*	*Df*	*p*	η_G^2
Admin		0.03	1, 89	0.860	0.000		**150.19****	1, 89	<.001	0.419
Side Effects		0.04	1, 89	0.844	0.000		**161.70****	1, 89	<.001	0.417
Efficacy		0.35	1, 89	0.554	0.001		**181.42****	1, 89	<.001	0.367
Duration		0.29	1, 89	0.595	0.000		**4.93***	1, 89	0.029	0.023

	Treatment (K vs M)				
Multivariate	*Pillai's Trace*	*F*	*df*	*p*	η_G^2
Affect	0.565	**27.87****	4, 86	<.001	0.207
Univariate		*F*	*Df*	*p*	η_G^2
Admin		**5.11***	1, 89	0.026	0.009
Side Effects		**32.08****	1, 89	<.001	0.054
Efficacy		0.96	1, 89	0.329	0.003
Duration		**88.71****	1, 89	<.001	0.193

$N = 91$. * $p < .05$, ** $p < .001$. η_G^2 = Generalized Eta-Squared.

Table F5: *Regression coefficients and statistics effects of coherence shifting and time on skin conductance level.*

Mild Disease Task

Parameter	*B*	*SE (B)*	*t (df)*	*p*	*95%CI B*
Intercept	1.83	0.26	**7.05** (14)**	**< .001**	[1.27, 2.38]
Time (Quadratic)	-0.20	0.05	**-4.03** (14)**	**< .001**	[-0.30, -0.09]
zACSoverall	0.08	0.15	0.55 (14)	0.59	[-0.25, 0.41]
Time (Linear)	0.84	0.19	**4.32**(14)**	**< .001**	[0.42, 1.25]
Time (linear) * zACS	-0.03	0.12	-0.23 (14)	0.82	[-0.27, 0.22]
Time (quadratic) * zACS	0.01	0.03	0.48 (14)	0.64	[-0.25, 0.08]

*N = 16. * p < .05. ** p < .001.* zACSoverall = overall absolute coherence shifting of affect and importance scores.

Serious Disease Task

Parameter	*B*	*SE (B)*	*t (df)*	*p*	*95%CI B*
Intercept	2.13	0.27	**7.85** (11.94)**	**< .001**	[1.54, 2.72]
Time (Quadratic)	-0.14	0.04	**-3.36* (5.97)**	**0.02**	[-0.24, -0.04]
zACSoverall	0.15	0.16	0.95 (11.94)	0.36	[-0.19, 0.50]
Time (Linear)	0.58	0.17	**3.51* (5.35)**	**0.02**	[0.16, 1.00]
Time (linear) * zACS	0.13	0.10	1.31 (5.35)	0.24	[-0.12, 0.37]
Time (quadratic) * zACS	-0.03	0.02	-1.41 (5.97)	0.21	[-0.09, 0.03]

*N = 14. * p < .05, ** p < .001.* zACSoverall = overall absolute coherence shifting of affect and importance scores.

Table F6: *Aversive feelings and overall strength of Coherence Shifting in the Study 2 mild and serious disease tasks.*

	Mild Disease Task	Serious Disease Task				
Feeling	*M (SD)*	*M (SD)*	*r (Between-Task)*	*Paired-samples t*	*df*	*P*
Anxious	4.49 (1.61)	5.88 (1.77)	**0.426****	**7.03****	83	< .001
Stressed	4.50 (1.84)	5.88 (1.95)	**0.496****	**6.64****	83	< .001
Unpleasant	4.49 (187)	5.68 (2.03)	**0.373****	**4.99****	83	< .001
Conflicted	4.55 (1.91)	5.27 (2.06)	**0.271***	**2.77***	83	0.007
Coherence Shifting						
Importance	0.084 (0.075)	0.308 (0.069)	0.136	-0.072	83	0.943
Affect	0.101 (0.076)	0.113 (0.082)	-0.179	0.940	83	0.350

$N = 84$. * $p < .05$, ** $p < .001$.

Table F7: *Descriptive statistics and correlations for Experiment 2 between-task and exploratory analysis.*

	M (SD)	1	2	3	4	5	6	7	8	9
1 Mild Imp. ACS	0.08 (0.08)									
2 Mild Affect ACS	0.1 (0.08)	.113								
3 Serious Imp. ACS	0.08 (0.07)	.136	-.062							
4 Serious Affect ACS	0.11 (0.08)	-.139	-.179	**.219***						
5 BEQ Negative Emotionality	22.85 (6.16)	.035	**.293****	-.020	-.102					
6 BEQ Positive Emotionality	20.75 (3.8)	.033	.109	.064	.019	**.633****				
7 BEQ Impulse Strength	27.05 (6.41)	.041	**.273***	.017	-.021	**.494****	**.575****			
8 BEQ Overall	70.64 (13.79)	.044	**.288****	.017	-.050	**.851****	**.826****	**.844****		
9 ERQ Cognitive Reappraisal	29.99 (5.21)	.018	.053	-.015	.054	**.263***	**.223***	-.089	.137	
10 ERQ Expressive Suppression	13.96 (5.24)	.044	-.106	.090	.070	**-.667****	**-.578****	**-.389****	**-.638****	**-.331****
11 Serious Choice Confidence	3.73 (0.83)	.107	.003	-.046	-.044	-.037	.066	-.084	-.037	-.101
12 Job Choice Confidence	3.38 (0.86)	-.068	-.031	-.031	.200	-.116	.136	-.095	-.058	.084
13 Serious Confidence Change	0.05 (0.73)	-.081	**-.267***	-.102	.181	-.085	-.013	-.148	-.110	.051
14 Job Confidence Change	0.12 (0.67)	-.186	-.084	-.152	.070	-.057	-.021	-.179	-.115	-.062

	10	11	12	13
11 Mild Choice Confidence	.037			
12 Serious Choice Confidence	.016	**.368****		
13 Mild Confidence Change	-.085	**.483****	**.317****	
14 Serious Confidence Change	.074	.038	**.318****	.063

$N = 84$. * $p < .05$, ** $p < .001$, ACS = Absolute Coherence Shifting, Imp. = Importance, BEQ = Berkeley Expressivity Questionnaire, ERQ = Emotion Regulation Questionnaire.

Appendix G: Aggregate Coherence Shifting Measures

This method for creating an aggregate coherence shifting measure was adapted from Carpenter et al., 2016. To create an aggregate variable for outcome affect ratings, all outcome ratings are first linearly transformed to a -1 to +1 scale and then combined using the following formula (with subscripts based on the serious disease task):

$$S_{Aff} = \frac{1}{8}\left(Aff_{Admin,K} + Aff_{Side\ Effect,K} + Aff_{Efficacy,K} + Aff_{Duration,K} - Aff_{Admin,M} - Aff_{Side\ Effect,M} - Aff_{Efficacy,M} - Aff_{Duration,M}\right)$$

where K is Treatment K, M is Treatment M, and "Admin" is Administration method. This will result in a value ranging from -1 to +1, with -1 representing maximum affect-based favorability towards Treatment M's outcomes and +1 representing maximum favorability to Treatment K's outcomes.

This score will be calculated for both pre-choice and mid-choice ratings, and a change score for affect between pre- and mid-choice times can be computed as:

$$CS_{Aff} = S_{Aff,Mid-choice} - S_{Aff,Pre-choice}$$

where negative scores would indicate a shift towards Treatment M and positive scores would mean a shift towards Treatment K. (Note that scores above +/-1 are possible but would require a large reversal of preferences between these two times.) These changes must be examined to ensure they are in the correct direction given the final choice. For participants where this is the case, the absolute value of the affect change score can be used as a measure of affect coherence shifting magnitude.

$$ACS_{Aff} = \left|S_{Aff,Mid-choice} - S_{Aff,Pre-choice}\right|$$

Similarly, a composite variable can be calculated for attribute importance weights after they have been linearly scaled to a 0 to 1 scale, using the formula:

$$S_{Imp} = \frac{1}{4}\left(Imp_{Admin} + Imp_{Efficacy} - Imp_{Side\ effects} - Imp_{Duration}\right)$$

where, in the serious disease task, Administration methods and Efficacy are favorable for Treatment K, and Side effects and Duration of symptoms are favorable for Treatment M (See Table 3). This measure also ranges from -1 to + 1 with positive scores being favorable towards Treatment K and negative scores favorable to Treatment M. This can also be used to calculate a pre-choice to mid-choice difference score:

$$CS_{Imp} = S_{Imp,Mid-choice} - S_{Imp,Pre-choice}$$

Then, if the direction of change is appropriate, an absolute measure of magnitude is calculated.

$$ACS_{Imp} = \left|S_{Imp,Mid-choice} - S_{Imp,Pre-choice}\right|$$

ACS_{Aff} and ACS_{Imp} are sufficiently aggregated to be used for between-tasks comparisons.

Additionally, a measure of overall within-task coherence shifting is necessary for comparisons with skin conductance. For this individual differences measure, z-scores will be taken for the two coherence-shifting variables and summed.

$$zACS_{Overall} = z(ACS_{Aff}) + z(ACS_{Imp})$$

REFERENCES

Abelson, R. P. (1995). *Statistics as principled argument*. New York, NY: Psychology Press.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csàki (Eds.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267-281). Budapest, Akadémiai Kaidó.

Alhakami, A. S., & Slovic, P. (1994). A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk Analysis*, *14*(6), 1085-1096. https://doi.org/10.1002/(sici)1099-0771(200001/03)13:1%3C1::aid-bdm333%3E3.0.co;2-s

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379-384. https://doi.org/10.3758/BF03192707

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336-372. https://doi.org/10.1016/j.geb.2004.06.010

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

Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, *6*(2), 215-225. https://doi.org/10.1093/cercor/6.2.215

Bickel, R. (2007). *Multilevel Analysis for Applied Regression: It's Just Regression! (1st ed.)*. New York, NY: The Guilford Press.

Boucsein, W. (2012). *Electrodermal activity* (2nd Ed.). Springer Science & Business Media.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59. https://doi.org/10.1016/0005-7916(94)90063-9

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261-304. https://doi.org/10.1177/0049124104268644

Byrne, K.A., Peters, C., & Willis, H.C. (2018, May). *Working hard or hardly working? How emotion affects effort-based decision-making*. To be presented at the Association for Psychological Science, San Francisco, CA.

Carpenter, S. M., & Niedenthal, P. M. (2017). Emotional processes in risky and multiattribute health decisions. *Psychology & Health, 33*, 58-76. https://doi.org/10.1080/08870446.2017.1314478

Carpenter, S. M., Yates, J. F., Preston, S. D., & Chen, L. (2016). Regulating emotions during difficult multiattribute decision making: The role of pre-decisional coherence shifting. *PloS one*, *11*(3), e0150873. https://doi.org/10.1371/journal.pone.0150873

Charpentier, C. J., De Neve, J. E., Li, X., Roiser, J. P., & Sharot, T. (2016). Models of affective decision making: How do feelings predict choice? *Psychological Science*, *27*(6), 763-775. https://doi.org/10.1177/0956797616634654

Cohen, J. (1988), *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, N.J.: Lawrence Erlbaum.

Cohen, J., & Cohen, J. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, N.J: L. Erlbaum Associates.

Ekstrom, R.B., French, J.W. & Harman, H.H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Elwyn, G., Laitner, S., Coulter, A., Walker, E., Watson, P., & Thomson, R. (2010). Implementing shared decision making in the NHS. *BMJ*, *341*, 971-975. https://doi.org/10.1136/bmj.c5146

Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., … Barry, M. (2012). Shared Decision Making: A Model for Clinical Practice. *Journal of General Internal Medicine*, *27*(10), 1361–1367. https://doi.org/10.1007/s11606-012-2077-6

Figner, B., & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), *A handbook of process tracing methods for decision research.* New York, NY: Psychology Press.

Goldberg, L. R. (2010). Personality, demographics, and self-reported behavioral acts: The development of avocational interest scales from estimates of the amount of time spent in interest-related activities. In C. R. Agnew, D. E. Carlston, W. G. Graziano, & J. R. Kelly (Eds.), *Then a miracle occurs: Focusing on behavior in social psychological theory and research* (pp. 205-226). New York: Oxford University Press.

Gross, J.J., & John, O.P. (1998). Mapping the domain of expressivity: Multi-method evidence for a hierarchical model. *Journal of Personality and Social Psychology, 74,* 170-191.

Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 72,* 435-448.

Hobson, N. M., Saunders, B., Al-Khindi, T., & Inzlicht, M. (2014). Emotion down-regulation diminishes cognitive control: A neurophysiological investigation. *Emotion, 14*(6), 1014-1026. http://dx.doi.org/10.1037/a0038028

Hastie, R., & Dawes, R. M. (2010). From preferences to Choices. In *Rational choice in an uncertain world: The psychology of judgment and decision making (2nd Ed.)* (pp. 217-236). Thousand Oaks, CA: SAGE Publications.

Hunt, E. (2002). *Thoughts on thought: A discussion of formal models of cognition*. Mahwah NJ: Lawrence Erlbaum.

Jarvis, B. G. (2012). MediaLab (Version 2012) [Computer Software]. New York, NY: Empirisoft Corporation.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291. https://doi.org/10.2307/1914185

Kilpatrick, D. G. (1972). Differential responsiveness of two electrodermal indices to psychological stress and performance of a complex cognitive task. *Psychophysiology*, *9*(2), 218-226. https://doi.org/10.1111/j.1469-8986.1972.tb00756.x

Kuykendall, D., & Keating, J. P. (1990). Mood and persuasion: Evidence for the differential influence of positive and negative states. *Psychology & Marketing, 7*, 1-9. https://doi.org/10.1002/mar.4220070102

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119-l37). Norwood, NJ: Ablex.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146-159. https://doi.org/10.1037//0022-3514.81.1.146

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267-286. https://doi.org/10.1037/0033-2909.127.2.267

Luce, M. F. (1998). Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research, 24*(4), 409–433. https://doi.org/10.1086/209518

Makoul, G., & Clayman, M. L. (2006). An integrative model of shared decision making in medical encounters. *Patient Education and Counseling*, *60*(3), 301-312. https://doi.org/10.1016/j.pec.2005.06.010

Marks, I. M., & Matthews, A. M. (1979). Brief standard self-rating for phobic patients. *Behavior Research & Therapy*, 17, 263-267.

Marr, D. (1982), *Vision: A Computational Approach*, San Francisco, Freeman & Co.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genetic, Social, and General Psychology Monographs, 121*, 339–361.

Payne, J. W., & Bettman, J. R. (2004). Walking with the scarecrow: The information-processing approach to decision research. In D. J. Koehler & N. Harvey (Eds.). *Blackwell handbook of judgment and decision making* (pp. 110-132).

Peters, E. (2006). The functions of affect in the construction of preferences. In S. Lichtenstein & P. Slovic (eds.) *The construction of preference*, (pp. 454-463). New York, NY: Cambridge.

Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. H. (2009). Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied*, *15*(3), 213-227. https://doi.org/10.1037/a0016978

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873-922.

Schlösser, T., Dunning, D., & Fetchenhauer, D. (2013). What a feeling: The role of immediate and anticipated emotions in risky decisions. *Journal of Behavioral Decision Making*, *26*(1), 13-30. https://doi.org/10.1002/bdm.757

Simon, D. (2004). A third view of the black box: Cognitive coherence in legal decision making. *The University of Chicago Law Review, 71*(2), 511-586.

Simon, D., & Holyoak, K. J. (2002). Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology review*, *6*(4), 283-294. https://doi.org/10.1207/S15327957PSPR0604_03

Simon, D., Krawczyk, D. C., Bleicher, A., & Holyoak, K. J. (2008). The transience of constructed preferences. *Journal of Behavioral Decision Making*, *21*(1), 1-14. https://doi.org/10.1002/bdm.575

Simon, D., Krawczyk, D. C., & Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychological Science*, *15*(5), 331-336. https://doi.org/10.1111/j.0956-7976.2004.00678.x

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K. J. (2001). The emergence of coherence over the course of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1250-1260. http://dx.doi.org/10.1037/0278-7393.27.5.1250

Simon, D., & Spiller, S. A. (2016). The Elasticity of Preferences. *Psychological Science*, *27*(12), 1588-1599. https://doi.org/10.1177/0956797616666501

Slovic, P. (1987). Perception of risk. *Science*, *236*, 280-285. https://doi.org/10.1126/science.3563507

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24*, 311-322. https://doi.org/10.1111/j.0272-4332.2004.00433.x

Stacey, D., Bennett, C. L., Barry, M. J., Col, N. F., Eden, K. B., Holmes-Rovner, M., ... & Thomson, R. (2011). Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews*, *2011*(10). Htpps://doi.org/10.1002/14651858.CD001431.pub3

Strecher, V. J., McEvoy DeVellis, B., Becker, M. H., & Rosenstock, I. M. (1986). The role of self-efficacy in achieving health behavior change. *Health Education Quarterly*, *13*(1), 73-92. https://doi.org/10.1177/109019818601300108

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33,* 529-554. http://dx.doi.org/10.1086/214483

Tiziani, A. (2017). Harvard's Nursing Guide to Drugs (10th ed.). Chatswood, NSW: Elsevier.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*, 281-299. http://dx.doi.org/10.1037/h0032955

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550.

van Smeden, M., de Groot, J. A., Moons, K. G., Collins, G. S., Altman, D. G., Eijkemans, M. J., & Reitsma, J. B. (2016). No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology*, *16*, 163. https://doi.org/10.1186/s12874-016-0267-3

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, *165*(6), 710-718. https://doi.org/10.1093/aje/kwk052

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton, NJ; Princeton University Press.

Weinstein, N. D., & Sandman, P. M. (2002). The precaution adoption process model and its application. In R. J. DiClemente, R. A. Crosby, & M. C. Kegler (Eds.) *Emerging theories in health promotion practice and research.* (pp. 16-39) San Francisco, CA: Jossey-Bass.

Wong, Y. N., Egleston, B. L., Sachdeva, K., Eghan, N., Pirollo, M., Stump, T. K., ... & Meropol, N. J. (2013). Cancer patients' trade-offs among efficacy, toxicity and out-of-pocket cost in the curative and non-curative setting. *Medical Care*, *51*(9). https://doi.org/10.1097/MLR.0b013e31829faffd.