The Pennsylvania State University

The Graduate School

The College of Information Sciences and Technology

WHY THE CHANGE OF HEART? UNDERSTANDING THE INTERACTIONS BETWEEN PHYSIOLOGY, AFFECT, AND COGNITION AND THEIR EFFECTS ON DECISION-MAKING

A Dissertation in

Information Sciences and Technology

by

Christopher L. Dancy II

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2014

The dissertation of Christopher L. Dancy II was reviewed and approved* by the following:

Frank E. Ritter
Professor of Information Sciences and Technology
Dissertation Co-Advisor
Co-Chair of Committee

Michael D. McNeese
Senior Associate Dean for Research, Graduate Studies, and Academic Affairs
Professor of Information Sciences and Technology
Dissertation Co-Advisor
Co-Chair of Committee

John Yen
University Professor of Information Sciences and Technology

Frank G. Hillary
Associate Professor of Psychology

Peter Forster
Assistant Dean for Online Programs and Professional Education

*Signatures are on file in the Graduate School

ABSTRACT

How do physiological and affective processes interact with cognitive processes to change the way we think? How can we better understand the processes that underlie decision-making and choice behavior? This dissertation presents a novel hybrid cognitive architecture, ACT-R/Φ, which extends the ACT-R cognitive architecture with an integrative model of physiology and a model of affect and emotion. Extending a cognitive architecture with representations of affect and physiology allows the straightforward development of more illustrative computational process models of human behavior that explain experimental results from physiology, neuroscience, and psychology. Computational models were developed that provide an account of how physiological change due to psychological stress or homeostasis can modulate cognitive processes. An experiment was also run to explore how subliminal visual emotional stimuli affect physiology and decision-making behavior during the Iowa Gambling Task. Results indicate that effects of subliminal affective stimuli were dependent on participant sex and personality differences. A computational process model was developed that performs the same task and behaves similarly to the participants.

Physiological and affective states continually interact with cognitive processes, biasing memory and, consequently, decisions. Evolved adaptations that support physiological and affective change of behavior (e.g., natural reactions to thirst or hunger) affect the way we learn and make choices. The ACT-R/Φ hybrid architecture, and the theoretical architectural models that define it, can be used to develop models of human behavior that include the necessary accounts of physiology and affect that describe what the body needs and how changes in behavior affects these needs. This improved understanding of the architecture that constrains our behavior gives us a better opportunity to comprehend why we make the decisions we do and how we can use this knowledge to make better decisions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Frank Ritter has been an advisor and mentor to me during these past 4 years. If I had not taken his class during my first semester in graduate school, I would not have found such an interesting topic of research that I continue to keep "picking at". He has also been a very public champion of my work.

Michael McNeese has also been an advisor of mine these past 4 years. He's given me good advice these past few years and I've enjoyed our off-topic Penn State football discussions.

My other committee members, John Yen and Frank Hillary have provided useful and diverse discussions related to my research and have been there to help out when I needed it. John Yen helped me obtain an EDA device to use. Frank Hillary allowed me to hang around his lab meetings and work with one of his grad students on a project that helped me learn more about brain imaging analysis.

Carleen Maitland somehow convinced my future employers to allow me to pursue my PhD and delay my commitment to work for them after graduating with my undergraduate degree. If it wasn't for that one call, I likely wouldn't have found this topic of research.

David Reitter gave more encouragement and positive feedback for this work than I could have expected.

Reginald Adams was willing to meet and discuss emotion theory with a wide-eyed grad student from IST.

Robert Hester has helped me with any problems and questions I have had related to the HumMod simulation system.

Martin Yeh, has always been available to provide useful feedback and may be the most humble professor I've ever met.

The IST administrative staff, including Rhonda, Sue, and Michelle, have always been there to answer my questions and tell me the right people to talk to.

Ryan, Changkun, Jeremy, Alex, Moojan, Yang, and John M. were always available to discuss theories and ideas related to computational modeling, cognition, and neuroscience.

Eric and Vince have provided useful wisdom during my graduate career and have been good at getting me out of the lab to go grab a bite to eat or a beer to drink.

The noontime Rec Hall basketball crew provided a place to play basketball and forget about my work for an hour or two.

My family (including Meg and Coleen) has always been there when I needed them, constantly encouraging me, and reminding me that I was succeeding even when I felt like my work wasn't good enough.

Finally, my wonderful fiancée, Lonni, put up with me doing work for "just another hour" for just another *couple of* hours. She has cooked and cleaned for me when I was too busy to bother, and has always been there to remind me that I can get through *this* and to tell me when I should take a break. I don't know how I got so lucky.

Chapter 1 Introduction

Introduction

"Psychology, as a scientific discipline, must be constituted by its very nature from an uncomfortable recipe: One third brain science, one-third behavioral sciences (including ethological approaches), and one-third experimental science."

Panksepp, 1998, p. 12

How do physiology and affect modulate cognition? What is a useful way to represent and simulate the interactions between physiology, affect, and cognition? This dissertation (a) provides a review of existing literature relevant to these questions and (b) presents some answers to these questions. This dissertation introduces a unified theory of human behavior that is instantiated in a novel hybrid computational architecture, models of mental and physiological activity that run within this architecture, and results from a study that was used to collect behavioral and physiological data. The development and application of these methods were driven by existing work in physiology, psychology, neuroscience, and artificial intelligence.

The work in this dissertation began with the exploration of ways one may add emotion to a cognitive architecture. Despite models of emotion being developed in the past, a definitive void emerged as the computational models of behavior and cognition remained largely empty of any affective account. The work also began with an underlying intuition that physiological modulation should be represented in a computational system that accounts for affective behaviors due to the largely influential modulatory systems involved during these behaviors (e.g., the hypothalamic-pituitary-adrenal, or HPA, axis). A survey of literature on cognitive and emotional behavior, as well as their physiological underpinnings, has revealed that a good portion of the work in separate disciplines is heavily related, but has yet to be connected into a coherent model that can be represented in a computational system. A more tractable way to understand the plethora of results in psychology, neuroscience, and physiology is needed to understand what results from these disciplines mean for understanding the different processes of human behavior and how these processes may interact.

What is emotion?

"*Everyone knows what attention is*."

James, 1890, p. 403

This dissertation does not particularly address attention; it does briefly address subliminal presentation of stimuli with an experiment, but it is not focused on the underlying mechanisms or how they can affect the way we make decisions. However, the quote above sheds light on an issue one can encounter when understanding the mechanisms and process of another topic in human behavior: emotion. One could also say that, "Everyone knows what emotion is" and, following that, "I know how emotion affects the way I make choices and decisions" without much blame; it is, in many ways, an intuitive concept. However different definitions or perspectives of emotion can lead to disagreements and confusion on how to interpret experimental results and how to generalize these results to real world issues.

There are several perspectives taken on emotion in literature. One may take a more physiological perspective, focusing on central and peripheral causes and effects of affective behavior (e.g., Damasio & Carvalho, 2013; LeDoux, 2012; Panksepp & Biven, 2012; Rolls, 2013). It is also possible to develop a conception of emotion on a higher, cognitive (e.g., Lazarus & Folkman, 1984; C. A. Smith & Kirby, 2009) or social (e.g., Feldman Barrett, 2006b; Russell, 2003) level; focusing on how subjective-appraisal and social-dynamics cause and result from emotional behavior.

The work in this dissertation uses a predominantly physiological perspective because it gives affective behavior a foundation in human physiology and grounds emotional theory in some known physiological processes. Within this perspective, “emotional behavior” and its antecedents is due to the interaction of several psychological processes realized by the interaction between several neural circuits. These processes can also be modulated by changes in central and peripheral physiology. Thus, “fearful” behavior (Figure 1-1) may be the result of learning to associate some condition with an unconditioned response (e.g., a painful shock) that activates an innate *FEAR* system (Panksepp & Biven, 2012) that may cause activation of a *Defense Survival Circuit* (LeDoux, 2012); this also may result in a change in peripheral physiology (e.g., an increase in peripheral cortisol and epinephrine).

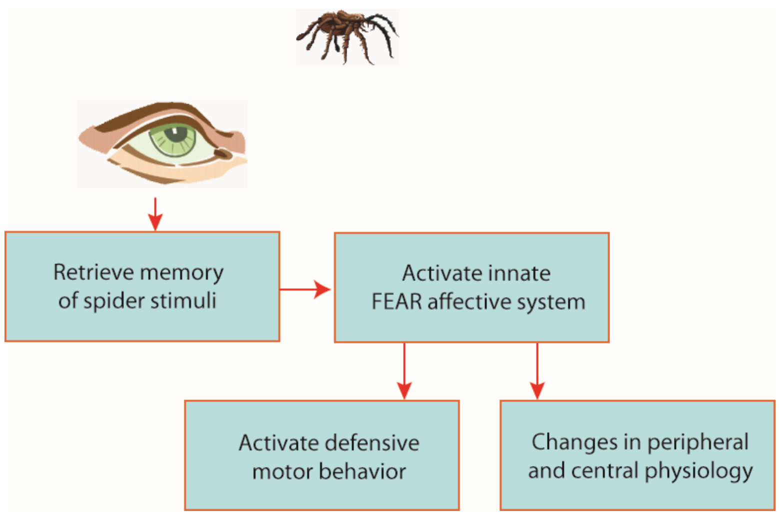

Figure 1-1. A simplified high-level view of the process of *fear* behavior

This physiological perspective also provides a foundation for answering the question “How can the human mind occur in the physical universe”, which was originally posited by Allen Newell (1991) and later used by Anderson (2007) to provide a focus on how cognition may be described by a computational architecture. With the integration of a systematic physiological model (Hester, Iliescu, Summers, & Coleman, 2011), the existing work by Panksepp and Biven (2012) and LeDoux (2012), and the existing work in unified theories of cognition (Anderson, 2007; Newell, 1990), a more representative architecture can be developed that can be used to understand behavior related to emotion by including an account for physiology.

Unified theories of cognition: A way to understand the architecture of cognition that constrains human behavior

Despite the high quality of work done in psychological experiments and theory, it has been previously argued that understanding these results and models together can be difficult (Newell, 1973). Unified Theories of Cognition (UTCs; Newell, 1990) attempt to bypass some of these potential issues by providing an architecture that can be used to describe different human behavior as the interaction of different psychological processes; thus, underlying a *decision* is the interaction of different memory and perceptual processes that precede any observed output. UTCs can allow a user to consider how seemingly unrelated processes can affect the behavior being studied without completely shifting the focus of the original research question. These unified theories can be important vehicles to move psychological models forward with more precision.

Newell (1990) discusses *bands* of behavior that characterize different theoretical focuses and time scales during the study of human behavior: the *Biological Band* (e.g., neural behavior), the *Cognitive Band* (e.g., symbolic memory retrieval), the *Rational Band* (e.g., complex behavior composed of cognitive band operations), and the *Social Band* (e.g., interaction with other people over a period of time). Though these bands (composed of levels of behavior) are useful, they are a reflection of a time when the physiological (especially neurological) processes modulating cognitive behavior were less understood and studied. Experimental data has been increasingly explored from a physiological perspective to further describe changes in psychological processes. Despite this increased understanding of the connections between Newell's cognitive and biological bands, a more unified account of these connections that can be developed into a computational system is still necessary.

Unified Theories of Cognition realized as computational cognitive architectures take a step beyond the theory and allow potential users to develop computational process models that can be run within simulation software. This step is important as it can give modelers behavior predictions of interactions between stochastic processes. It also allows a potential user to formalize a model more strictly as the model should be able to reliably run within the computational cognitive architecture. Other researchers can then better understand the model by running it themselves and can extend the model to explore other research questions or tasks.

A computational cognitive architecture can also be a useful tool for representing multiple levels of behavior more explicitly. These simulations can be used to explore some possible effects of epinephrine on cognition (e.g., Dancy, Ritter, Berry, & Klein, in press), the effects of sleep deprivation on behavior (e.g., Gunzelmann, Gross, Gluck, & Dinges, 2009), or to explore how cognitive limitations may affect social behaviors and structure (e.g., Zhao et al., Accepted).

ACT-R (Anderson, 2007) is a cognitive architecture that has been used to represent a wide-range of human behavior. The architecture has also gone through several extensions that have partially been driven by research topics explored by users of the architecture (e.g., Byrne, 2001; Fleetwood & Byrne, 2002; Fu & Anderson, 2006). Despite the progression of the architecture over time, it has yet to provide unified and systematic accounts of affective and physiological modulation of cognition. Other computational architectures provide accounts of affective and physiological modulation (e.g., Bach, 2009; Silverman et al., 2012), but have yet to match the human-like results of ACT-R models. Thus, a notable gap exists with modeling cognition, affect, physiology, and emergent behaviors that may arise due to the interaction between these processes.

Research objectives and overview of dissertation

The goal of this research is to further the understanding on ways physiology, affect, and cognition interact to mediate memory and choice processes that underlie decision-making. This dissertation approaches that goal by reviewing related work from neuroscience, psychology, artificial intelligence, and computational physiology and integrating that work into a new hybrid computational architecture. A computational process model is also presented that runs within the new architecture and performs a decision-making task. Electrodermal activity, behavioral, and questionnaire data were collected from participants who completed the same task. This task was run to explore the effects of non-integral affective stimuli on physiological and cognitive behavior.

This dissertation consists of five additional chapters. In chapter two, a review of perspectives in emotion and decision-making are presented. Within the same chapter, computational cognitive architectures, computational physiology systems, and past attempts at integrating computational representations of physiology or emotion and cognition are also reviewed. The review in chapter two outlines what is known in these areas, and some of the ways a hybrid computational architecture that represents behavior on the physiological, affective, and cognitive levels should constrain behavior.

The extensions made to the ACT-R cognitive architecture to create the ACT-R/Φ hybrid architecture are then outlined in the third chapter; this includes details of the additional modules added to ACT-R and connections to the HumMod physiological simulation system. In the fourth chapter, an ACT-R/Φ computational process model that performs a modified version of a decision-making task (the Iowa Gambling Task) is presented. The fifth chapter explains the methodology and results from a study run to collect electrodermal activity, task-related behavior, and questionnaire data during (before and after for the questionnaire data) the same computerized modified Iowa Gambling Task that was performed by the process model. The sixth chapter of this dissertation presents comparisons between model predictions from chapter four and experimental data recorded in chapter five.

The final chapter provides a discussion and conclusion to the dissertation. In this chapter, results are further discussed, and limitations of the study and development of the computational system are discussed. Finally, the dissertation reviews contributions made with the work and presents potential future directions of the work.

Preview of areas of contributions

In line with the interdisciplinary nature of the College of Information Sciences and Technology, this work contributes to multiple research communities. The new hybrid cognitive architecture, the decision-making model, and the experimental results positively contribute to work in emotion, cognitive science, computational physiology, neuroscience, and decision-making.

A cognitive architecture with physiology and affect

ACT-R/Φ is a novel hybrid computational architecture that extends a widely used cognitive architecture (ACT-R) with representations of physiology and affect. It combines a cognitive architecture, an integrative model of human physiology (HumMod), and a major theory in affect and emotion (primary-process affect theory). This is the first and only cognitive architecture to use an integrative model of human physiology to dynamically change cognitive behavior. The affect representation in ACT-R/Φ provides one of the few accounts of affect/emotion in a cognitive architecture and is the first to connect cognitive, affective, and physiological representations into one system. The architectural model that is implemented in the ACT-R/Φ computational architecture is also the first to suggest that the ACT-R architectural model and the primary-process affect model can be combined into a coherent theoretical model based on work in neuroscience, cognitive science, and psychology. The ACT-R/Φ hybrid architecture can be used to model how the interactions between physiology, affect, and cognition may result in differences in human behavior. Because ACT-R/Φ is an extended version of the ACT-R cognitive architecture, ACT-R process models can be used within the ACT-R/Φ architecture with little or no modification; thus, previous ACT-R models can be directly compared to similar models that use the new components of ACT-R/Φ. This architecture is freely available, so that the ACT-R community, or anyone else wishing to develop a cognitive model, can use the architecture to develop interesting models and further extend the architecture.

A computational process model of the Iowa Gambling Task

I have developed a computational process model that completes the Iowa Gambling Task (IGT). This process model communicates with a Matlab version of the IGT. The Matlab version of the IGT was also completed by participants in a study; the results from this study are presented in this dissertation. This process model runs within the ACT-R/Φ architecture and is the first known computational process model that uses perceptual, memory, and affective processes to complete the IGT. The IGT model can be expanded to examine how changes in the IGT (e.g., decision-making complexity) and changes in physiology (e.g., a rise in cortisol during the task) may change behavior during the task. The model code is currently available by request so that it can be used and extended by other researchers.

A study examining the effects of subliminal images on decision-making

A modified version of the Iowa Gambling Task was run to examine how subliminally presented images may affect decision-making related behavior. This is the first experiment to explore the effects of subliminally presented images on decision-making behavior during the IGT; the images were obtained from the International Affective Picture System (IAPS). This is also the first study to explore associations between ratings on the Affective Neuroscience Personality Scales (ANPS; personality scales based on primary-process affect theory), ratings on the positive and negative affect scales (PANAS), and decision-making behavior with or without subliminal image presentation. Thus, both the experimental treatment (a set of subliminal images), and using these affect measurement instruments (PANAS and ANPS) are contributions to the literature. The computerized version of the IGT was developed with scripts that use the Psychtoolbox library (Brainard, 1997) in Matlab and are freely available. This software also represents a contribution as no existing software Matlab-based version is freely available and the scripts used to run the study presented in this dissertation are particularly useful because they can be used to conduct versions of the IGT that require millisecond timing precision (Kleiner, 2010).

Chapter 2 Review

In understanding human behavior, it is important to have a grasp on the psychological processes that mediate behavioral output and the physiology that implements and modulates these processes. It is also important to understand how human psychological and physiological processes may interact to create *systems* of processes that mediate human behavior. Simon (1996) argues that "… we do not have to know, or guess at all, all the internal structure of a system but only that part of it that is crucial to the abstraction". In this way, one does not necessarily need to know all of the physiological, affective, and cognitive processes to understand aspects of human behavior. Still, it is useful to understand how physiological systems may interact with affective and cognitive systems to more realistically grasp the ways these interactions result-in and constrain behavior.

When understanding the ways physiology, affect, and cognition interact, it is important to have a fairly clear definition (or way of defining) the underlying process of behavior. To this end, I review literature related to affect and decision-making. Affect is conceptualized as encompassing what has traditionally been referred to as mood, emotion, and motivational drives; this is similar to the position and definition chosen by Panksepp (1998). Following this section, I provide a brief overview of some theories in decision-making that use accounts of affect or emotion to explain irrationalities of human decision-making to review some of the processes that describe decision-making and choice behavior. I also briefly review current cognitive architectures and models of computational physiology to highlight some of the computational systems that have been developed to simulate different levels of human behavior. Lastly, I review some past attempts of representing physiological or affective change of cognitive behavior in a computational architecture and call attention to some of the lessons one can gain from the development of these unique computational architectures.

Affect and emotion

> "Stepping back from the overarching concept of emotion and focusing instead on key phenomena that make emotion an interesting topic may be the best way out of the conceptual stalemate that results from endless debates about what emotion is."
>
> LeDoux, 2012, p. 654

The quote above from Joseph LeDoux, a researcher in the study of affect and emotion, highlights the difficulty with understanding emotional behavior. The domain remains vast and segmented largely due to the separate perspectives taken; one could potentially study and conceptualize physiology underpinning emotional experiences, or the cognitive or social aspects of the same experience. However, theoretical perspectives of emotion that focus on different perspectives can have complementary effects on one another.

It is possible to conceptually separate research on emotion into three perspectives: physiological-anatomical, cognitive, and social. Those in the physiological-anatomical perspective are typically interested in physiological systems enacted during processing of conditioned and unconditioned stimuli (e.g., LeDoux, 2012; Panksepp & Biven, 2012), those in the cognitive perspective focus more so on processes relating to appraisal (e.g., Lazarus & Folkman, 1984; Scherer, Schorr, & Johnstone, 2001; C. A. Smith & Kirby, 2009), and those in the social perspective focus on the social and cultural definitions that change how a person processes and responds to stimuli (e.g., Feldman Barrett, 2006b; Russell, 2003). Theories explored from these different perspectives often acknowledge others; however, the choice of methodologies used for studying human emotion will likely be determined by the perspective of interest and thus the data obtained can be difficult to compare between perspectives.

The conceptualization of levels used by Panksepp (2011) makes integrating the primary-process affect theory (Panksepp & Biven, 2012) with existing accounts of cognition more straightforward. With this perspective, many emotional behaviors can be conceptualized as an interaction between affective processes on the *primary-process* level and cognitive processes often explored on the *secondary* and *tertiary* levels that involve (among other things) learning and memory behavior (e.g., see Figure 2-1). This perspective is also useful because it grounds the affective categorization with behavior and neurological structures, while also attempting to separate unconditioned affective behavior from other conditioned behavior that also involves memory processes; not surprisingly, this theoretical framework would fall under the physiological-anatomical perspective.

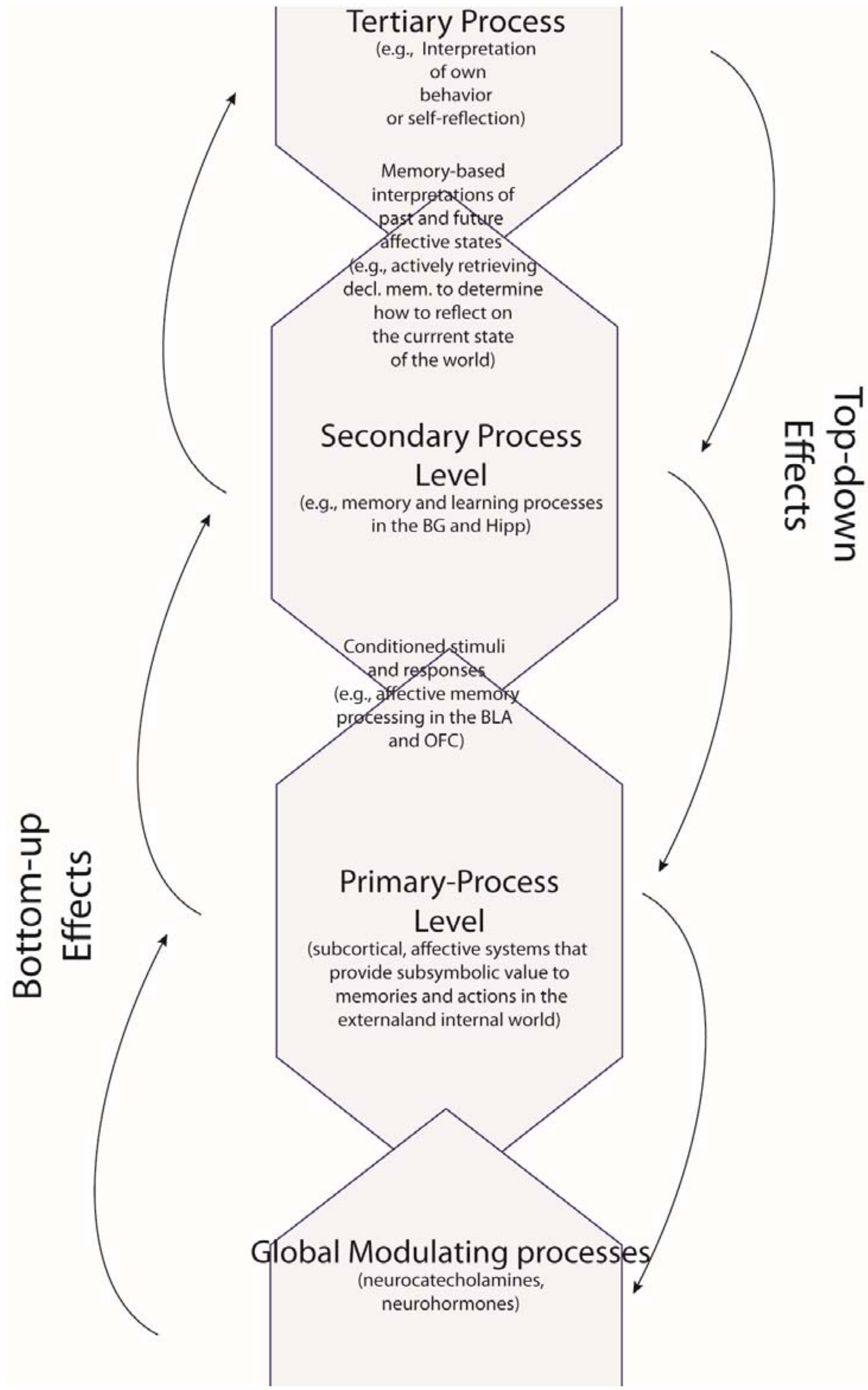

Figure 2-1. Levels of behavioral processes from Panksepp, Fuchs, and Iacobucci (2011)

In the next section, I review aspects of the primary-process affect theory and some ways it relates to learning and memory processes.

Primary-process affect

Panksepp and Biven (2012) posit primary affective systems that cause certain unconditioned behaviors when activated. Four of the major systems theorized to exist have been explored more deeply by Panksepp and other researchers: SEEKING (e.g., appetitive behavior), FEAR (e.g., flight behaviors), RAGE (e.g., affective attack), and PANIC (e.g. distress vocalization). These terms are capitalized in an attempt to avoid the confusion created by its usage as common vernacular.

The work in this dissertation focuses on SEEKING and FEAR systems because the SEEKING functional system is the most foundational of all the systems and both have been widely studied with varying terminology and perspectives (e.g., Alcaro & Panksepp, 2011; Berridge, Robinson, & Aldridge, 2009; Mobbs et al., 2007; Öhman, Carlsson, Lundqvist, & Ingvar, 2007; Raio, Carmel, Carrasco, & Phelps, 2012; J. Wright & Panksepp, 2012; Zhang, Berridge, Tindell, Smith, & Aldridge, 2009).

The SEEKING system (Figure 2-2) is associated with anticipatory-appetitive behaviors and is driven by activation of several neurological structures, including the ventral tegmental area (VTA) and the lateral hypothalamus (LH). This system should not be seen as one that also directly mediates consummatory behaviors (e.g., drinking a beverage when thirsty), SEEKING is the preceding system to consummatory behavior (Panksepp, 1998; Panksepp & Biven, 2012). Consummatory behaviors are hypothesized to lead to a reduction in activation of this basic appetitive system. Thus, this reduction in stimulation of the appetitive system may be more linked with associative reinforcement than inception of the stimulation (Panksepp, 1998). This result does fit a cognitive perspective, because one would expect a stimuli that leads to the completion of a goal (and the extinction of SEEKING behavior) to be reinforced in memory. The location of some of the structures in the SEEKING system (lateral hypothalamus) also make it an obvious target of homeostatic affective modulation (e.g., hunger and thirst) via interoceptors.

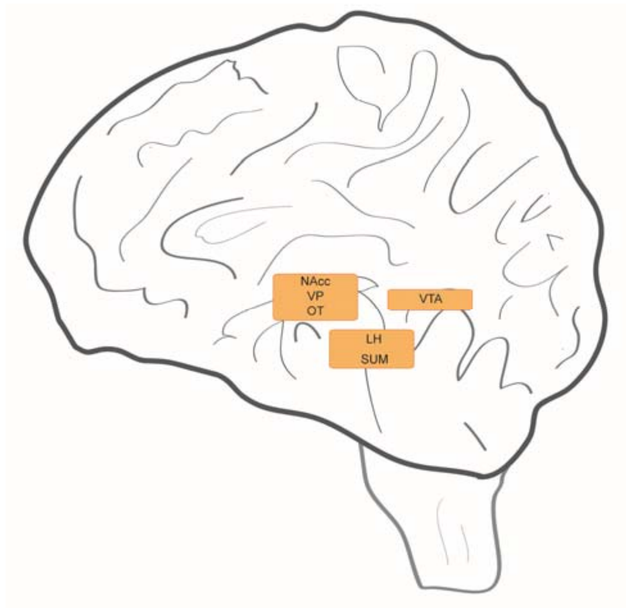

Figure 2-2. Neural structures contained in the SEEKING circuit. Abbreviations: LH – lateral hypothalamus; NAcc – nucleus accumbens; OT – olfactory tubercle; SUM – supramammilary nucleus; VTA – ventral tegmental area

One would expect the SEEKING system to be involved during gambling tasks (e.g., the Iowa Gambling Task) due to its relation to consummatory behavior and rewards. Past neuroimaging studies have been conducted that observe activation in the structures associated with the SEEKING system and reward (Ballard & Knutson, 2009; Haber & Knutson, 2009; Knutson, Adams, Fong, & Hommer, 2001; Knutson & Greer, 2008). In the case of the Iowa Gambling task one may expect a decrease in the activation of SEEKING structures when making more rewarding decisions because it is hypothesized to deactivate in response to rewarding consummatory behavior, though activation of any of these structures may be independent of SEEKING behavior (the circuit of connected structures is implicated in this behavior) depending on experimental conditions. This system conceptualization is descriptively useful and functionally relevant to the procedural memory systems in many cognitive architectures.

Another primal affective system posited by Panksepp and Biven (2012) is the FEAR system. Evidence suggests that the innate FEAR system consists of neural circuitry including the amygdaloid complex (commonly simply referred to as the amygdala), anterior/medial hypothalamus, anterior insula, and periaqueductal gray, Figure 2-3 (Johansen, Tarpley, LeDoux, & Blair, 2010; LeDoux, 2012; LeDoux & Phelps, 2008; Liotti & Panksepp, 2004; Mobbs et al., 2007; Öhman et al., 2007; Öhman & Mineka, 2001; Panksepp, 1998); the system represents an evolutionary development of innate responses to unconditioned behavior and connects to memory systems to facilitate learning of fear behaviors and stimuli that cause unconditioned responses. Hypothalamic and periaqueductal gray activity in this system is important as it is responsible for the autonomic response often observed when the FEAR system is activated. The amygdala is the most well-known (and experimentally tested) of these structures due to its importance to triggering responses to stimuli and learning to associate new stimuli with unconditioned responses (Johansen et al., 2010; LeDoux, 1996, 2012; Panksepp & Biven, 2012) .

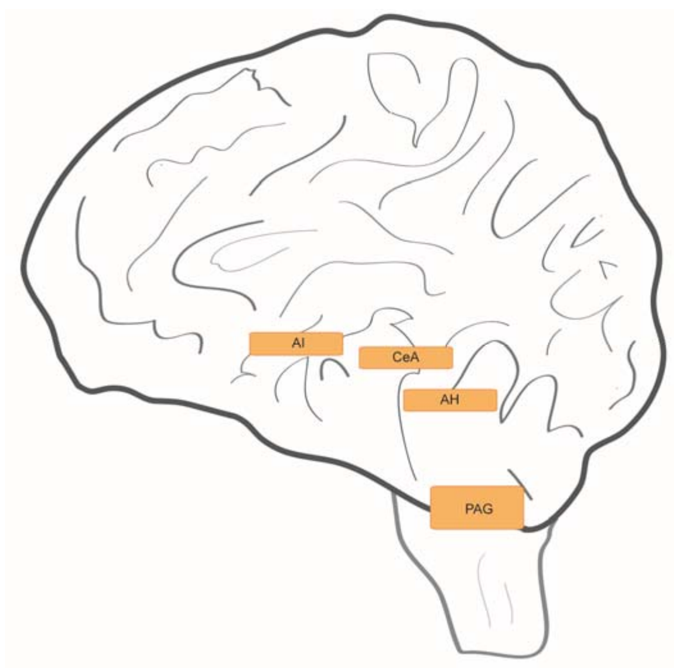

Figure 2-3. Neural structures contained in the FEAR circuit. Abbreviations: AH – anterior hypothalamus; AI – anterior insula; CeA – central amygdala; PAG – periacqueductal gray

Using animal studies of this system (e.g., rats' response to the smell of a cat) as a foundation for studies of this system in humans has been attractive due to the lack of both (human-level) cognitive and social influence on responses to FEAR system-activating stimuli. Similar to the SEEKING system, Panksepp and Biven (2012) posit the specific neural circuit involving the structures mentioned above using results of electrical stimulation studies, though the mentioned structures have also been implicated in FEAR system responses in human imaging studies (e.g., Mobbs et al., 2007; Öhman et al., 2007). While the concept of a primary FEAR system remains important, many of the studies on fear responses observe the interaction of the above mentioned structures with other physiological structures and psychological processes, that is, study fear above the primary level and do not specifically target unconditioned responses. Many studies focus on learned, conditioned responses to fear stimuli.

In studying affect and emotion on a primary-process level, it becomes clear that changes in physiology can have range of effects on affective behavior. Therefore, it is important to have a unified theory of human physiology, in addition to a unified theory of cognition, when creating a computational theory of affect and emotion.

Learning and affect

It is useful to consider the separate neural systems involved in unconditioned responses, because this allows a more coherent and accurate account of the psychological processes underlying conditioned affective and emotional behavior. The SEEKING and FEAR systems described above have often been explored from a secondary, i.e. learning and memory, perspective. Typically, experimental data collected from this perspective will explore conditioned responses or conditioned stimuli. Studies related to learning and the SEEKING system have traditionally used concepts such as the *reward system* or incentive salience (e.g., Richard, Castro, DiFeliceantonio, Robinson, & Berridge, in press).

The incentive salience, or wanting, framework (Berridge et al., 2009) is heavily related to the work on the SEEKING system, but focuses on learning on a secondary-level (e.g., as shown in Figure 2-1). Many of the neurological structures posited to be involved in SEEKING behavior are also included in wanting/incentive salience behavior. The work with the incentive salience theory is also useful as researchers have used the perspective to tease apart *liking* and *wanting* in the brain (Berridge et al., 2009; McClure, Daw, & Read Montague, 2003; K. S. Smith, Berridge, & Aldridge, 2011) and researchers have developed computational models of incentive salience and its effect on reinforcement learning (Zhang et al., 2009).

Incentive salience is mediated by a dopaminergic system that includes the ventral tegmental area (VTA), substantia nigra, nucleus accumbens shell, and dorsal striatum (Richard et al., in press); activity in the dorsal striatum has also been related to the procedural memory component in cognitive architectures (e.g., Anderson, 2007). Other neural substrates like the basolateral amygdala (BLA) and the ventromedial prefrontal cortex (VMPFC) are also known to affect the wanting component of incentive salience. Therefore, it is useful to consider systematic accounts of the effect of incentive salience on human behavior given the breadth of structures involved that also mediate other characterized behaviors (e.g., the BLA and VMPFC in fear and goal behavior).

Fear memory research typically involves the study of conditioned (previously neutral) stimuli and responses that are learned as a consequence of being paired with certain aversive unconditioned stimuli (e.g., pain from a shock or a loud noise). The amygdala is the neural substrate that is most widely associated with fear behavior and learning. Though studies indicate that the amygdala does in fact play a large role in fear, studies in this area also indicate that the amygdala can be separated into substructures that are connected, but serve different functions in the cause of and response to fear stimuli (Delgado, Jou, & Phelps, 2011; Maren & Quirk, 2004). The central amygdala (CeA) is the substructure of the amygdala that is primarily involved with behavioral expression in response to aversive stimuli. CeA connections to the anterior and medial hypothalamus, and consequently the periaqueductal gray (PAG), facilitate this expression of certain behavioral (e.g., fleeing or freezing) and physiological (e.g., an increase in sympathetic nervous system activity) changes (Moscarello & LeDoux, 2013; Panksepp et al., 2011); these substrates are theorized to be key to the feeling of fear (Panksepp, 1998; Panksepp et al., 2011) and have been shown to increase in activation as aversive threats become more imminent (Mobbs et al., 2007). In addition to the structures just discussed, conditioned learning of fear behavior and stimuli have been associated other portions of the amygdala, particularly the lateral and basolateral amygdala substructures (LA and BLA, respectively; Fanselow, 2010; LeDoux, 1996, 2012; Moscarello & LeDoux, 2013; Phelps, 2006).

The study of fear and its effects on memory are heavily related to the studies of stress as aversive stimuli often cause an activation of *stress systems* (Joëls, Fernandez, & Roozendaal, 2011; Öhman, 2008; Öhman & Mineka, 2001; Panksepp et al., 2011). The typical systems involved in stress response discussed are the hypothalamic-pituitary-adrenal axis (HPA-axis) and the sympathetic-adrenal-medullary axis (SAM-axis); these two systems also serve other physiological systems not discussed here, but that may be related to affect and cognition (e.g., digestion and sleep). Activation of these systems is known to affect memory processes (e.g., see Joëls et al., 2011; Wolf, 2009) and these effects are typically associated with changes in epinephrine, norepinephrine, and cortisol that are general modulators of neural structures like the hippocampus and ventral tegmental area (Cahill & Alkire, 2003; De Quervain, Aerni, Schelling, & Roozendaal, 2009; Sara, 2009). The VTA is one of the main neural modulators of dopamine, which is known to affect SEEKING (Alcaro & Panksepp, 2011), wanting/incentive salience (Berridge et al., 2009), and procedural learning (Anderson, 2007).

It is clear that the conceptualization of emotional effects on cognition (in particular, learning and memory) likely should include a representation of how underlying physiological effects can modulate cognitive and affective processes. These processes can potentially become independent of the original causes of noticeable change in the physiological system, i.e., responses can become conditioned and operate irrespective of any unconditioned stimuli. A systematic account of how these processes can affect each other should be considered and, in particular, it is important that the system can allow one to break apart the structural and functional systems involved in a coherent way. A computational account of affective modulation of cognition and physiology should also include an affective memory system, composed of the separate emotional memory systems, that is distinct of other forms of memory (see Eichenbaum, 2010; LaBar & Cabeza, 2006; Squire, 2004 for reviews on memory systems).

Grounding conceptualizations of emotion

Differences in emotion conceptualization make it difficult to compare approaches from an empirical point of view and a computational point of view. Studies taken from a neuroscience perspective seem a natural fit in grounding emotional theory in behavioral and physiological data that have been collected together with a common goal. Several imaging studies have been conducted to link different aspects of emotional processing and behavior with physiological changes (Phan, Wager, Taylor, & Liberzon, 2002). Studies of physiological reactions to both conscious (e.g., Adams et al., 2011; Britton et al., 2006; Büchel et al., 1998) and unconscious (e.g., Liddell et al., 2005; Morris et al., 1998) stimuli have been conducted.

When comparing emotional processing of stimuli to other studies it is not only useful to have stimuli with similar properties, but also have stimuli that have shown to evoke consistent emotional responses. Some researchers may use primitive sounds, which are distinguishable via frequency (e.g., Bradley & Lang, 2007a). Another way to elicit change in affect is to use a common set of visual stimuli (e.g., images of snakes and spiders). The International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1997) particularly provides an attractive set of visual stimuli due to its correlation with human self-report data collected using the Self-Assessment Manikin System; this system assigns values in arousal, valence, and dominance dimensions to the pictures. While the dimension perspective is typically at odds with any conceptualization that assigns discrete emotional labels (e.g., Feldman Barrett, 2006a), IAPS has been used to bridge the two concepts, thus allowing for integration of the distinct theoretical conceptualizations (Mikels et al., 2005).

One also may find a bridge between conceptualizations if one separates emotion into the processing of emotional stimuli and the behavior in response to such stimuli. The dimensional account fits easily into processing of emotional stimuli as it seems to be a relatively straightforward way to quantify stimuli values during this stage. However coherent behavior, feelings, and correlated activity in certain circuits within the brain in response to both emotional stimuli and stimulation of certain brain structures, provide evidence for primal discrete emotions; in a similar fashion, Bradley and Lang (2007b) conceptualize primal networks within the brain for *appetitive* and *aversive* stimuli processing.

IAPS presents an attractive methodological approach for further learning structural and functional aspects of emotional processing. Presenting visual stimuli with a known normalized effect below consciousness (i.e., subliminally) allows one to observe behavior that can be considered more innate due to the assumed faster (quick and dirty; LeDoux, 2008) processing of this stimuli that occurs. One can observe and model how unconscious emotional processing, i.e. immediate emotions (Rick & Loewenstein, 2008), can interact with anticipatory affect and conscious cognitive processing of information (e.g., decision-making in gambling).

Measuring affect and emotion

To ground conceptual theoretical stances and observe experimental interactions of affect and cognition, it is useful to have standardized and validated methods of quantification. Often methods employed can fall into the personality scale and psychophysiological categories. Using a personality scale of some kind to quantify participants that would otherwise be assumed to be within the same category for experimental tests provides a manner to further explain variance amongst participants. Psychophysiological methods allow one to connect behavioral output and functionality to a physical cause and effect (or correlation). Psychophysiology can improve one's power of explanation for observed outputs from a human system. Below I discuss the Affective Neuroscience Personality Scales, a personality measurement tool, and the Positive and Negative Affect Scales (PANAS), a brief questionnaire that can be used to assess the current affective state of a person. I also discuss psychophysiological measurement methodologies, specifically paying attention to functional magnetic resonance imaging (fMRI), pupillometry, and electrodermal activity (EDA).

Affective Neuroscience Personality Scales

The Affective Neuroscience Personality Scales (ANPS, Davis & Panksepp, 2011) is a personality questionnaire with a theoretical foundation in primitive emotions, i.e. affective neuroscience (Panksepp, 1998). These scales are useful if one chooses to use a basic emotion theoretical framework (as discussed previously from a physiological perspective). This scale has been compared to the five factor model (Davis & Panksepp, 2011), which is a standard personality model (Digman, 1990). ANPS is an attractive supplement to any experimental procedure with a theoretical grounding in primary-process affect theory, and even tertiary emotional representations that fail to significantly conflict with the primitive conceptualizations.

Positive and Negative Affect Schedule

The Positive and Negative Affect Schedule (PANAS, Watson et al., 1988) is a scale for current binary mood measurement, i.e. positive or negative mood. The PANAS is meant to measure mood from the perspective of a certain time scale (e.g., how one may feel at the moment of administration or how one felt two weeks prior). The PANAS provides a relatively fast way to assess a participants' mood directly after an experiment is conducted. This scale has also been used to measure the effects of baseline affective state on behavior in the IGT (Suhr & Tsanadis, 2007).

Psycho-physiological measurement

There are several methods used to record activity in the brain depending on research needs, i.e. spatial and temporal resolution requirements. Electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET), and functional magnetic resonance technology (fMRI) have all been used to measure brain activity during the presentation of some stimuli (phasic activity) or in absence of the presentation of some stimulus (tonic or spontaneous activity). Of these techniques, those using EEG will achieve the greatest temporal resolution, because EEG is used to directly measure radial and tangential electrical signals. FMRI has the greatest spatial resolution of this group of methods, something especially useful when studying behavior related to neural subcortical structures.

Functional Magnetic Resonance Imaging

Though the use of blood oxygen level dependent (BOLD) functional magnetic resonance imaging achieves a temporal resolution that is lower than EEG (e.g., seconds vs. milliseconds, respectively), the spatial resolution one can achieve from its use is highly valuable for observing subcortical activity (e.g., Figure 2-4). Subcortical areas of the brain are often implicated in affective behaviors. The temporal resolution limitation of fMRI is due to a reliance on the hemodynamic response function (hrf).

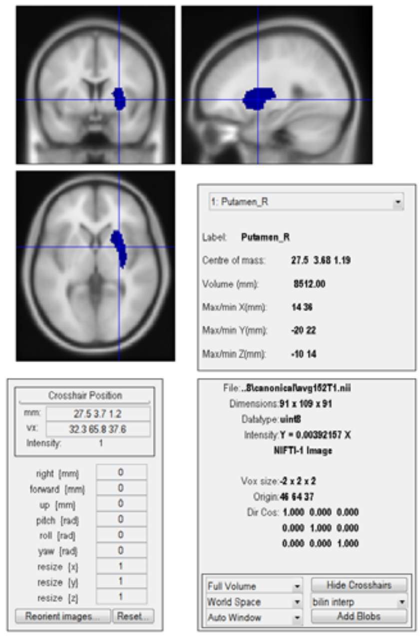

Figure 2-4. An example image output using the SPM fMRI analysis package (Wellcome Trust Centre for Neuroimaging). This image reflects activity in the Putamen

The hrf reflects a change in oxygenated blood in area of the brain due to activity in that area. Due to an initial dropout of signal in the first few seconds, temporal resolution is often limited to 3-4 seconds. It may be possible to achieve a lower temporal resolution depending on the type of experimental design one chooses to employ, i.e. event-related or block design.

An event-related design allows one to record activation in the brain in response to specific stimuli. Thus event-related designs may be useful when used to record activation during appetitive or aversive behavioral responses. Using a block design has an advantage of a decreased study time and an increase in signal to noise ratio. However, a huge disadvantage to using a block design is the use of cognitive subtraction that results in a high loss of temporal resolution

Despite the increased temporal resolution of event-related designs, concerns have been raised over the possibility of having too high of a granularity for some affective behavior. It has been posited that it may take up to 10s for activation in the brain to begin due to an emotional response (Liotti & Panksepp, 2004), possibly making an event-related paradigm inappropriate. One may also raise an opposite problem, processing could be too rapid (e.g., LeDoux, 1996) to correctly capture in even an event-related design. Though these concerns and others (e.g., Kosslyn, 1999) about fMRI and its use in experiments are rightfully raised, fMRI still seems the most attractive option for observing brain activity during affective and cognitive processing given its relative availability (as compared to MEG), low invasiveness (as compared to PET), and adequate spatial resolution (as compared to EEG).

Measurements of peripheral physiology are often conducted in behavioral experiments to measure somatic and autonomic activation (e.g., muscle contraction, or changes in hormone levels). Several peripheral measurement systems rely on changes in electrical activity due to neuronal action potential spikes. Electromyography (EMG), electrooculography (EOG), and electrocardiography (EKG) are used to detect bioelectrical activity during muscle contraction, eye movement, and heart rate (respectively); all of these methods have been used to detect peripheral changes to stimuli. EMG for example has been used in studying of facial activity during emotional reactions (P. Winkielman & Cacioppo, 2001) as has EOG (Tobin et al., 2000).

Physiological measurements not centered on human-generated bioelectrical activity also are used in measuring affective stimuli. Pupillometry can be used to record eye movement and pupil size while a participant is processing stimuli. Hormone levels can also be measured during a study to assess change in different stress systems as a result of the presentation of stimuli. Finally, electrodermal activity (EDA) is often used to measure general autonomic nervous system (ANS) response.

All of the mentioned peripheral physiology measurements have their drawbacks, but their differing levels of measurement often allow one to combine measurements to triangulate the physiological processes that are activated to implement, and in response to, behavioral changes. Below I describe two useful methods of measurement which can be used simultaneously within a study: pupillometry and EDA.

Pupillometry/Eye-Tracking

Eye behavior can be used to measure the effect of certain emotional stimuli (e.g., Partala & Surakka, 2003; Wadlinger & Isaacowitz, 2008; Nummenmaa & Calvo, 2008). Pupil size change is related to different environmental stimuli (Stern et al., 2001); for example, the pupils will become more dilated when processing a higher load of information (Beatty & Lucero-Wagoner, 2000). Pupil size is also theorized to be correlated with activity in the locus coeruleus (LC) (Aston-Jones & Cohen, 2005), a structure that increases neuro-endocrine (norepinephrine) response when activated. The measurement of pupil size can be accomplished by taking video of the user's pupils during stimuli presentation (e.g., using an eye-tracking system). Using an eye-tracking system for pupillometry has the advantage of being able to record user eye movements related to visual focus of attention. Example headgear for an eye-tracking system is presented in Figure 2-5.

Figure 2-5. An example of eye-tracker head gear

Though other methods, i.e. EOG, may arguably provide a superior method of measurement for eye movement, the video taken in pupillometry is often a sufficient measurement for eye-movement detection. Whether or not this video is sufficient is likely tied to the placement of the video camera used. Several eye-tracking systems in use today place a camera close to the eye to allow pupil tracking while also tracking overall eye movement. These cameras also can be mounted on head-gear of some kind to allow head motion and a (participant) point of view video. Thus, while there may be other options, pupillometry (as provided by many eye-tracking systems) is an attractive method of measurement when considering availability, ease of use, and functionality.

Electrodermal activity

Electrodermal activity (EDA) is used to measure peripheral physiological activity in reaction to stimuli. This method has been used many experiments and is useful due to its relatively low effect on the participant. EDA primarily records electrical activity in the eccrine sweat glands that are concentrated in the hands and feet. Eccrine sweat glands are innervated by the ANS (solely the sympathetic branch).

EDA measurement is relatively cheap, unobtrusive, and easy to achieve in an experimental setting. The usual placement of instrumentation (hands or wrist) makes it a viable measurement tool to combine with other behavioral measurement tools like an eye-tracker and/or instrument to conduct pupillometry. These peripheral tools give an experimenter the opportunity to run more experiments on participants due to the cost (usually) being independent of time and likely related to the price of the actual equipment. Thus a researcher may find it useful to combine (CNS and ANS) experimental methodologies to see the different effects a given stimuli or task has on a participant.

Though EDA is a tool commonly used to determine anticipatory emotion during the IGT, others (e.g., Colombetti, 2008) have noted the need for alternative non-invasive methods, like pupillometry, to observe autonomic response during the task.

Summary

There are many useful ways to conceptualize affect, it can be understood from a physiological, cognitive, or social perspective. Though these perspectives are all useful independently, they also serve a complimentary role to one another and help to give a more unified account of human behavior. Following this, experimental data from each of these perspectives can also serve a complimentary role if the limitation of the method, and the theoretical focus that is used when interpreting data, is understood.

Understanding affect and affective behavior from a physiological perspective is useful for this dissertation work because it can constrains theory to physiological plausible explanations and potentially gives us a more realistic view of how the architecture of the human mind may occur in the physical universe. Theoretical frameworks from affective neuroscience (e.g., LeDoux, 2012; Panksepp & Biven, 2012) provide useful guides for understanding how different levels of human behavior (e.g., Panksepp, 2011) result in and from changes in the external and internal environment. The conceptualization of human affective behavior offered by Panksepp and Biven (2012) and LeDoux (2012) also fit well with some unified theories of human behavior (e.g., Anderson, 2007) to provide a useful understanding of how physiological, affective, and cognitive processes may interact and modulate decision-making behavior over time.

The architectural model of human behavior presented in this dissertation combines existing theory from affective neuroscience with a unified theory of cognition. In turn, the architectural model describes how physiological, affective, and cognitive processes may interact and result in dynamic affective behavior over-time. A computational system based on the model is also presented that can be used to better understand how the interaction between physiological, affective, and cognitive processes can result in quantitative and qualitative differences in affective behavior.

Decision-Making

Theories of decision-making in classic economics are often of the expected utility variety and a normative nature (Stanovich, 2009). These theories provide an optimal decision in the face of given (or determined) probabilities and values. While theories of this variety may provide an optimal solution given optimal information, they fall short of actual human decision-making, because decision-related information is often incomplete, and, even with a fairly complete set of information, humans tend to exhibit several logical fallacies; much of the time valuable information is only obtained after the decision has been made.

Logical fallacies that cause a deviation from the optimal solution can be conceptualized if one looks at axioms of choice (Stanovich, 2009). Many fallacies stem from a breakdown in choice transitivity. Equation 2-1 shows the basic axiom behind transitivity; if one chooses A over

B and B over C, then one should consequently choose A over C; this conceptual idea behind the law of comparative judgment (Thurstone, 1927).

Equation 2-1

$$If\ A > B\ \&\ B > C\ \rightarrow A > C$$

This rule has been shown to be violated when a person is presented with separate, albeit irrelevant, choices after the decision has already been made (e.g., Sen, 1993). Many other common deviations from the axioms exist, like precluding effects of context (failing to follow the axiom of independence), the endowment effect (a tendency to overvalue an object that is already owned; Kahneman et al., 1991), and loss aversion.

Tversky and Kahneman (1981) discuss the effects of framing the perspective question and possible answers on a person's decision outcome. They theorize that the perception of choices can change based on how one chooses to frame the choice outcomes and also that value and decision-weights (i.e., expected utility) follow a non-linear distribution; thus they offer a theory aimed at explaining logical fallacies like those mentioned above. *Figure 2-6* shows an example set given by the researchers in a study.

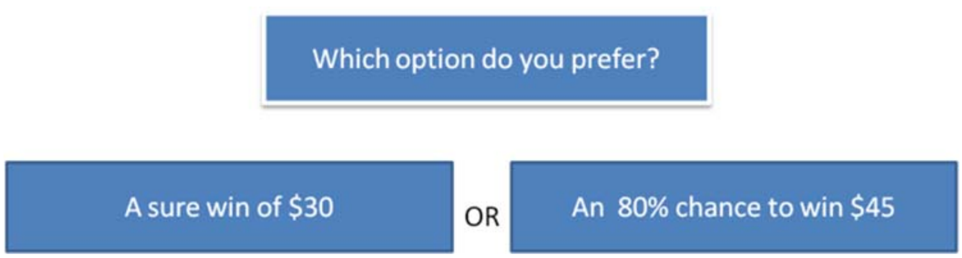

Figure 2-6. A question-answer set often incorrectly answered (adapted from Tversky, 1981).

When presented with these choices, users are more likely to select the first over the second option, even though the expected utility of the second option is actually higher than the (Tversky & Kahneman, 1981). While they note that bounded rationality (e.g., inconsistencies arising due to a limit on mental effort and consequently the amount of considered information; Simon, 1972) could be in play for some decision choices, they offer prospect theory (Kahneman

& Tversky, 1979) as a more reasonable explanation for the given experiments. According to them, the evidence also raises doubts of an eventual choice coherence criterion. Framing can even affect one's emotional response to the problem and/or choices and consequently effect perception of the choices.

A classical economic perspective on decision-making is often insufficient in explaining human behavior. While these theories may prove helpful from an optimality point-of-view, theorists like Kahneman et al. (1991) have pointed out several deviations of human behavior from these normative perspectives on decision-making. In light of the shortcomings of this classic logic based decision-making, some researchers have chosen to view decision-making with an affective perspective (e.g., Bechara & Damasio, 2005; Loewenstein et al., 2001), often using evidence from brain imaging studies to show the role of affect decision-making (e.g., Knutson et al., 2001; Knutson et al., 2007; Li et al., 2010). In the following three sections, I provide an overview of the influence of memory on decision-making and of two useful theories that focus on the role of affect during decision-making.

Memory and decision-making

Before discussing two useful theories that describe how emotion and decision-making may interact, it is beneficial to understand some of the processes that underlie and mediate decisions. In this section, I discuss human memory and how it is related to the choices and decisions we make. Indeed, learning and memory modulate the decisions we make under varying physiological, environmental, and temporal conditions.

Human memory can be described as distinct systems that are used to modulate and support human behavior (see Squire, 2004, for a brief history on this depiction of memory). The memory systems have been described using evidence from animal studies, human brain lesion studies, and (more recently) human studies with more intricate ways to measure the type of memory being used for a behavior (e.g., Schwabe & Wolf, 2013). Memory systems are typically described as some form of declarative memory (e.g., facts and events), procedural memory (e.g., habits), and emotional memory (e.g., emotional responses exhibited during certain pavlovian conditioning).

When making a decision, any of these memory systems may be in use, and the way a person (or animal) uses the memory to make a decision may be described as a certain strategy. Navigation-based tasks involving rats (e.g., Packard & McGaugh, 1996) and humans (e.g., Schwabe & Wolf, 2013) are often used as evidence for these different memory-based strategies; these studies have also been used to further understand how basal ganglia (procedural) and hippocampus (declarative) memory systems may interact. Other lesion studies have been used to understand decisions and emotional memory systems, and the brain structures typically involved in related responses (e.g., the amygdala, Fanselow, 2010; Maren & Quirk, 2004).

The understanding of these separate memory systems has led to the understanding of how destruction of these systems can lead to deficits in decision-making (e.g., Bechara, Damasio, Damasio, & Lee, 1999). From previous lesion studies, it is clear that the strategy used to learn and make a decision (i.e., the predominant memory system) can have consequences for decision-making performance. Computational simulation has been used as a tool to explore how strategies can affect decision and choice behavior given a different set of conditions (e.g., see Marewski & Link, 2014). Understanding what strategies are potentially being used during a task and being able to explore what it means to use varies memory systems over time to make a decision is paramount to understanding why we make the decisions we do and how we may make more optimal decisions.

Risk-as-Feelings hypothesis

Breaking away from the classic-economic expected utility perspective, Loewenstein et al. (2001) highlight the role of affect in judgment/decision-making in their risk-as-feelings hypothesis. They draw a distinction between the effect of feelings before a choice (anticipatory emotion) and anticipated emotions (outcomes) of the task in decision-making under risk. Anticipatory emotions are the immediate reactions to risks whereas anticipated emotions are those predicted to be experienced in the future. Figure 2-7 shows that these two types of emotions along with task-based subjective probabilities and current mood state feed into one's cognitive evaluation and feelings. Cognitive evaluation and feeling have a bidirectional relationship and together determine the behavioral outcome. Thus, they posit a process separate of (but not necessarily independent of) cognitive evaluation that affects decision-making.

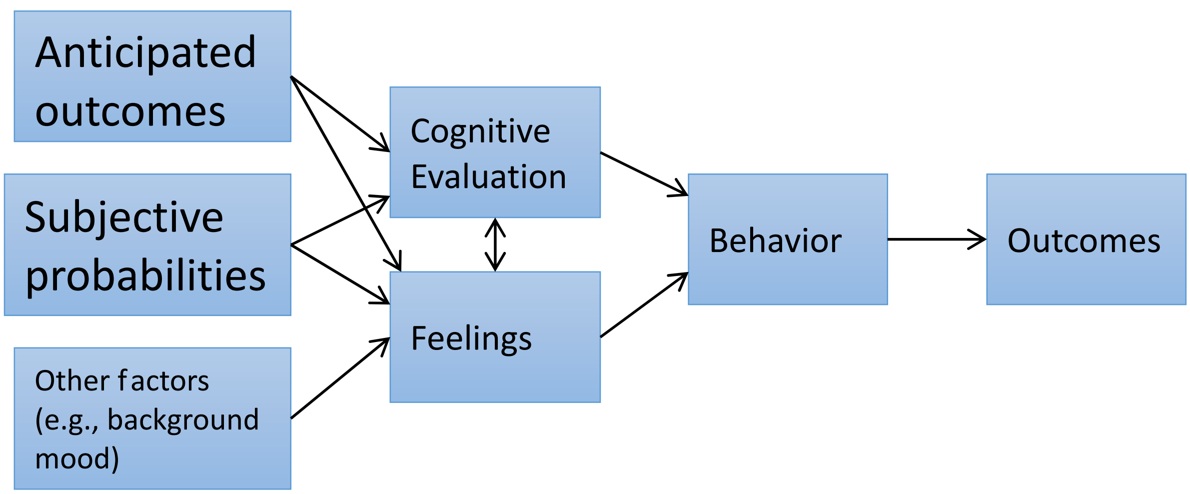

Figure 2-7. A conceptual model of the Risk-as-Feeling Hypothesis (adapted from Loewenstein et al., 2001)

The separate feeling category draws on results from past research on affect (e.g., Zajonc, 1980). The possible direct influence of non-cortical structures on behavior is represented by the paths from feelings to behavior in *Figure 2-7* that do not include cognitive evaluation. The importance placed on feelings in the risk-as-feelings hypothesis is similar to other theories such as the somatic marker hypothesis (discussed below), however this theory is argued to have a diverging focus by highlighting the direct effect of feeling on behavior (Loewenstein et al., 2001).

Though the risk-as-feeling hypothesis does present an attractive descriptive model of decision-making and behavior in risky situations, the theory itself seems to have less descriptive power than others like the somatic marker hypothesis. This is likely due to the reasoning behind the introduction of the theory and its focus on risky behavior. Loewenstein et al. introduced the theory as a reaction to typical expected utility and consequent-based theories in judgment/decision-making under risk literature. Thus this theory seems to have been well suited for its audience base and time, but as others (e.g., those using neuroimaging data) continue to evolve, it is likely that lines will become blurred and theories attempting to generalize will incorporate the ideas presented in this theory on behavior under risk.

Somatic marker hypothesis

Research on the effects of lesions to the ventromedial prefrontal cortex led Damasio (1994) to develop the somatic marker hypothesis (SMH), a theory on the effects of emotions, feelings, and somatic states on decision-making. The theory posits that emotion is advantageous when related to the task, but may be disruptive otherwise (Bechara & Damasio, 2005). The theory hinges on the concept of somatic states, which are collections of physiological responses that characterize an emotion. The SMH uses the Iowa Gambling Task (IGT) to experimentally test participant decision-making under ambiguity (Bechara, Damasio, Tranel, & Damasio, 1997).

The IGT is a gambling task meant to engage the processes involved in real-life decision-making. Participants must pick a card from 4 decks, two of which will give them a net loss (if they were to continuously select from that deck). The participants are never given the distribution of the decks and are told to just pick the cards from the deck they think will allow them to win the most money. Over the course of the experiment healthy-control participants seem to go through 4 stages: pre-punishment, pre-hunch, hunch, conceptualization (Bechara & Damasio, 2005). These stages are separated using performance and anticipatory somatic signal data (recorded in the form of electrodermal activity).

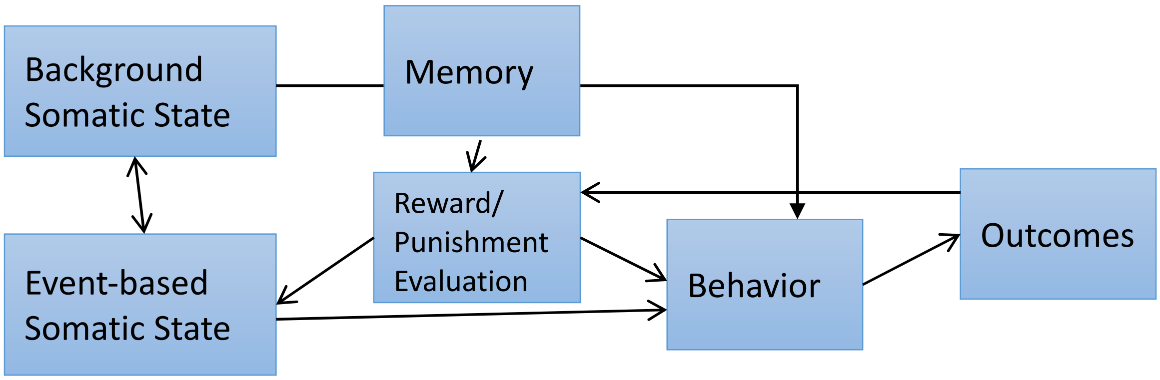

Figure 2-8. A conceptual Model of the Somatic Marker Hypothesis

Somatic states are caused by inducers that are regarded as either primary or secondary inducers; the difference between the two is determined by the core cause of the state. Primary inducers are stimuli (learned or innate) that cause a response state (aversive or pleasurable). On the other hand, secondary inducers are stimuli internal to the mind, i.e. recalled memories or imagined situations. Both primary and secondary inducer-related processing can be caused by one single external stimulus, e.g. a picture may elicit an immediate aversive (primary) response and corresponding recall of memories related to objects in the image, causing an additional (secondary) response (Bechara & Damasio, 2005). Secondary inducers are initially dependent on primary inducers, however, once learned, the dependence on the primary begins to decrease, i.e. elicitation of a somatic state due to recall of memory becomes independent of any external stimuli.

The validity of using the IGT as evidence for the SMH may be in question given that in healthy controls it has been more recently found that there may be sex-related differences in task-related decision-making behavior (Van den Bos, Homberg, & De Visser, 2013)[1]. Despite this and other criticisms of its theoretical underpinnings (e.g., Dunn, 2006; Maia & McClelland, 2004), the SMH remains a major theory of decision-making. More recently, several functional neuroscience studies have been carried out using the IGT that have given further credibility to aspects of the SMH (e.g., Li et al., 2010; Fukui, 2005). Others have used the SMH to describe the processes involved in psychologically pathological conditions (e.g., Verdejo-Garcia, 2009).

Summary

It is clear that purely rational expected-utility decision-making based theories provide a normative description of decision-making. The *process* of decision-making is important to consider is often the result of several cognitive (e.g., Gonzalez & Dutt, 2011), physiological, and affective processes. Physiological states (Bechara & Damasio, 2005) and affective feelings during risk (Loewenstein et al., 2001) affect decision-making even if the information in the task is fairly ambiguous; remembered states can also differ from actual experienced affective states, thereby making affect-based choice (e.g., Kahneman & Thaler, 2006) constrained by memory. Current emotional states as well as anticipatory states can change how one perceives the available choices and change out expected outcomes in decision-making. Rick and Loewenstein (2008) discuss the need for further research not only on expected (i.e., anticipatory) affect, but also immediate emotions. They highlight the “pressing need” for research that focuses on the consequences of immediate affect (whether integral or incidental) and its interaction with expected affect; consequently, they also discuss the need for further research on both integral and non-integral,

[1] This is intriguing, because many of the recent studies that have used the IGT have failed to explicitly factor in any sex-related differences in their populations (e.g., Aïte et al., 2013; Davies & Turnbull, 2011; Stocco, Fum, & Napoli, 2009; Turnbull, Evans, Bunce, Carzolio, & O’Connor, 2005) and it seems that this may be an important factor in the behavior exhibited during the task.

immediate affect. Moreover, affective and emotional states affect decisions through effects on the cognitive process making it important to understand how affect can change cognition, and consequently decision-making, over time. On the topic of computational simulation, Belavkin (2001) also discusses the need for an automatic way to implement "emotional changes" to information processing.

The experiments conducted and computational models developed in this dissertation research contribute to the filling of this void in the literature. A process model has been developed to make a more precise theory of behavior during the IGT than is typically provided by descriptive theories of human decision-making (see Colombetti, 2008; Marewski & Mehlhorn, 2011, for discussion on the issue of precision in descriptive models of decision-making); the model integrates and uses perceptual, declarative memory, procedural memory, and affective (including physiological) processes. With a more precise computational process model, it is more straightforward to have the model complete the same computerized version as actual participants: (a) to compare the model data with participant data; (b) to understand why participants may display certain behaviors during the task; and (c) to predict how decision-making behavior will change in response to environmental, physiological, and affective changes.

Computational cognitive architectures

Cognitive architectures provide a system in which to develop process models of human behavior and cognition with the foundation of a unified theory of cognition. These models provide an avenue to test theories that force a modeler to make quantitative selections about his or her theories and allow comparisons between theories. The complexities of the theories and the computationally implemented architectures suggests a separation of cognitive architectures into categories. Useful manners in which to divide cognitive architectures have been offered previously (Bach, 2009; Kennedy, 2011; Pew & Mavor, 1998; Ritter et al., 2003). *Figure 2-9* presents an adaptation made from categorical separation previously used to highlight differences among cognitive architectures.

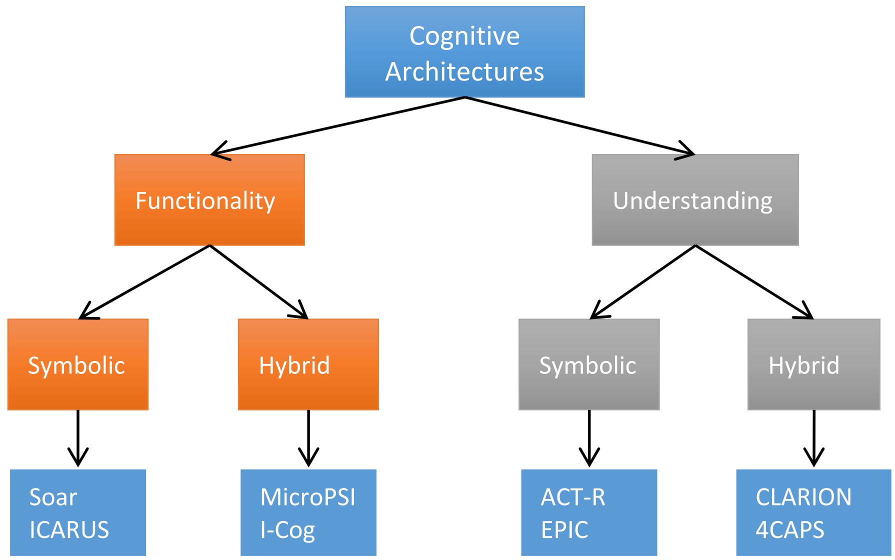

Figure 2-9. A conceptualization of some current cognitive architectures

At the highest level in his hierarchy, Kennedy (2011) conceptualizes cognitive architectures as *understanding* and *functionality* oriented. This likely has a foundation in not only the original ideas behind the theory and corresponding architecture, but also how the specific cognitive architecture has been progressed since its first computational realization. Within the understanding hierarchy are architectures classified as either *overt behavior* or *brain structure*; this further highlights the different focus of the separate category of architectures. Within the functionality hierarchy are classifications for *psychology*, *engineering*, and *philosophy*; this categorization represents the differences in design inspiration of the original architectures.

Bach (2009) takes a different approach to classifying cognitive architectures, he chooses instead to classify architectures based on the underlying representation structure. Bach separates cognitive architectures into *symbolic*, *hybrid* and *distributed* categories. The symbolic category houses architectures with a localist symbolic representation (e.g., rules), while distributed architectures are those with global representations and distributed processing (e.g., neural networks). As one may expect, hybrid cognitive architectures are those that implement some form of both representation and/or processing.

As previously discussed, Kennedy (2011) develops an initial separation of cognitive architectures based on the goal of the architecture. I use this same distinction to develop the same two initial categories: understanding and functionality (Figure 2-9). While the hope and overall intention of these architectures is to integrate the two, current and past design choices begin to add up; though understanding and functionality certainly are not mutually exclusive, one may have to sacrifice one for the other. Within these two categories I further categorize architectures with distinctions between symbolic and hybrid, similar to (Bach, 2009)[2]. I discuss ACT-R, Clarion, Soar, and MicroPSI – one architecture for each union of the categories previously mentioned.

ACT-R

ACT-R (Anderson, 2007) is a modular, symbolic architecture with the foundation of a unified theory of cognition. ACT-R was developed out of a perspective in psychology traditionally concerned with memory and reaction timing; thus the architecture is often compared with this type of experimental psychology data when researchers attempt to validate cognitive models developed in the architecture. Long-term memory in ACT-R is represented with both declarative (lists of facts) and procedural memory (realized as if-then style rules) implementations. The modules in ACT-R (6.0) typically used for the development of models of cognition are[3]: procedural, imaginal, declarative, goal, manual, audio, vision, and speech.

[2] Though many would consider ACT-R a *hybrid* cognitive architecture due to its symbolic and subsymbolic representations, I place it under the symbolic category to more closely match the definition given by Bach, 2009.

[3] Another module noted later in Figure 2-10 is the Temporal Module. Though it is not as typically used, I do think it deserves consideration for use by those who may be interested in modeling how time is internally perceived.

The modules in ACT-R encapsulate functionality to represent different aspects of human behavior and cognition. The imaginal module represents internal representation of perceptual information. The declarative module represents human declarative memory in the form of lists; this module is one of the two methods of long-term memory present in the system. The goal module keeps track of the models current goal state, often affecting the model's subsequent internal choice of procedural memory. The aural, visual, manual, and vocal modules all either receive information or send information to (or both) the external environment. All of the modules discussed connect into a central production system run by production rules. A higher-level view of the ACT-R system can be seen in Figure 2-10.

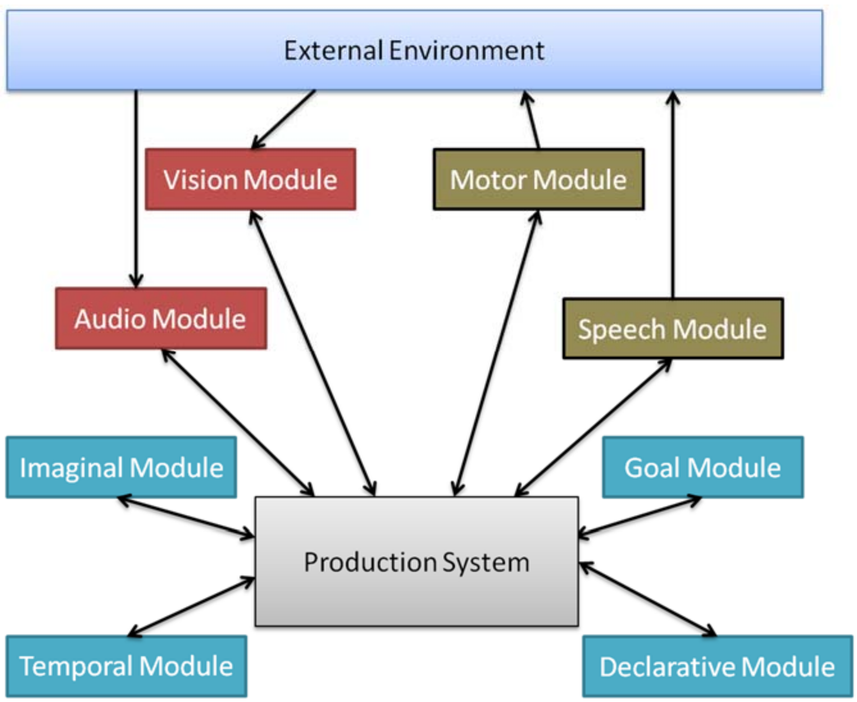

Figure 2-10. Overview of the ACT-R architecture.

Long-term memory in ACT-R is represented by procedural memory and declarative memory (the production system and declarative module, respectively). While both of these systems have grown out of a symbolic representation, i.e. rules and fact-lists, they also has subsymbolic components. Declarative memories have a base-level activation that determines the strength of the memory, independent of context and also strength value that is associated with the current context (what is in the buffers).

The base-level learning value (Equation 2-2) is determined using the time *t* since an item has been practiced. This learning value is further controlled by a decay value *d*. The decay value is typically set to 0.5 in models but has been lowered in models to simulate memory effects on a longer-time scale (Anderson, 2007); the summation of time *t* since *k*th practice of an item has been previously compared to the more neural concept of long-term potentiation (LTP, Pavlik Jr & Anderson, 2005).

Equation 2-2

$$B_i = \ln\left(\sum_{k=1}^{n} t_k^{-d}\right) + \beta_i$$

Equation 2-3

$$P(R_i) = \frac{1}{1 + e^{-(A_i - \tau)/s_i}}$$

Probability of retrieval of a declarative memory is represented by Equation 2-3. This probability equation uses the values of memory activation *A*, threshold activation value *τ*, and association strength *s*. Any memory with an activation value below the threshold will fail to be retrieved under any circumstances. The associative strength is set by default to represent effects found in the fan experiment (Anderson & Reder, 1999).

To connect the theoretical underpinnings of the ACT-R architecture with (central) physiology, modules in ACT-R have been correlated to structures in the brain (Anderson, 2007; Anderson, Fincham, Qin, & Stocco, 2008). Perhaps most important for the purposes of this dissertation are the brain regions related to internal module activity, i.e. the declarative, procedural, imaginal, and goal modules. Figure 2-11 shows the brain regions that have been associated with activity in ACT-R modules.

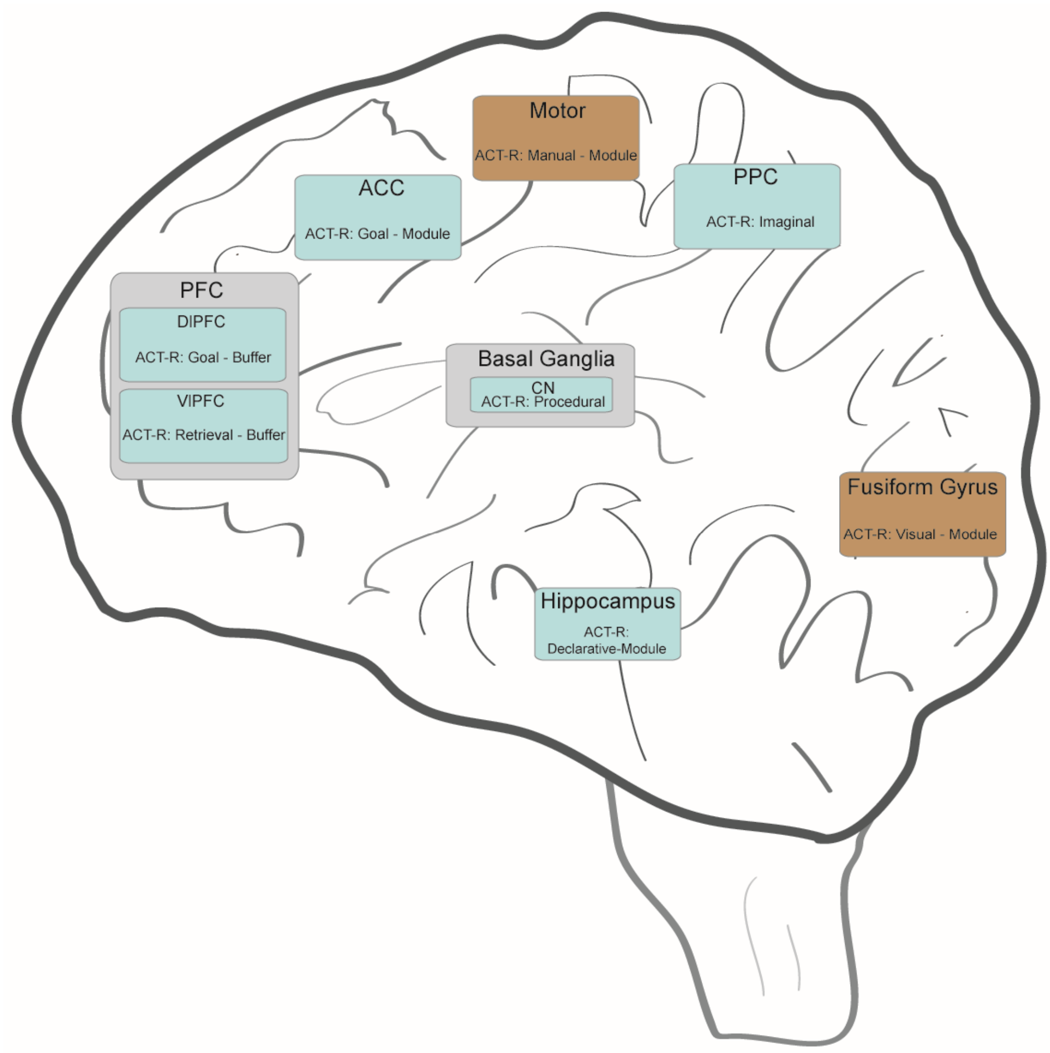

Figure 2-11. Brain regions labeled with their associated ACT-R modules. Abbreviations: ACC – anterior cingulate cortex; CN caudate nucleus; DlPFC dorsolateral prefrontal cortex; PPC – posterior parietal cortex; VlPFC ventrolateral prefrontal cortex

Activity in the retrieval buffer (declarative module) corresponds to activation in the ventrolateral prefrontal cortex. This area of the brain is hypothesized to hold cues to information elsewhere in the brain (Anderson, Fincham, et al., 2008). Thus, one would expect that the continual retrieval of information, say due to an incorrect answer, necessitates an increase in the activation of this area, because it must continue to hold cue information for retrieval of the correct memory; consequently, this effect has been found by Anderson and colleagues in previous studies.

Goal module activity is connected to activation in the anterior cingulate cortex. This module allows one to control cognitive activities independent of the external environment as it represents an internal control state. The anterior cingulate cortex has been argued to play a key role in conflict monitoring (Kerns et al., 2004; Walsh, Buonocore, Carter, & Mangun, 2010), this matches results found that indicate an association between ACC activation and task difficulty (Anderson, Fincham, et al., 2008). The dorsolateral prefrontal cortex is also associated with activity in the Goal buffer (Anderson et al., 2004).

Activity in the procedural module corresponds to activation in the thalamus and the caudate nucleus of the basal ganglia (BG). The BG are comprised of the caudate nucleus, putamen, nucleus accumbens, globus pallidus, and olfactory tubercle. The specific portions of the BG found corresponding to activity in the procedural module and production system are the caudate nucleus, putamen, globus pallidus and substantia nigra (Anderson, Fincham, et al., 2008). The afferent connections to the basal ganglia from the cortex and efferent connections (from the BG) to the thalamus (which has efferents to the cortex) provide an obvious connection to the procedural module; the basal ganglia has long been thought of as the key set of structures for skill learning (e.g., see LeDoux, 1996).

The imaginal module (problem representation) is associated with activity in the posterior parietal cortex. Change to internal representations of the problem is assumed to correspond to activation in this area of the brain. Furthermore, this particular region of the parietal cortex seems to respond to alterations of difficulty of problem representation (Anderson, 2007; Anderson, Albert, & Fincham, 2005; Anderson, Byrne, Fincham, & Gunn, 2008).

The associations between computational modules and areas in the brain provide a method to connect physiological and cognitive perspectives of human cognition and behavior. The open-source and modular setup of the architecture also allows one to develop new modules that affect only certain parts of the system; a good example of the use of this modular setup is one of the modifications made in the transition from ACT-R 5 to ACT-R 6.

These associations between neurological structures and psychology processes make representing and simulating the interactions between peripheral physiology and central physiology more tractable. Thus, one can now to integrate theories of cognition and those that involve cognitive modulators (e.g., affect) in a more straightforward manner with a more clear understanding of how physiological constraints may affect this modulation of behavior.

CLARION

CLARION (Sun, 2006) differs from ACT-R in that it employs what Bach (2009) would call a *true* hybrid architecture, i.e. localist and globalist representations. Each of what Sun describes as subsystems have both implicit and explicit representation structures. There are four separate subsystems: action-centered subsystem (ACS), non-action-centered subsystem (NACS), motivational subsystem (MS), and meta-cognitive subsystem (MCS). Though the CLARION architecture is centered on simulating human intelligence with a goal of understanding human behavior and cognition, it currently does not provide the same mapping from architecture to brain as ACT-R (e.g., Anderson, 2008b). It may be argued that one could accomplish a (albeit different) mapping as long as the representation is sufficient for the given level of abstraction in Clarion (as discussed by Simon, 1996).

Clarion's action-centered subsystem controls any action taken by the agent, whether the action is external or internal). As with all subsystems in CLARION, the ACS has explicit and implicit representations and each are combined after the ACS decision-making process has finished. The current state of the world, represented symbolically (explicit) and as inputs to a neural network (implicit), is used to compare possible implicit and explicit actions. An action is then selected from the actions available with each representation. The non-action-centered subsystem holds general knowledge about the world not represented by action knowledge (Sun, 2006). This subsystem would relate more to the declarative memory system in ACT-R (but again it differs in that it provides local-symbolic and global-neural-network representations). A chunk node (symbolic representation) is connected to a distributed set of features represented in a neural network.

The MS is used to simulate underlying drives, similar in principle to the urges presented in MicroPSI (Bach, 2009; discussed below), and goals. Drives receive input from perception and are separated into low-level primary, high-level primary, and secondary goals; the goal structure receives input from goal actions determined by the ACS and MCS. The low-level drives are *Get-Food*, *Get-Water*, and *Avoid-Danger* (Sun, 2006). The architecture also has high-level drives that represent more social motivations including esteem and belongingness. Secondary drives are those acquired after an agent experiences the world, i.e. non-innate drives.

The CLARION architecture provides a different conceptualization of cognition than ACT-R and it perhaps could be argued that it abstracts problem representations at different levels than other architectures like ACT-R and Soar. Especially noteworthy is the motivation representation subsystem, something that is represented in some (e.g., MicroPSI), but certainly not all (e.g., Soar and ACT-R) cognitive architectures. Though the theoretical abstraction of the system provides an attractive representation to validate with behavior modeled after experimental data, the lack of a fully functioning architecture (e.g., the NACS subsystem has yet to be realized in the actual software implementation) makes it difficult to really assess the CLARION architecture and compare to others. This unfinished computation system also makes it difficult to speculate about possible extensions as implementation can differ from theory in nuanced ways.

Soar

Soar (Laird, 2012) is a cognitive architecture that is centered on the goal of providing agent functionality with cognitive plausibility; matching human data is not a primary goal of Soar. Thus, while the Soar cognitive architecture does have an underlying unified theory of cognition (Newell, 1990), it also has significant influence from goals in the AI domain (e.g. models of general intelligence). In the past Soar has relied on production rules as its sole form of long-term memory, however more recently it has been extended to provide explicit symbolic representations for long-term declarative memory (Laird, 2008).

Traditional long-term memory structures in Soar (i.e., pre Soar 9.0) consisted of procedural knowledge implemented as if-then style rules. Soar models select an operator based on a condition that (usually) results in an action that retrieves information to be placed into a short-term memory buffer. The amount of information that can go into this buffer is not inherently limited by the architecture and thus must be limited explicitly by the modeler.

Short-term memory is represented in a symbolic graph structure that allows the representation of symbol properties and relations. The short-term memory structures are used by the model for checking conditions of the current state of the models representations within a given decision-cycle. Within Soar, rules are used to propose, evaluate, and apply operators. Thus a rule is used to propose an operator which creates a symbolic structure in the short-term memory that has an associated condition for selection.

To provide a computational representation of emotion based on an appraisal theory (Scherer et al., 2001), Soar also provides a new module for representing the functional changes emotion has on cognition based on an appraisal theory (Marinier III, Laird, & Lewis, 2009). This module presents the only full integration of appraisal-based emotion (into a cognitive architecture) of this kind; though others have developed computational models of emotion with

Soar (Marsella & Gratch, 2009), it has been argued that these do not provide the full integration presented with the Soar 9 appraisal detector module (Marinier III et al., 2009).

Soar provides an architecture that straddles the line between a focus on human cognition and behavior and a focus on general agent intelligence. Though it has a foundation in human cognition, it is less concerned with simulating and replicating human experimental data than other architectures like ACT-R. Nonetheless, Soar continues to present a cognitive architecture that is attractive for developing intelligent agents organized based on a representation of human cognition. Of particular interest for this research is the appraisal module that should be used for insights for modeling cognitive appraisal of emotional stimuli.

MicroPSI

MicroPSI is an agent architecture developed based on the PSI theory by Dietrich Dörner and realized by Bach (2009). Its components can be separated into six main groups: sensors, execution space, long term memory, access memory, maintenance script space, meta-management, and external behaviors. I discuss the sensory and memory facilities of the architecture.

External sensors provide information that is compared to internal expectations and mapped into memory accordingly. If there is no expected internal representation that can be matched to the external situation, then a new object schema is created (Bach, 2009). Agents obtain an account of abstraction by comparison of commonalities between separate schemas.

The sensor components not only provide a picture of the external world, but also provide a representation of internal sensory components. A group of emotional modulators exist that specifically modulate how internal behavior is carried out. These three modulators (arousal, resolution level, and selection threshold) affect allocation of mental resources given the context of a specific situation.

There also exists a motivation system that affects the internal behavioral (e.g., goals, planning, and/or behavior selections) component of the architecture. This motivational system provides representations for physiological-based and cognitive-based *urges* as well as a social-based urge. Physiological-based urges include intactness, food, and water. The urges on the cognitive and social level are competence and reduction of uncertainty, and affiliation (respectively). Causing an increase in these urges creates an associated negative signal within the architecture while a reduction in the urge will create a corresponding positive (i.e., pleasure) signal.

This sensory modulator component provides an explicit representation of important moderators to cognition. While the physiological urges are fairly simple representations of the homeostatic motivations they represent, the urges provide fully integrated moderation of (functionally) cognitive mechanisms. These (and other) urges also provide a mechanism of autonomy, limiting problems commonly associated with models developed in cognitive architectures, like perseveration.

Memory in MicroPSI is represented with networks of nodes stored in a hierarchy. Activation of networks is controlled by *actuator* nodes. Nodes can have both sub and super relationships with other nodes in memory. Memory nodes on same level of the hierarchy instead have temporal or spatial relationships. Links between nodes also exhibit time and usage-based decay, partially determined by the *node-space*, i.e. different node-spaces inherently exhibit a stronger or weaker decay speed.

Long-term and access memory systems within the architecture have similar representations, the only difference being a *node-space* distinction; this just allows one to create groups of nodes with a (architecturally) functional similarity. The basic unit of a MicroPSI representation is a *net-entity*. Thus, net-entities are used to represent all actual node implementations in the system. Each net-entity is connected to another net-entity via a gate (output) and a slot (input).

The hybrid representation in MicroPSI (i.e. symbolic and subsymbolic) is presented as a more implicit distinction; some representations are more *local* (symbolic) while others are predominantly *global* (subsymbolic) within the network (Bach, 2009). A hybrid symbolic-subsymbolic network representation has consequences for the manner in which one must program actual agents. This implementation would likely prove to be very difficult in a text-based editor for more novices. MicroPSI agents are developed using a graphical language, with pictorial representations and links that mirror the ideas presented above. This is in contrast to architectures like ACT-R that use text-based editors to represent models; though higher-level languages to represent certain symbolic architectures have been developed separate of the actual architecture development (e.g., Cohen, Ritter, & Haynes, 2010).

MicroPSI is a symbolic-subsymbolic hybrid architecture that provides motivation-based and cognitive representations. Though it is self-admittedly not likely to result in an "accurate model of the mind" (Bach, 2009), it does a fairly good job at straddling the middle ground between higher-level cognition and underlying affect that moderates and modulates cognition. A more integrative representation of physiology would be desirable to simulate a wider range of physiology's modulation of cognition. Still, the integration of the representations of physiology is an important step in the evolution of cognitive and agent architectures; this leads modelers to more consistently acknowledges the pivotal role physiology plays in the mind. MicroPSI shows that one can model diverse behavior with autonomous intelligent agents using physiological and emotional motivations in non-cognitive (agent) architecture.

The next logical steps for anyone wishing understand how systematic changes in physiology affect cognition are to incorporate these same motivating and emotional factors in a cognitive architecture based on an underlying UTC and extend this cognitive architecture with a more complex representation of physiology. This type of extension would move the architecture towards the ability to represent more diverse behavior and psychological theories computationally.

Summary

These architectures all accomplish a level of agent intelligence with differing abstractions and representations of cognitive processes, motivation, and behavior. If one is only concerned with providing intelligent agents with a representation with psychologically plausible foundations, those built in architectures like Soar and MicroPSI will likely provide the facilities for simulating intelligent behavior along with interesting differences depending on conditions like strategy choice and parameter settings. However, if one is concerned with developing computational models that match human experimental data, then it is likely one would select an architecture like ACT-R or CLARION that are centered on understanding human behavior with models built within the architecture.

Due to its wider community base, longer maturity, and it having a system more fully functioning with the theory, ACT-R is the most attractive choice for modeling human behavior and cognition. Motivating and emotional factors from other architectures like Soar and MicroPSI could be used to extend ACT-R. Models in computational physiology are attractive to provide a more complex representation of physiology to ACT-R.

Past accounts of physiology, affect, and cognition in computational architectures

CoJACK

CoJACK is an extension of the Java Agent Construction Kit (JACK), a beliefs, desires, and intentions architecture (Ritter et al., 2012). This extension added activation and noise, i.e. cognitive constraints, to knowledge representations within JACK. Belief-sets (CoJACK) are similar to declarative memory chunks (ACT-R) and consequently CoJACK agents can retrieve incorrect (or even fail to retrieve) belief-sets; a similar addition is made to the intention representation (roughly mapped to ACT-R's rule representations). CoJACK adds a representation of time and implements two overlays that moderate the altered representations: caffeine and fear.

Both moderators have a reservoir to represent the amount of effect the moderator should have on its connected components (Ritter et al., 2012); these reservoirs decrease over time based on a decay function. How the respective reservoir amount affects cognition and behavior is determined by mathematical models built into the architecture. Though these moderators can produce interesting behavior when used in simulations (e.g., Evertsz et al., 2009) the behavior resulting from the moderators have yet to be validated.

CoJACK offers several microcognition (low-level) elements with an overall macrocognition (high-level) representation. This combination may improve the ability to develop expert agents within psychological plausibility. CoJACK also offers behavioral tracing tools that are likely useful for providing explanations of agent behavior (Evertsz, Ritter, Busetta, Pedrotti, & Bittner, 2008; Ritter et al., 2012). Though it has its shortcomings (e.g., lack of validation of implementation) the architecture offers insights on how one may proceed when extending any existing architecture.

EMA

The Emotion and Adaptation (EMA) model is a computational process model of emotion that uses appraisal theory (C. A. Smith & Lazarus, 1990) to constrain and provide its definition of emotion and how the dynamics of emotional appraisal affect and result in a different behavior (Marsella & Gratch, 2009). EMA is interesting because it represents one of the most integral process models of emotion and uses the Soar architecture (discussed in the section on Computational cognitive architectures) and consequently adopts the constraints inherent in that architecture. The EMA model has two levels of representation that determine an agent's emotional state, how that state affects cognition, and how this interaction results in dynamic behavior over time: (a) representation on an appraisal level that determines moment-to-moment changes in an agents emotional state; (b) representation on a higher mood level that describes a background state that affects and biases appraisal dynamics and consequently behavior.

On the appraisal level, there are several *appraisal variables* that are used to describe an event and cope with the event: relevance, perspective, desirability, likelihood, causal attribution, controllability, changeability, and expectedness. Continuously active feature detectors map events to the different appraisal variables. Features of an event are mapped based on time-based propositions (past, present, and future) that result in a causal interpretation of the event in terms the specific propositions (e.g., *I know, from the past, that bears, which are found to be in close proximity to humans, can harm the humans*). Each of these causal interpretations are *appraisal frames* that all compete in parallel for the attention of the coping response.

The higher-level mood representation describes an aggregate of the original appraisal frame that changes relatively slowly and is disconnected from the actual events that are the subject of the appraisal frames. The mood state is represented by explicit emotional labels (e.g., fear, joy, etc.) with intensity values. The background mood-level for each emotional label changes an appraisal frame by adding to the appraised value of that emotional label (Marsella & Gratch, 2009).

The EMA model is interesting in that it is very similar in some ways to conceptualizations made by theorists from the physiological-anatomical perspective (e.g., LeDoux, 2012; Panksepp & Biven, 2012), but it explicitly does not address the ways physiological changes can result from and affect behavior (Marsella & Gratch, 2009) which can have notable effects over time (e.g., see Berthoud & Münzberg, 2011; Joëls et al., 2011; McMorris, 2009). Thus, the perspective is a higher-level take on affective behavior and does not provide a very straightforward mapping between neurophysiological data from studies of affective/cognitive behavior and its own representations; essentially EMA's atomic representation are on a higher, more psychological, level than those provided in many neurophysiological studies of behavior (e.g., see Feldman Barrett, 2006a; Russell, 2003 for a discussion on atomic representations of emotion). However, EMA should not be seen as competing with models that provide a more physiological-anatomical perspective, it should be seen as complementary. The EMA model provides an example of how one may develop an affective process model of behavior (albeit at from a different perspective than used in this dissertation) that changes over time and is continuously affected by the environment.

ACT-R with CNPA

Gunzelmann and colleagues have worked on cognitive models (developed in ACT-R) that simulate effects of fatigue and circadian rhythms on cognition (Gunzelmann et al., 2009). The model changes procedural rule activation values based on a linear equation mapping to data obtained from a mathematical model of alertness. A psychomotor vigilance task (PVT) was modeled to compare to existing experimental data to validate predictions made by the model.

A manual parameter sweep[4] on the *G* parameter (conceptualized as alertness) and the utility threshold (determined if a procedural rules utility is high enough to be considered for retrieval) provided parameters to match performance under maximum and minimum alertness. The authors then matched these with corresponding high and low outputs from the two biomathematical models used to develop a linear transformation from the alertness model output to corresponding parameter value (Gunzelmann et al., 2009). These equations were then used to change parameters over time based on predetermined output values obtained from running the biomathematical models; a linear relationship was assumed between values. These changes yielded results very comparable to human experimental data in the conducted PVT.

This work displays that one can successfully have output from biologically based mathematical models modulate different aspects of a cognitive architecture. The method has also been used to model the effect of sleep and fatigue on more complex tasks such as driving (Gunzelmann, Moore Jr, Salvucci, & Gluck, 2011). Though the model is promising, it does lack certain generalizability due to its use of a mathematical model representing a single specific aspect of functioning (and high-level) physiological behavior. There also is a lack of feedback between the ACT-R model and the biomathematical models, meaning an actual real time connection between systems was not instituted. The biomathematical model output was determined independent of the actual model behavior; this may explain why Gunzelmann et al. (2011) were unable to simulate certain spikes in driver alertness.

It would be useful to apply some of the same ideas and principles used in this line of research to the connection of ACT-R and a more comprehensive mathematical model of physiology that gives output and receives certain feedback in real time, thus a two-way connection. As an example, one could use the CNPA model to drive HumMod physiological

---

[4] Here, by manual parameter sweep, I mean Gunzelmann et al. (2009) ran the model with different parameter values until they found the values that resulted in the lowest RMSE between the model and human performance for different times of the day.

variables to simulate changes in physiology due to circadian rhythms. This would allow one to model the same processes captured by the previous ACT-R CNPA model, but also potentially capture certain spikes in physiological variables (e.g., due to being exposed to a highly stressful stimulus) that would affect behavior for a short period of time.

Summary

As cognitive architectures continue to mature and move past software implementation issues, it is likely that more robust representations of moderators will begin to appear. Gunzelmann et al. (2009) offers insight on the possibility of tying ACT-R parameters to an external biomathematical model while Ritter et al. (2012) display the power of extending an existing higher-level architecture with cognitive limitations based on a cognitive architecture and adding moderators to this implementation. Another particularly useful representation in CoJACK is the representation of caffeine's effect on cognition with an implementation that is informed by existing pharmacokinetics research.

However, these implementations do have flaws. The Gunzelmann model fails to implement a bidirectional connection with the biomathematical model, while CoJACK has yet to be validated in a manner similar to ACT-R. Despite these drawbacks, all architectures offer insight into ways one may implement a connection between physiology and cognition. If one wishes to provide a fairly robust representation for the implementation of moderators, the connection of a mature cognitive architecture to an integrative computational physiological model seems the most useful way to implement a system that connects cognitive and physiological processes.

Computational physiology

While cognitive architectures can be used to develop a wide range of agents and models, the underlying physiology of the behavior often remains, at most, an afterthought. While one may accomplish some behaviors using overlays (e.g., Evertsz et al., 2008), parameter changes (e.g., Ritter et al., 2009Gunzelmann et al., 2009), or more specific urges/motivations (e.g., Sun, 2006; Bach, 2009), these methods can be improved so that models can better generalize between tasks physiological representations or overlays into one complex architectural conceptualization. Computational physiological models present a possible way to represent underlying functional physiology in cognitive architectures.

Though several models of physiology exist, to be useful as an extension to a cognitive architecture, the model should be fairly robust and integrative, i.e. it should represent an amount of functional physiology that would be difficult to build as an architectural overlay. I briefly review two possible candidate physiological model/simulation representations to be used to modulate architectural behavior: CellML and HumMod.

CellML

CellML (Hedley et al., 2001) is an XML-based language architecture that allows modelers to represent physiological components and connections as a complex system and provides a standard for the physiological modeling and simulation community. The CellML language representation allows modelers to combine the CellML model with other XML-based modeling languages like ChemML (Cuellar et al., 2003). Because it is a language standard designed to be used with multiple simulation systems, CellML does not provide an explicit simulation system, instead other simulation systems (e.g., those being developed by the International Union of Physiological Sciences Physiome project) must be used to simulate the model developed in

CellML; an API also exists designed to compile CellML model code into a schema more simulation ready (Miller, 2010).

Though CellML models and corresponding simulation systems, like the mentioned physiome project, present a potentially fruitful idea that could grow to provide a very rich and robust combined model and simulation package, at its current stage the projects involving CellML fail to provide integrative models of physiology. This lack of an integrative model makes it more difficult to provide a fairly full representation of functional human physiology to connect to an existing cognitive architecture.

HumMod

HumMod (Hester et al., 2011b) is a simulation system that provides an integrative model of human physiology that is a derivative of the Guyton model of integrative physiology (Guyton et al., 1972). This model provides over 6,500 state variables and over 1,500 linear and non-linear equations to represent physiology in a top-down fashion, i.e. the model begins representation at a gross anatomical level and provides further abstraction as needed. The HumMod model is built using a form of XML, making it open to modification, extension, and external validation. The model also provides several points of access to the physiological variables through parameters that can change many aspects of the physiological output including output related to both the endocrine and nervous systems. Figure 2-12 presents some of the major systems in HumMod and example variables.

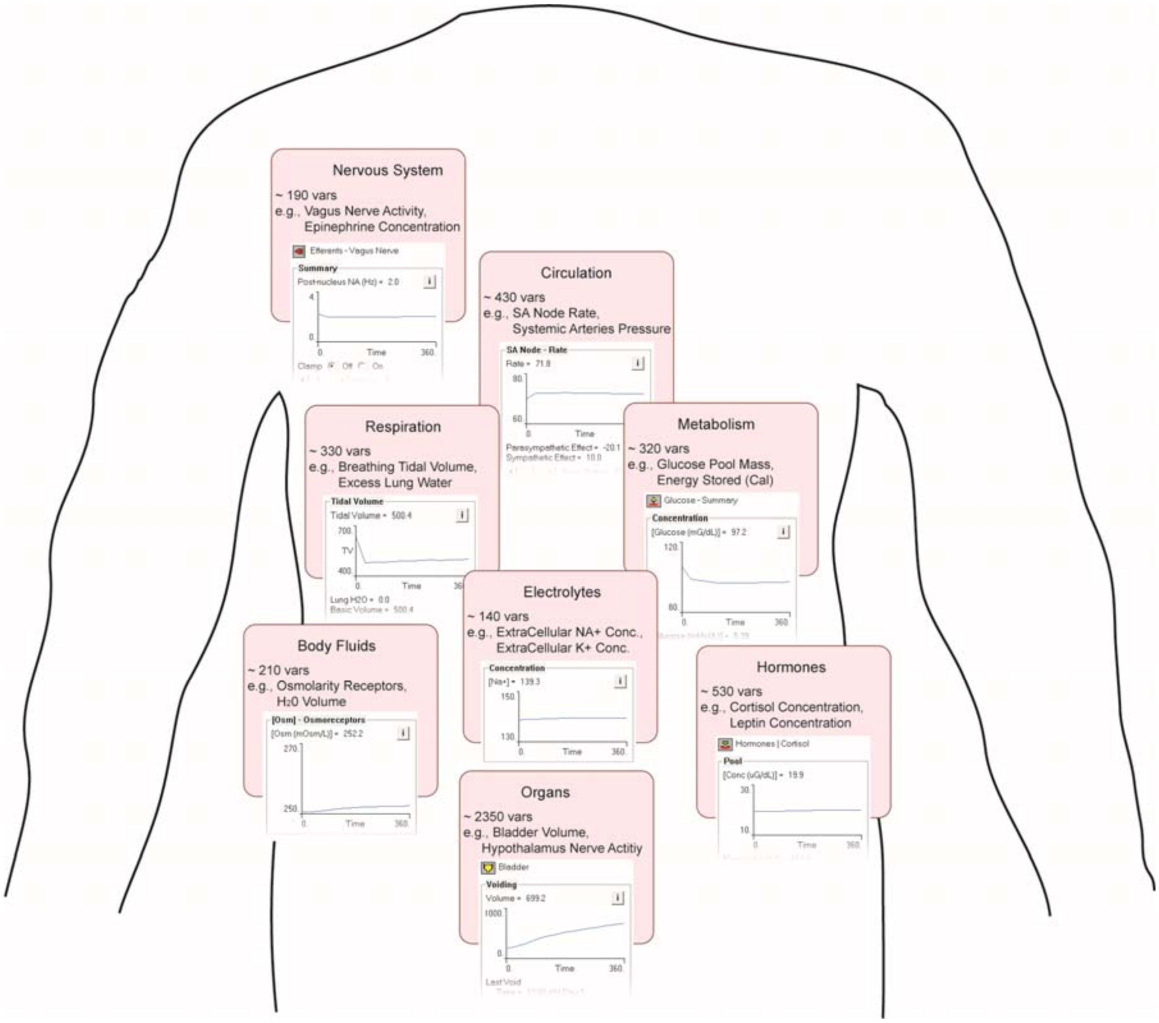

Figure 2-12. An illustration of some of the physiological systems and example variables of those systems in HumMod

The HumMod simulation system allows one to simulate physiological changes over time using a processing (within the system) called time slicing. Using time slicing, one can run a model from increments of one second to three months. Thus one can compare how time dependent physiological responses occur in a short amount of time (e.g., seconds) to a relatively long amount of time (e.g., months).

If one wishes to skip the interface and manipulate the variables directly, message passing can be used to communicate directly with the system's model solver. One can pass text-based messages to the model solver and it will parse these messages. Using strictly message passing allows the user to increment the model's independent variable, i.e. time, in a custom manner. One can also use this message passing to get and set any variable the modeler wishes to affect, though it is likely that the user will want to follow the basic outline of the normal HumMod interface, only changing variables that are not directly affected by others. If one does wish to affect a variable that has inputs from other variables it would be possible to add a new variable to the model that connects directly to the state variable that one wishes to affect. One could also potentially alter initial variable values within the model's XML to simulate certain effects.

Summary

The HumMod system offers the only integrative model of human physiology at its level of comprehensiveness (Hester et al., 2011a). Other projects involving the CellML modeling language seem promising but have yet to produce a model and simulation system comparable to HumMod. The top-down representation of HumMod also allows a more smooth transition from experimental results to simulation interpretation. The integrative nature of HumMod that encompasses functionality within an interconnected system (as opposed to micro models of computational physiology) moves towards the unified argument made in psychology by Newell (1973). At the current state of computational models of physiology, HumMod offers the most attractive human physiology model and simulation system for the exploration of physiological and cognitive connections within a cognitive architecture.

Review summary

When understanding the ways physiology and affect can potentially modulate cognition, it is important to have a fairly clear definition of (or way of defining) the underlying process of behavior. The primary process affect theory by Panksepp and Biven (2012) is useful for understanding how physiology and affect mediate and modulate cognition as it is founded in neuroscience work that grounds the theory to the constraints of physiology. This theoretical framework is also useful because it describes systems that are predominantly separate from the cognitive systems (e.g., memory systems) that are typically explored in neuropsychology. The modularization of affective and related cognitive systems leads to an understanding of the process behind affective and emotional behavior, in addition to the actual outcomes of the behavior.

In understanding decision-making and choice behavior, it is also important to understand the process behind the behavior. Similar choice outcomes can potentially be the result of drastically different processing, whether in physiological, affective, or cognitive systems. It is useful to have an understanding of the physiology that implements and modulates the affective and cognitive processes that underlie decision-making behavior, because this can lead to a better ability to predict future choices and decisions given certain internal and external environments.

Cognitive architectures provide a unified theory of cognition (UTC) that can be used to understand and explore the processes that lead to decision-making and choice behavior. These architectures provide certain constraints (dependent on the theory) and are typically instantiated in computational software that makes theoretical constraints and predictions more explicit and testable. ACT-R is the most attractive UTC because it has a strong foundation in experimental data and the computational system that implements the theory is stable and can be modified in any way permitted by the programming language used to implement the software. ACT-R does not have any explicit representation of physiology or affect, but it is modular, which means physiological and affective components can be added to the system in pieces without changing every aspect of the underlying processing. The open-source nature of ACT-R and the modularity of its components make the extension of the architecture with representations of physiology and affect relatively straightforward.

Computational models that integrate some affective and physiological representations with a computational system have been developed in the past. In addition to the MicroPSI architecture that has been built with some simple physiological and motivation representations, EMA integrates a model of emotion (appraisal) with the Soar architecture, and CoJACK adds representations of fear and caffeine to JACK (a BDI architecture). ACT-R has also been connected to a model of fatigue (CNPA) to simulate some effects of sleep deprivation on cognition. All of these computational systems work well for the given task, but these models remain difficult to generalize and further extend (especially the physiology) because the underlying representations have developed with a specific task or goal in mind. An integrative model of physiological, affective, and cognitive processes would be more useful and easier to extend as more experimental data from physiology, neuroscience, and psychology domains becomes available.

Several computational models of human physiology exist, however, very few provide an integrative model of physiological processes that result in and from human behavior. HumMod is a useful computational simulation system that provides an integrative model of human physiology because the model provides a top-down representation (e.g., organs and hormones) that makes modeling experimental data in the system more straightforward. The system can also be used to see how changes in parameters (e.g., epinephrine levels) can have systematic effects on physiology.

The next chapter presents a hybrid computational architecture that models how physiological processes implement and modulate affective and cognitive processes. The architecture (ACT-R/Φ) extends the ACT-R cognitive architecture with modules that represent some of the affective systems posited by Panksepp and Biven (2012) and a representation of physiology from the HumMod simulation system. I describe the underlying descriptive model of the architecture as well as different computational components of the architecture.

Chapter 3 Integrating physiological, psychological, and neurological theories into a computational architecture

Introduction

In understanding how physiology, affect, and cognition may interact and result in dynamic behavior, it is useful to construct an integrative model of how these different levels of behavior may interact. It is especially useful when considering the range of disciplines tackling similar aspects of behavior, but from different, though often related, perspectives. Given the complexity of the behavior represented on these levels, it can be difficult to develop a model that's coherent and tractable. Fortunately, descriptive and computational models of human behavior on the physiological, affective, and cognitive levels exist; typically models on these different levels have not been integrated, especially into integrated computational models. As reviewed in the previous chapter, several candidate models exist that would be useful to integrate.

I suggest extending a unified theory of cognition (UTC) with an integrative model of human physiology to allow straightforward and tractable connections between physiology and cognition. Though some direct connections between physiological and cognitive processes do exist, I also propose that extending this same UTC with a neuropsychological theory of affect to provide an architectural account of affective and emotional behavior that also mediate some physiological influences on cognition is useful. Connections between these three models are also supported by existing models from neuroscience, psychology, and stress research.

The architectural model behind ACT-R/Φ

The architectural model underlying the ACT-R/Φ hybrid architecture provides a synthesis of work on the systematic effects of stress and arousal (e.g., Aston-Jones & Cohen, 2005; De Quervain et al., 2009; Joëls & Baram, 2009; Joëls et al., 2011; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012), work in affective neuroscience (e.g., Berridge et al., 2009; LeDoux, 2012; Panksepp, 1998; Panksepp & Biven, 2012), and work in cognitive architectures (e.g., Anderson, 2007; Anderson, Fincham, et al., 2008; Ritter, Reifers, Klein, & Schoelles, 2007) to construct a useful architecture that can be used to understand specific human behavior (Figure 3-3).

Several physiological structures and modulatory hormones that modulate memory, learning, and decision-making are also involved in stress systems (Figure 3-1). As discussed earlier, two major systems implicated in stress are the HPA and SAM axes.

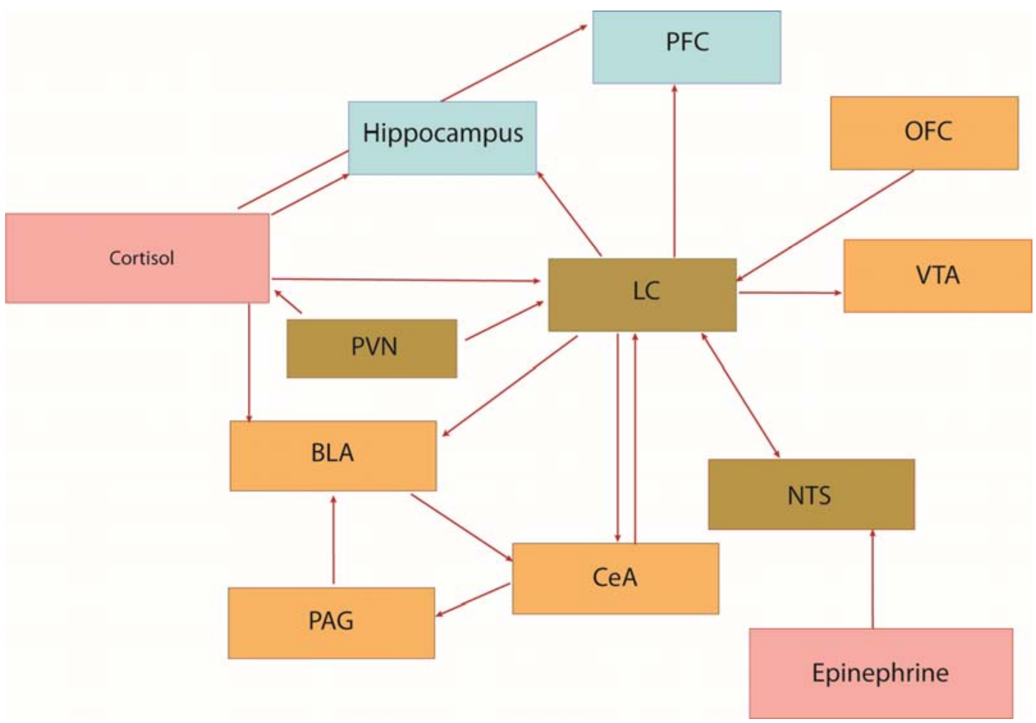

Figure 3-1. Some neural structures that are involved in cognition, affect, and stress response/regulation, and functional connections between the structures. Acronyms: BLA – basolateral amygdala; CeA — central amygdala; LC — locus coeruleus; NTS — nucleus of tractus solitaries; OFC — orbitofrontal cortex; PAG – periaqueductal gray; PFC — prefrontal cortex; PVN — periventricular nucleus; VTA – Ventral Tegmental Area

Activation of the HPA axis is characterized by a cascade of central and peripheral hormonal changes principally beginning with activity in the periventricular nucleus (PVN) that results in release of corticotropin-releasing hormone (CRH), and ending in the peripheral release of cortisol that is caused by a peripheral increase of adrenocorticotropic hormone (ACTH). Cortisol can pass through the blood-brain barrier and thus directly affects both the peripheral and central nervous systems. After it crosses the blood-brain barrier, cortisol affects declarative memory formation by interacting with neural norepinephrine (from the LC and nucleus tractus solitarus or NTS) and modulating declarative retrieval and encoding (Schwabe et al., 2012).

Activation of the SAM axis is characterized by central release of norepinephrine via the LC and NTS, as well as peripheral release of epinephrine. Importantly, this release of epinephrine affects norepinephrine release via afferent vagal pathways that modulate the NTS and LC (Cahill & Alkire, 2003; Miyashita & Williams, 2006). The LC is also affected by CRF primarily from connections originating from the CeA and PVN (Reyes, Carvalho, Vakharia, & Van Bockstaele, 2011; Sara & Bouret, 2012), and the CeA is an important structure in the FEAR system (Panksepp & Biven, 2012). Norepinephrine facilitates memory encoding via connections between the BLA and interacts with cortisol to suppress memory retrieval (Roozendaal, Hahn, Nathan, de Quervain, & McGaugh, 2004). The Locus Coereulus also affects procedural memory by interacting with reinforcement learning mechanisms (Sara, 2009), perhaps most notably the dopaminergic system that primarily involves VTA dopamine projections and is important in the SEEKING system. Goal or task directed behavior, i.e. *exploitation* (Aston-Jones & Cohen, 2005), is characterized by phasic activity in the LC while behavior more reminiscent of exploration is characterized by more tonic LC activity.

Thus, stress systems, the primary-process affect systems, and cognitive systems overlap and interact in several different ways (Figure 3-1 and Figure 3-2). Activation of the stress systems can occur with an increase in activity in both the FEAR (e.g., Johansen et al., 2010; Reyes et al.,

2011) and SEEKING (e.g., Roesch, Calu, Esber, & Schoenbaum, 2010; Sara, 2009) systems via unconditioned or conditioned stimuli; this causes a cascade of changes along the stress axes. Stress-based changes in hormones and catecholamines cause an increase in arousal (with the LC as an *arousal center*) that modulates declarative memory retrieval (suppressing retrieval of old memories) and encoding (facilitating the encoding of new memories). This arousal also affects procedural memory by interacting with reinforcement learning-based behavior and facilitating exploration or exploitation behavior (e.g., Aston-Jones & Cohen, 2005). Magnitude of the utility-based error of reinforcement (e.g., Glimcher, 2011; Sutton & Barto, 1998) causes an increase in phasic LC activity, that is, phasic LC activity decreases when prediction error remains low (for example, during overtraining) and increases when reward reinforcement differs from expectation[5].

[5] In a computational system that uses a single value to represent this tradeoff between tonic and phasic activity in the LC, as opposed to a network of nodes (e.g., McClure, Gilzenrat, & Cohen, 2006), one would likely need a type of piecewise function that favors an optimal value and decreases selection threshold when approaching the nominal value, but increases noise as the current value increases above nominal value (e.g., a function that would give a U-shaped noise transformation curve and an inverted U-shaped exploitation curve). Assumedly, the transformed value would represent a form of phasic/tonic ratio.

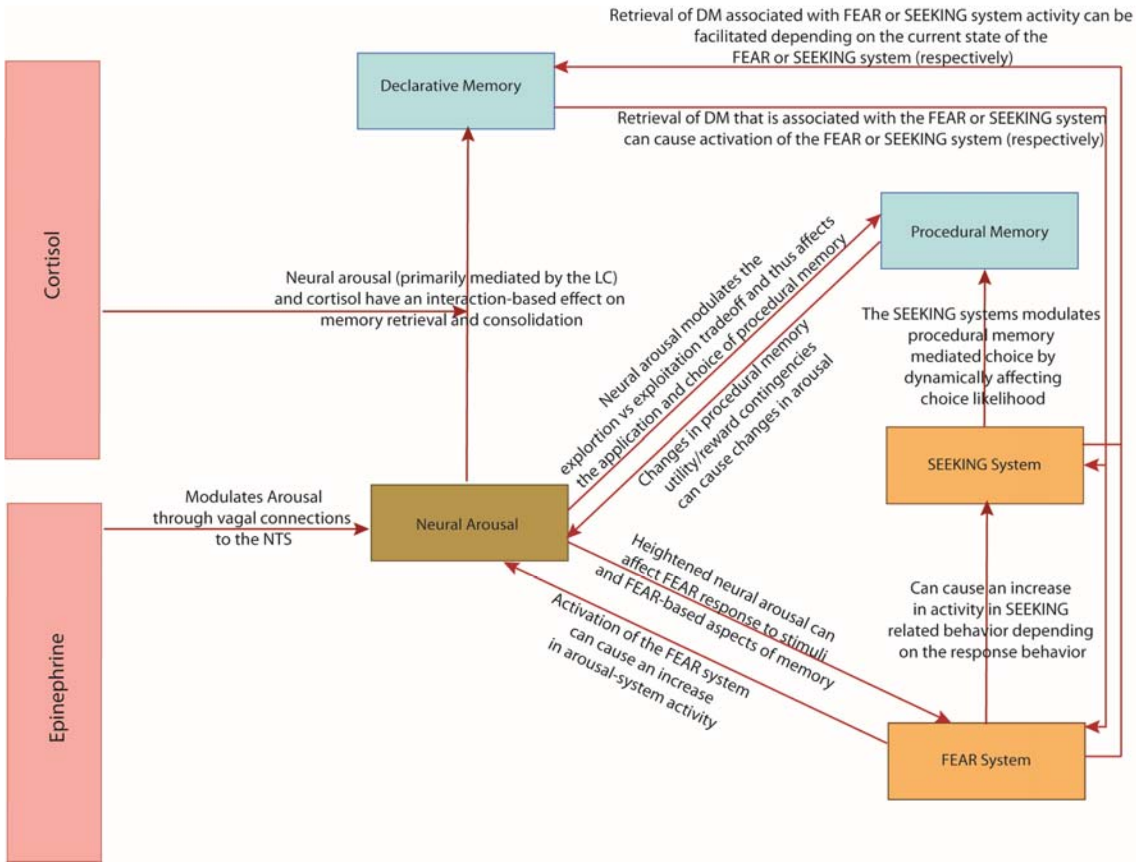

Figure 3-2. A high-level functional model of interactions between neural-cognitive/affective systems. This model provides a unified account of other models in cognition, stress, and affective neuroscience research (e.g., Anderson, 2007; Panksepp & Biven, 2012; Sara & Bouret, 2012; Schwabe et al., 2012)

It is clear that activation of affective systems due to contextual cues and stimuli can have implications for behavior, especially in terms of learning and decision-making. The breadth of the architectural model of cognition presented above makes the development of a corresponding computational system that implements the descriptive model useful, because it allows further specification and exploration of the model itself and what it means for decision-making behavior.

This descriptive model of an architecture has been developed into a computational system that allows the development of computational process models of human behavior that provide representations on the physiological, affective, and cognitive levels. The HumMod simulation system (Hester, Brown, et al., 2011) and the ACT-R cognitive architecture (Anderson, 2007) have been connected to make the ACT-R/Φ computational architecture. Computational process models have been developed to represent various effects of stress and thirst on cognition (Dancy, 2013; Dancy & Kaulakis, 2013; Dancy et al., in press). The ACT-R/Φ architecture can also be expanded to include more connections between the physiological, affective, and cognitive levels.

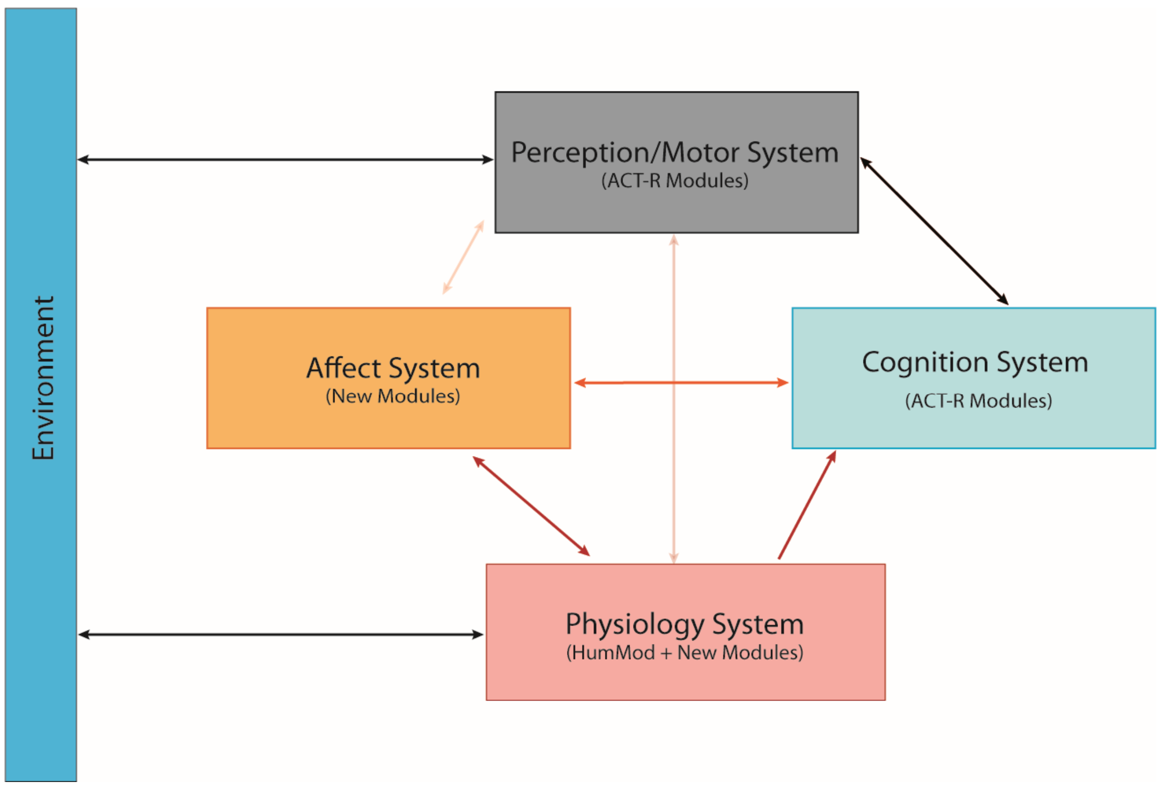

Figure 3-3. High-level description of the systems used in ACT-R/Φ

The ACT-R/Φ computational architecture

ACT-R/Φ extends the ACT-R architecture with new functional modules that connect the HumMod system to ACT-R and can facilitate physiological modulation of cognition. The physio module communicates with the HumMod system. The SEEKING, FEAR, and affective-association modules were added to add a basic theory of affect and emotion (Panksepp & Biven,

2012) to the architecture. These modules all have internal equations that are used to connect module behavior to other modules in the ACT-R architecture. These equations generally come from previous work and are informed by experimental data, but these equations have not been explored from a systematic perspective, that is, they often are not applied to a process within an overall system. Thus, the most useful equations for these processes have not necessarily been agreed upon and remain an interesting and open problem.

Physio module

In understanding how human behavior can change over time, both acutely and subtly, it is useful to understand how physiological variables may serve a modulatory function. Furthermore, the ability to manipulate physiological variables during the thought and behavior process can give insights on how small changes in physiology can eventually have large non-linear effects. ACT-R/Φ uses the *physio* module to serve as a functional communication between the physiological system (HumMod) and the affective/cognitive systems (traditional ACT-R and the new modules discussed later in this chapter).

The physio module synchronizes the two systems so that they are operating at the same point in simulated time. The physio module also holds the values of physiological variables over-time and can be used to change parameter values of the physiological system at any point in a simulation. The module updates the internal physiology variables periodically depending upon a *delay* parameter. Thus, every *delay* seconds, the module updates the *current* state of the physiological system and these values are used by various functions in equations to change the state of the affect and cognitive systems in ACT-R/Φ.

Physiology used by ACT-R/Φ

Values from the physiological system currently modulate subsymbolic properties of the ACT-R/Φ architecture. Specific aspects of physiology were chosen as a starting point for representing how physiology can modulate behavior. ACT-R/Φ contains connections between physiological variables related to stress and homeostatic-based behaviors. These connections are far from the final word in ways physiology can modulate behavior, but provide a useful starting point for simulation.

The two physiological systems typically associated with stress are the Hypothalamic-Pituitary-Adrenal (HPA) and sympathetic-adrenal-medullary (SAM) axes. The HPA axis runs from the hypothalamus to the pituitary gland and then to the adrenal cortex with corticotropin-releasing hormone (CRH), adrenocorticotropic-hormone (ACTH), and glucocorticoids (cortisol) acting as the major modulators along this pathway. The SAM axis describes activation of the autonomic nervous system that involves the peripheral release of epinephrine (adrenaline) from the adrenal medulla and also release of norepinephrine both peripherally via sympathetic nerves and centrally via the Nucleus of the Solitary Tract (NTS); the NTS has efferent projections to several areas of the brain including the Locus Coeruleus (LC) that is largely responsible for neural release of norepinephrine to areas including the hippocampus (Sara, 2009; Ulrich-Lai & Herman, 2009).

To represent the modulatory effects that components of these stress systems can have on cognition, epinephrine, CRH, ACTH, and cortisol have been identified as modulatory variables that are in the HumMod physiological system and can modulate subsymbolic properties of the cognitive and affect systems in ACT-R/Φ. These variables exist in the HumMod system, so the integration is fairly straightforward with an understanding of the HumMod and ACT-R systems and a way to communicate between the two, i.e. with the physio module in ACT-R/Φ. Some of the interesting effects that one can get from connecting these variables has been shown through a model of stress and mental serial subtraction (discussed later; Dancy et al., in press).

Homeostatic-based modulation of behavior has been another area of exploration for the ACT-R/Φ architecture. The pervasive and modulatory effects of physiological needs on affect and cognition (e.g., Aarøe & Petersen, 2013; Mogg, Bradley, Hyare, & Lee, 1998; Tuk, Trampe, & Warlop, 2011; N. D. Wright et al., 2012) make homeostatic-based behavioral change an attractive target for connections between the physiological, affective, and cognitive levels. Glucose and insulin have been identified as physiological variables that can modulate the effects of acute hunger on affect and cognition while leptin has been identified as having both acute and long-term effects on behavior (Berthoud & Münzberg, 2011; Harris, 2000). Osmolarity is the physiological representation that is used to modulate thrist-related behavior in the architecture as an increase in osmolarity can be treated as an objective increase in thirst and typically also results in a subjective thirsty feeling (Johnson, 2007; N. D. Wright et al., 2012). A preliminary model has been developed to run in ACT-R/Φ that has rule utility affected by thirst based on the state of osmolarity (Dancy & Kaulakis, 2013).

Below are equations that transform values from physiological variables into values that are used to change subsymbolic values of memory in the architecture (these values are used by some equations in the affective-associations module). Equation 3-1 gives an arousal value that can be used to modulate memory systems. Equation 3-2 is used to gauge the current state of physiological thirst sensors, so that when a model becomes thirsty (or the opposite) the architecture will begin to increase the chances of thirst related memory being used.

In Equation 3-1, the variables marked *base* (e.g., $cortisol_{base}$) represent baseline physiological values (as determined by HumMod), while variables marked *max* (e.g., $cortisol_{max}$) represent an expected ceiling value that can be set as a parameter, so a modeler could set them as a value that would be considered extremely high in normal conditions or a modeler could set it as the maximum cortisol value found during an experiment. The same variables that are not marked represent the current state of that variable (e.g., *cortisol*). Cortisol is treated as a multiplicative value in the arousal equation here due to evidence that its effects are most evident on arousing memory (e.g., Roozendaal, Okuda, Van der Zee, & McGaugh, 2006).

In Equation 3-2, $osmo_{base}$ represents the baseline osmolarity value (as determined by HumMod), while $osmo_{max}$ represents the ceiling osmolarity value that causes the HumMod physiological model to begin to *shut down*. As osmolarity increases models get thirstier and both thirst-related declarative memory (retrieval) and thirst-related procedural memory (rule firing) show increased subsymbolic values. The other homeostatic-based equations (hunger and skin temperature) are built the same way as the Equation 3-2.

Module equations

Equation 3-1

$$Arousal_{Phys} = \left(\frac{cortisol - cortisol_{base}}{cortisol_{max} - cortisol_{base}}\right) * \left(\frac{crh - crh_{base}}{3 * crh_{max} - crh_{base}} * \frac{epi - epi_{base}}{3 * epi_{max} - epi_{base}}\right)$$

Equation 3-2

$$\text{Thirst} = \frac{(osmo - osmo_{base})}{osmo_{max} - osmo_{base}} \; \forall \, osmo > osmo_{base}$$

SEEKING module

The SEEKING module is based on the SEEKING system, a low-level system in the primary-process affect framework (Alcaro & Panksepp, 2011; Panksepp & Biven, 2012). This system is key to appetitive motivation and behavior (e.g., feeling thirsty and approaching a cup of water as a result). The SEEKING module modulates the included memory systems subsymbolically by affecting the likelihood of memories being recalled (declarative memory activation) and the likelihood that certain actions will be carried out (procedural memory utility). This system can also be affected by the memory systems and this can cause a cycle of increased activation of systems if not interrupted.

In addition to the top-down memory changes in the SEEKING module, bottom-up physiological changes in ACT-R/Φ can also affect SEEKING module activation; an increase in physiological thirst will cause a change in the SEEKING system that can affect subsymbolic connections to procedural and declarative memory. Thus, the SEEKING module acts as an intermediate layer between physiological and cognitive processes. The connections between the SEEKING module and memory components is based on existing neuropsychological evidence of functional connections between the SEEKING circuit and memory systems. The most related and systematic evidence may be from the work by Berridge and colleagues on incentive salience (e.g., Berridge et al., 2009; Richard et al., in press; Zhang et al., 2009). Work related to the incentive salience framework provides evidence for a separation wanting motivation and behavior from hedonic liking and suggests useful mathematical equations that can be used in computational systems (K. S. Smith et al., 2011; Zhang et al., 2009).

Related neural substrates

Using known psychologically functional connections between neural and physiological systems is a useful way to ground architectural structure and function. The SEEKING circuit is composed of several neural structures that are related to psychological behavior.

Activity in the Lateral Hypothalamus (LH) modulates homeostatic-based behavior, most notably, hunger (Berthoud & Münzberg, 2011; Jennings, Rizzi, Stamatakis, Ung, & Stuber, 2013) and thirst (Bourque, 2008; Johnson, 2007). The peripheral hormones leptin and ghrelin, as well as peripheral levels of glucose (Friedman & Stricker, 1976), affect hypothalamic-based change in behavior. Thus, the inclusion of the lateral hypothalamus in the SEEKING circuit allows a straightforward addition of bottom-up homeostatic-affect to cognitive function in the ACT-R/Φ architecture by connecting the values of changing peripheral hormones, housed in the physio module, to functions in the SEEKING module that can modulate subsymbolic memory properties.

The ventral striatum, which is composed of the nucleus accumbens (NAcc) and olfactory tubercle (OT), may be the most widely known structure in the SEEKING system across the neuroscience, psychology, and cognitive science disciplines given its association with reward-related processing (e.g., Haber & Knutson, 2009; Knutson et al., 2001; Kober et al., 2010). The neurotransmitter dopamine is also heavily implicated in the SEEKING system as a neuromodulator and primarily comes from projections from the ventral tegmental area (VTA); opioids are regarded as more important for hedonic liking behavior (K. S. Smith et al., 2011). Given the correlation of basal ganglia structures and procedural memory system in ACT-R (both stimulus-response pairs and reinforcement learning; Anderson, 2007; Anderson, Fincham, et al., 2008) it seems straightforward that a module encapsulating SEEKING system behavior should affect procedural memory. The incentive salience framework strengthens the evidence for a functional subsymbolic connection and provides a complementary perspective, especially given its use to develop a reinforcement equation that includes basic affective modulation (McClure et al., 2003; Zhang et al., 2009).

Module equations

The SEEKING module serves as a main interface between changes in the physio module (especially homeostatic-based changes) and changes in the cognitive system. A physiological homeostatic-based change modulates the SEEKING affective state via internal values and these values are compared in a winner-take-all manner to determine the SEEKING *urge* that will affect architecture memory systems over a given period of time. The equations below are an example of the equations that are used to transform physiological change into homeostatic-based behavior.

U_A in Equation 3-3 represents the Utility of some production rule after an affective component ($\log(k)$ in this equation) has been added during procedural memory conflict resolution. Equation 3-3 is heavily inspired by the work by Zhang et al. (2009) on developing a mathematical model of incentive salience; k is a quantified representation of a SEEKING or wanting that affects rule choice and $sVal_{Thirst}$ is a representation of subjective thirst that is calculated by adding objective thirst (change in baseline osmolarity, or $osmo_{base}$) to a noise value (ε). The result of Equation 3-4 (k) is only used if it is highest k value when a periodically scheduled function polls all of the SEEKING values.

Equation 3-3

$$U_A = U + \log(k)$$

Equation 3-4

$$k = \ sVal * e^{reward_{max}} \ \ s.t. \ sVal \in [0,1]$$

Equation 3-5

$$sVal_{Thirst} = \frac{reward_{max} * \ (osmo - \ osmo_{base})}{osmo_{base} - \ osmo_{min}} + \ \varepsilon$$

Affective values attached to production rules are updated periodically when a conflict resolution, i.e. when the utility values of matching and partially matching production rules are compared; only the production rule selected during conflict resolution has its value updated.

In Equation 3-6, the delayed reward parameter k and the learning rate α can be set as a parameter in a model[6]. The affective (SEEKING) value of the next production from the most

[6] The equation $\frac{x}{1+kt}$ is also called a discount function (e.g., Fu & Anderson, 2006) that serves to discount the new value hyperbolically as the elapsed time since the previous value update increases.

recent production trace is represented by $S_{i+1}(n-1)$ and the variable t is the time elapsed since the same rule was last fired; thus, $S_i(n)$ is the SEEKING value saved when Equation 3-6 is used. The *reinforcer* (as computed in Equation 3-7) is determined using the current state of the SEEKING system.

Equation 3-6

$$S_i(n) = S_i(n-1) + \alpha(reinforcer + \left(\frac{S_{i+1}(n-1)}{1+kt}\right) - S_i(n-1))$$

Equation 3-7

$$reinforcer = (\log(SEEKING_{current} * e^{SEEKING_{max}}))$$

Many of the functions and equations that one may attribute to the SEEKING system that involve appetitive systems and memory systems (e.g., functions that would likely be discussed from a conditioned stimulus perspective) are contained within the affective associations module. The affective-association module represents an interface between the primary affect systems and architectural declarative and procedural memory systems. In the next section, I discuss the other module that represents a primary affect system, the FEAR module.

FEAR module

The FEAR module in ACT-R/Φ is based on the low-level affective FEAR system hypothesized by Panksepp and Biven (2012). The FEAR module is meant to encapsulate the low-level processing of aversive stimuli (e.g., a painful shock) done by the intrinsic FEAR system; in animals, the FEAR system can cause freeze and flight reactions when activated via direct electrical stimulation (Panksepp & Biven, 2012). The module is directly affected by nociceptive stimuli or aural stimuli and indirectly affected by visual stimuli. Visual stimuli do not directly change the state of the FEAR module as many of the principal neural substrates involved in low-level processing of aversive stimuli (including the amygdala), and their behavior, are represented with the *affective-associations* module and discussed in a later section.

The FEAR module principally operates independent of the learning systems, but can be affected by them depending on the affective content in the specific memory elements. As with the SEEKING module, the state of the FEAR module subsymbolically affects the declarative and procedural memory systems. The FEAR module also has a direct connection with the physio module and affects the SAM and HPA stress systems. Evidence for the deep connection between the FEAR system and stress systems (e.g., see Öhman & Mineka, 2001; Panksepp et al., 2011) make this connection interesting as the FEAR system not only changes physiological variables, but some of these variables (e.g., CRF and epinephrine) can affect both memory systems and the FEAR system itself, potentially modulating the FEAR state of the architecture on multiple levels at once[7].

Related neural substrates

The typical neural structure with the *fear emotion* is the amygdala (e.g., LeDoux, 1996). This is partially due to the multitude of studies involving a fear-conditioning paradigm of some kind to explore the cognitive and physiological effects of a conditioned stimuli (the amygdala is important for affective processing of emotional stimuli). Though the amygdala certainly does play a large role in the FEAR system described by Panksepp and Biven (2012), it is not necessarily the *principle* component of the unconditioned system. The most principle component of the FEAR system may be the Periaqueductal Gray (PAG). Several other neural structures are

[7] This deep interrelation among systems is a reason I chose to partially focus organization of architectural components based on physiological systems in addition to psychological systems. Depending on the perspective an article may take, the same portions of the amygdala could be part of a *stress*, *fear*, or *learning* system. It is useful to have a solid grounding in data and representations from neuroscience/physiology.

implicated to be involved in this system and it is important to try to understand the processing system as well as the principle components.

The intrinsic FEAR system principally runs through the central amygdala and the anterior and medial hypothalamus to the (more dorsal) PAG. The central amygdala (CeA) acts as an important structure in that it communicates with the PAG, the lateral hypothalamus (LH), and the paraventricular nucleus of the hypothalamus (PVN), to affect freezing behavior, autonomic nervous system response, and the release of certain hormones (respectively). Unconditioned responses (e.g., a response to a shock or a startling loud noise) do not need the amygdala to be present to occur, instead they more directly cause an activation of the PAG.

Module equations

The FEAR module only has a few equations as it is not a memory system *per se* but instead it is a system that is meant to represent a low-level aversive affective system that can cause changes to memory systems. Two main functions/equations are used to update values within the FEAR module. One function updates values attached to procedural memory (rules) that cause a rule to be preferred in certain states. Another function decays the internal FEAR value that is used by other modules.

Equation 3-8 is the major equation in the module, it is the production updating equation that is identical to the production-affect value update equation in the SEEKING module (Equation 3-6). In Equation 3-8 the delayed reward parameter k and the learning rate α can be set as a parameter in a model. The affective (FEAR) value of the next production from the most recent production trace is represented by $F_{i+1}(n-1)$ and the variable t is the time elapsed since the same rule was last fired. In Equation 3-9, $FEAR_{current}$ is the current state (value) of the

FEAR system and $FEAR_{max}$ is a parameter that can be set by the modeler that represents the maximum FEAR value a model can have.

Equation 3-8

$$F_i(n) = F_i(n-1) + \alpha(reinforcer + \left(\frac{F_{i+1}(n-1)}{1+kt}\right) - F_i(n-1))$$

Equation 3-9

$$reinforcer = (\log(FEAR_{current} * e^{FEAR_{max}}))$$

The other important (albeit simple) equation within the FEAR module is the value decay equation shown in Equation 3-10. Over time (varying by a set of timing-delay parameters that can be set by a modeler), the architecture's values for nociceptive (from the physio module), aural, associative (affective-associations module), and visual connections will decay at a constant rate[8].

Equation 3-10

$$val_{vision} - (decay_{vision} * val_{vision})$$

This equation is repeated for all of the internal values in the module that are listed in the previous paragraph. The equation is used under the assumption that the FEAR system within an architecture of the mind should return to normal with no outside influences. Many of the functions and equations that one may attribute to the FEAR system that involve aversion and memory systems (e.g., functions that would likely be discussed from a conditioned stimulus perspective) are contained within the affective associations module as this module represents an interface between the primary affect systems and architectural declarative and procedural memory systems.

---

[8] It could be argued that these equations should be more complex in some ways and possibly non-linear over time. For now, I stick to a simple decay function as the architecture as a whole is very complex and there aren't any straightforward computational or mathematical equations for this decay.

Affective associations module

The Affective Associations module acts as an interface between architectural memory and affective systems. The module is meant to be more representative of the processes that result in and react to learned (secondary) stimuli (e.g., a monetary reward; Delgado et al., 2011). In a way, the module is similar to the way the affect-system, as a whole, provides an interface between several physiological and cognitive systems. Unlike the SEEKING and FEAR modules, this module is not explicitly based on a primary affective circuit and could be considered to be between the primary and secondary levels described in Figure 2-1.

This module has several internal memory systems that link the then current affective state to visual, auditory, or declarative representations. Affective pairings were chosen to be represented separately based on visual, auditory, and declarative memory systems due to the evidence for multiple affective memory systems (Fanselow, 2010). These internal memory systems can change the affective state of a model and are tied to subsymbolic changes to procedural and declarative memory systems made by the module (see the equations section below for more detail on how these internal memory systems and the other memory systems interact).

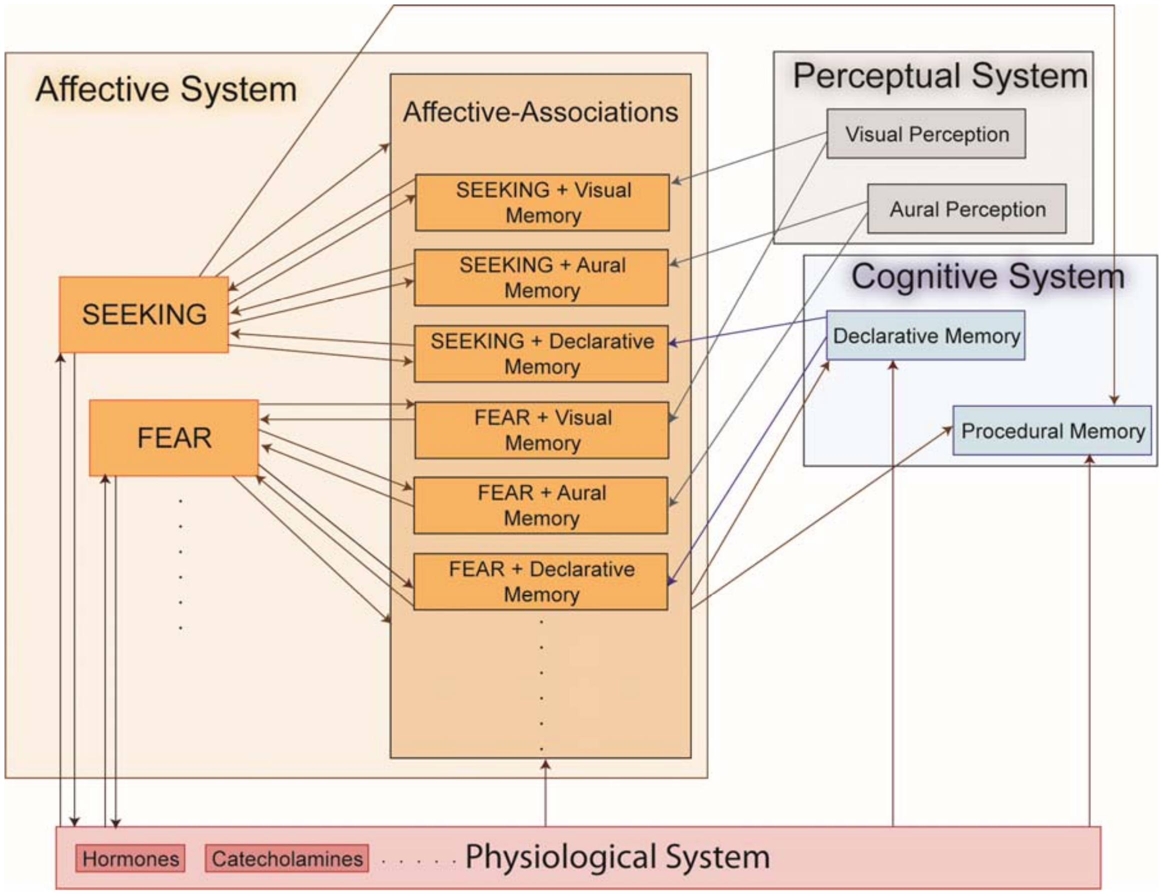

Figure 3-4. A high level view of ACT-R/Φ that shows the connections between the affective-associations module, the primary-process systems in the Affect system, the perceptual systems, and the cognitive systems.

Related neural substrates

The main structures implicated in the functionality of this module are the amygdala, ventromedial prefrontal cortex (VMPFC), and Orbitofrontal cortex (OFC). The amygdala has been shown to be important for affective learning for both appetitive and aversive events (Delgado et al., 2011; J. Li, Schiller, Schoenbaum, Phelps, & Daw, 2011; Phelps, 2006; Roesch et al., 2010). The VMPFC and OFC are also important for coupling affective experience with memory (Rolls, 2013), but are hypothesized to be particularly important for the processing of associativity between stored memories and affect (Bechara, 2005; Verdejo-García & Bechara, 2009).

As discussed earlier, the amygdala is composed of several sub nuclei that have different functions. The lateral amygdala (LA) and basolateral amygdala (BLA) are important for conditioned response to stimuli that predict an unconditioned stimulus of some kind (e.g., stimuli that cause increased activity in the FEAR system discussed earlier). These nuclei are not responsible for the explicit knowledge of conditioned and unconditioned stimuli mappings, but instead the affective reaction and anticipation related to a stimulus. The LA and BLA serve as an input to the amygdala structure and receive affect connections from neural structures related to declarative memory, visual perception, and auditory perception (among others).

The VMPFC and OFC (hereafter referred together as VMPFC/OFC) serve a function very similar to the amygdala and in fact one could consider these structures apart of their own system depending on perspective. Nonetheless, it is clear from lesion studies that the amygdala does generally serve a different purpose than the VMPFC/OFC (e.g., Bechara et al., 1999). The VMPFC/OFC seems to be important for assessing and normalizing affective value to a common scale (Levy & Glimcher, 2012) and for associating that affective value with declarative memory representations (Bechara, 2005). Thus, the VMPFC/OFC can be seen as partially mapping to the affective-declarative memory system in the affective associations module and the processes that lead to and result from this system.

Module equations

The memory systems in the affective-associations module use a TD-style equation (as shown in Equation 3-11), very similar to those in the SEEKING and FEAR modules, to update the affective value associated with visual, aural, and declarative-memory based stimuli[9]. Equation 3-11 and Equation 3-12 are used to update the FEAR affective value for objects in each memory system. $A_i(n-1)$ is the affective value (paired with a chunk from a perceptual system or the declarative memory system) before the update equation is called. The α variable is the learning rate that controls how much each instance of a chunk-value pair affects the overall chunk-value memory of that chunk. The $reinforcer$ variable is some primary reinforcer (e.g., a loud noise or a painful shock; Equation 3-12) and the FEAR variable represents the current state of the FEAR system (given by the FEAR module). As with the similar TD-inspired equations discussed in the previous two sections, $\frac{1}{1+kt}$ is the discount equation that functions to decrease the weight of the new chunk-value pair on the current memory as time between the update of the chunk-value pair increases pair increases.

Equation 3-11

$$A_i(n) = A_i(n-1) + \alpha\left(reinforcer + \left(\frac{FEAR}{1+kt}\right) - A_i(n-1)\right)$$

Equation 3-12

$$reinfocer = \max(FEAR_{aural}, FEAR_{nociceptive})$$

Equation 3-12 describes the reinforcing value and is determined by the maximum of unconditioned aural and nociceptive stimuli values in the FEAR module; that is, either an alarming loud noise could or a painful shock/tissue damage could reinforce the value, depending on which value is higher.

The affective-associations module also adds affective values to chunks that will make them more likely to be retrieved from declarative memory if the current affective state is more identical to the state when they were encoded, that is, being afraid makes a model more likely to

[9] These equations can also be overridden using a specific *hook* parameter (i.e., like the *hooks* that can be specified in the original ACT-R architecture) to specify the alternative function the modeler would like to use.

retrieve memories of other times it was afraid. During a given retrieval request chunk activation values are offset with Equation 3-13. In Equation 3-13 $FEAR_{max}$ is a parameter that represents the max FEAR value one would expect to occur within the system, while $FEAR_{chunk}$ is the FEAR value attached to the chunk and $FEAR_{current}$ is the value that represents the current state of the FEAR module.

Equation 3-13

$$Foffset = \log(abs\left(FEAR_{max} - abs(FEAR_{chunk} - FEAR_{current})\right) * e^{FEAR_{max}})$$

When chunks are combined in declarative memory, the state value that is attached to the new combined chunk is set to the average of the values from the two previous chunks.

The ACT-R/Φ hybrid architecture combines several theories of human behavior into one coherent computational system that has representations on the physiological, affective, and cognitive levels. This system provides behavioral predictions on all three levels from an architectural point of view. Physiology modulates the affective systems to cause a *bottom-up* change in cognition, while direct changes in the cognitive system (e.g., the retrieval of a stressful memory) can cause a cascade of physiological effects. The modules represented in the architecture are far from the final word, but provide a foundation that can be built upon to make more interesting predictions later, including those related to choice and decision-making.

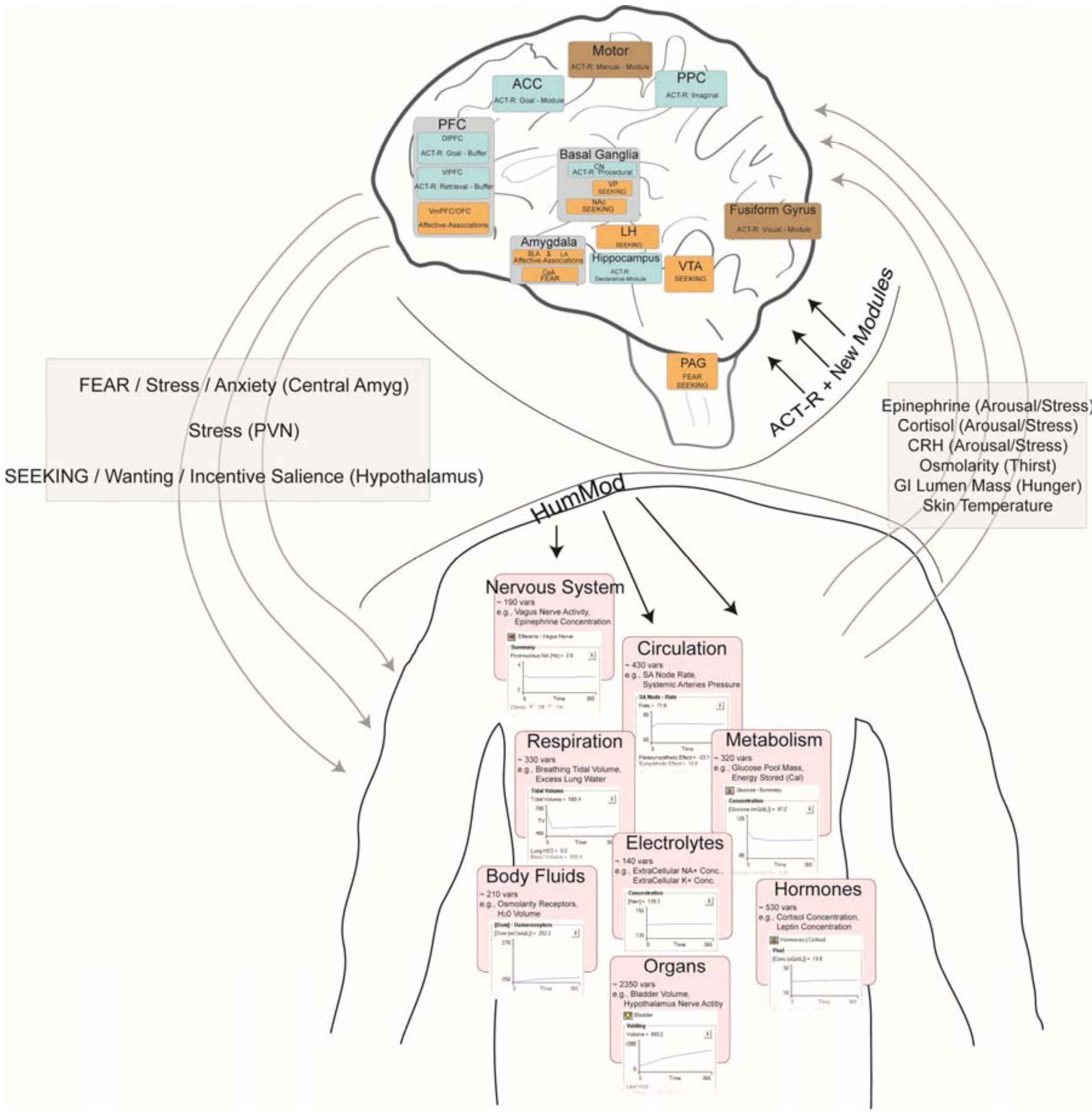

Figure 3-5. Modules in ACT-R/Φ mapped onto different brain areas (brain), the physiology represented by the HumMod system (body), and the connections between modules and physiology (directed lines). Abbreviations: ACC anterior cingulate cortex; BLA basolateral amygdala; CeA - central amygdala; CN - caudate nucleus; DlPFC dorsolateral prefrontal cortex; LA – lateral amygdala; NAcc nucleus accumbens; PAG periaqueductal gray; VlPFC ventrolateral prefrontal cortex; VmPFC ventromedial prefrontal cortex; VTA ventral tegmental area

Past models that use ACT-R/Φ components

Throughout the evolution of the architecture, models have been developed within ACT-R/Φ both to test out architectural mechanisms and to better understand the process of physiology modulating affective cognitive processes. Below, I present two models that run in the architecture. One is a model of serial subtraction that predicts effects of an increase in epinephrine over time. The other, is a model that simulates the effects of physiological thirst (osmolarity) on memory and choice in a *thirsty* version of the ultimatum game.

Stress model

Stress can be explored from several different perspectives and on several levels of behavior and has been widely studied (e.g., see Charmandari, Tsigos, & Chrousos, 2005; Joëls & Baram, 2009; Joëls et al., 2011; Kirschbaum, Pirke, & Hellhammer, 1993; Lazarus & Folkman, 1984). Despite the amount of literature on stress, relatively few computational process models of how stress affects cognition exist. Ritter, Kase, Klein, Bennett, and Schoelles (2009) explored what parameter values in the ACT-R architecture resulted in simulated behavior that was similar to like human behavior in a stressful task. They developed a model that performed mental serial subtraction while under stress. They chose three parameters to vary: syllables per second (SYL), base-level constant (BLC), declarative memory noise (ANS); ANS and BLC directly affect declarative memories. The work by Ritter et al., 2009 was useful and moved understanding forward, however, the model did not make clear how physiological and cognitive variables may interact to change cognition, their model was represented purely on a cognitive level.

An initial version of the ACT-R/Φ architecture was used to run a modified version of the Ritter et al. (2009) model. This new model expanded upon the previous model by connecting continuously varying physiological variables to cognitive parameters (Dancy, Ritter, & Berry, 2012; Dancy et al., in press). This meant that now, when the model would get stressed, its behavior would be continuously changed over the course of the task. Thus, a stressful experience in block one would cause a non-linear change in the physiology that affected cognitive behavior. The model only had to be slightly modified (a new rule, and a few more architectural parameters) and this presented an advantage of extending a modular architecture that has been already built to include a physiological substrate: models developed to be run within the architecture are likely to work within the new extended system with relatively little (or even no) change.

The new model added a connection between the declarative memory noise parameter and epinephrine level in the physiological substrate (represented with HumMod, discussed earlier, as the physiological model). This connection means a rise or fall in epinephrine levels can make it easier or harder for the model to retrieve declarative facts. A stressful stimulus (a loud noise in this model) would cause an activation of the sympathetic nervous system, causing an increase in peripheral epinephrine (Figure 3-6). In turn, epinephrine continuously affects declarative memory noise, making it easier or harder to retrieve the correct declarative memory chunks (depending on the epinephrine level). As the model is completing the mental serial subtraction task, a stressor is aurally perceived and this causes a brief change of attention, as well as an activation of the CNS integration nerve (in HumMod) causing an increase in peripheral epinephrine. The stressful experience only stops affecting peripheral physiology when the subtraction problem must be re-setup, which occurs when the model gets a problem incorrect. The rule that is fired to setup the subtraction problem also deactivates the CNS integration nerve (to the nerve's baseline state) which causes the epinephrine to begin to decrease towards baseline (e.g., see Figure 3-7). Therefore, declarative memory noise is continuously changing over the course of the task, dependent on the models current physiological state, as determined by HumMod.

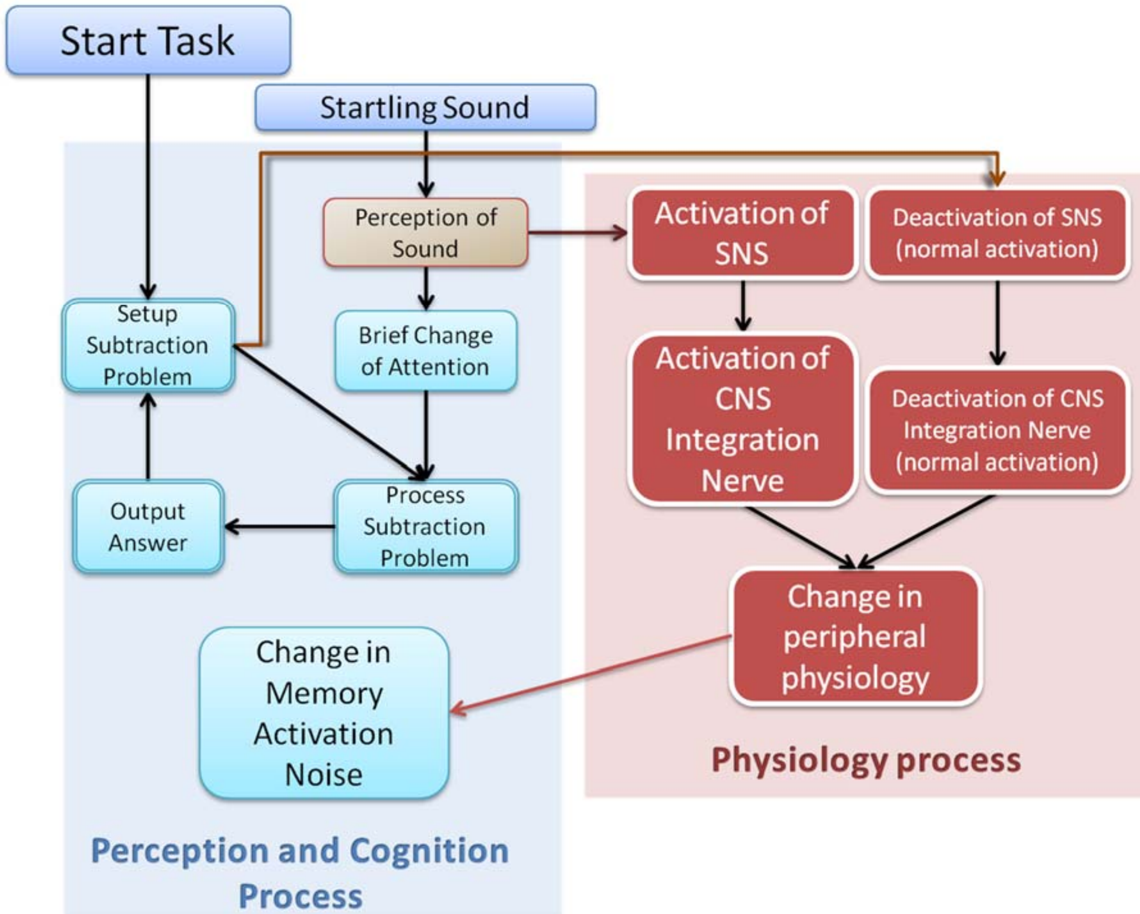

Figure 3-6. A diagram of the ACT-R/Φ serial subtraction model

As shown in Figure 3-7, the physiology of the model never actually reaches its original state throughout the task because the temporal dynamics of the task cause the model to be stressed again before epinephrine levels can reach baseline. Thus effects of stress from the first block of the task not only carry over into the second task, but also into the third and fourth blocks.

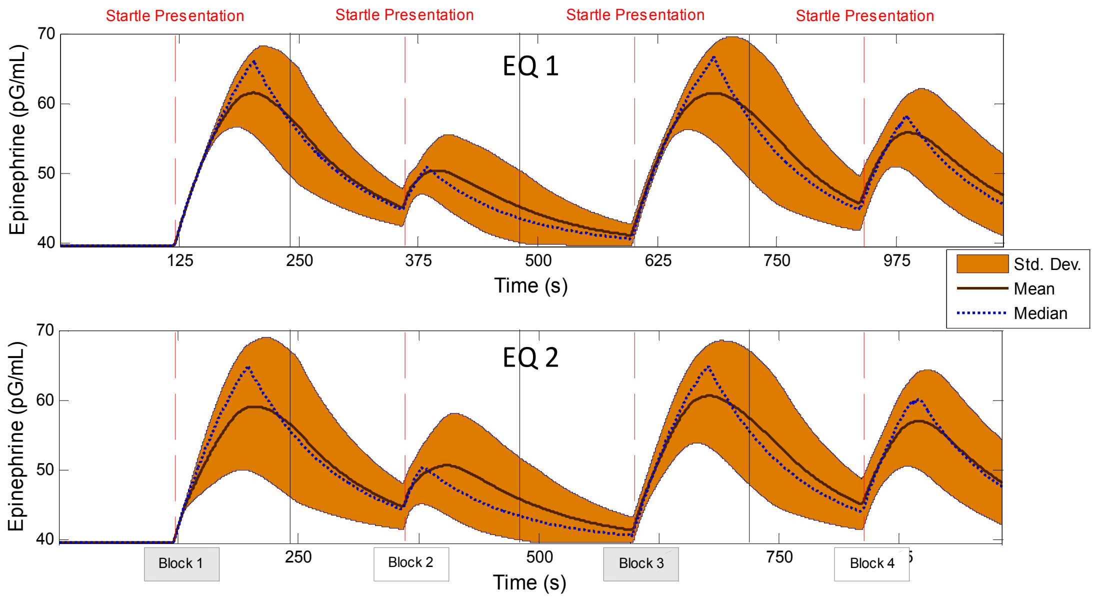

Figure 3-7. Epinephrine levels over the course of the task. The black solid line gives the average epinephrine level at that point in time and the dotted line gives the median at that point in time. The filled in area around the solid and dotted lines gives the standard deviation. ($n = 200\ runs$)

Two equations were explored and used to connect epinephrine to declarative memory noise. Equation 3-14 provides a simple linear transformation, while Equation 3-15 is a piece-wise linear equation meant to cause an inverted U-shaped performance curve (e.g., Yerkes & Dodson, 1908). The default value used for the $ansMultiplier$ variable was determined to be 21.4.

Equation 3-14

$$:ans \ = \frac{ansMultiplier * (Epi_{CurrentValue} - Epi_{Baseline})}{Epi_{Max}}$$

Equation 3-15

$$:ans = \begin{cases} \dfrac{-2 * ansMultiplier * (Epi_{CurrentValue} \ - \ Epi_{Baseline})}{Epi_{Max}} + :ans_{ACT-R-All} \ \forall \ Epi_{CurrentValue} \ < 50.25 \\ \dfrac{2 * ansMultiplier * (Epi_{CurrentValue} \ - \ Epi_{Baseline})}{Epi_{Max}} - :ans_{ACT-R-All} \ \forall \ Epi_{CurrentValue} \geq 50.25 \end{cases}$$

The second equation causes an agent to have a higher memory noise value at the onset of the task (with lower levels of epinephrine in the blood) and as the levels increase up towards the *50.25(pG/ml)* value, declarative memory noise decreases, making it easier for the model to retrieve declarative subtraction facts and thus improving the models performance. However, as the epinephrine levels eclipse the set threshold value, memory noise begins to increase again, making retrieval more difficult. Figure 3-8 shows the declarative memory noise values found using the different equations over time.

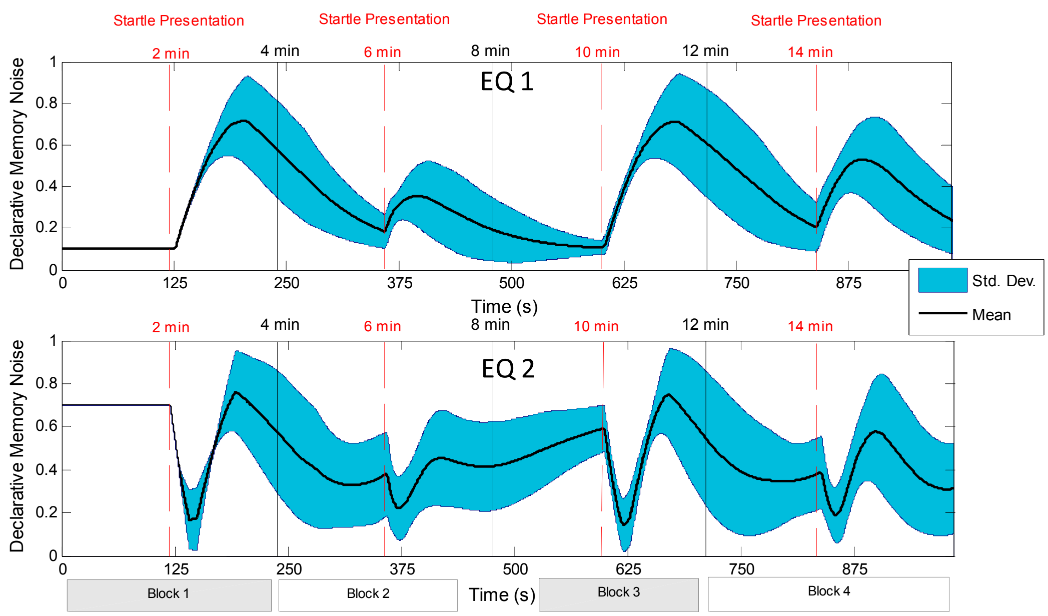

Figure 3-8. Declarative memory noise over the course of the task. The black solid line gives the average declarative memory noise at that point in time and the dotted line gives the median at that point in time. The filled area around the solid and dotted lines gives the current standard deviation. ($n = 200\ runs$)

Despite equal rules, and stimulus onset times, average epinephrine levels and variation differed across time (Figure 3-7). This is because the *deactivation* of the sympathetic nervous system (returning activation to baseline levels) happened after a certain rule fired. There existed (and still exists) a cyclic relationship between the physiological and cognitive levels in ACT-R/Φ.

The percent of subtraction problems correct achieved by the model while using Equation 3-15 (EQ 2) was lower than both the human data and the models performance while using Equation 3-14 (EQ 1). In fact, between the three models (a) the original ACT-R model used by Ritter et al. (2009), (b) the ACT-R/Φ model using Equation 3-14, and (c) the ACT-R/Φ model using Equation 3-15, the Equation 3-14 model had a mean closest to human data (79.2% correct) while both Equation 3-14 and 3-15 models had considerably more number of subtraction problem attempts than the ACT-R models; the human data had a higher number of attempts than all of the models (Dancy et al., 2012).

While there are several ways in which this ACT-R/Φ model could be improved, it represented an important first step in connecting the HumMod system to ACT-R. No other model had yet used a continuously changing integrative model of physiology to modulate cognitive parameters (though Gunzelmann et al., 2009, did use a mathematical model of *alertness* to alter cognitive parameters). This model and work showed that in connecting a physiological system to a cognitive architectures, one can begin think about human behavior from another perspective and simulate interesting ways physiology can affect cognition, as well as interesting ways cognition may affect physiology.

Thirsty model

N. D. Wright et al. (2012) conducted an experiment that examined how human responses to unfairness change when the unfairness is related to their own homeostatic change. They artificially induced some participants to be thirsty using intravenous hypertonic saline (saline with a high NaCl concentration) and had the participants complete a modified version of the Ultimatum Game (Güth, Schmittberger, & Schwarze, 1982). In the Ultimatum Game, participants are normally assigned to be either a *proposer* or *responder*. In a typical game, the proposer can offer to split a monetary reward a certain way and the responder can either accept or reject that offer with a rejection resulting in neither player getting any of the reward. N. D. Wright et al. (2012) modified the game in two important ways: (1) Participants were told that the game would be split up into proposers and responders when in actuality participants were only assigned to the responder role and the proposal was experimentally fixed; (2) Instead of the typical monetary reward, glasses of water were used to see how they would affect a response in the game from a participant who has been artificially made to be thirsty.

This experiment was a useful experiment to help form the architectural model and flush out architectural model components principally due to the wide-range of data presented. In addition to the behavioral measures (the decision to accept or reject the predetermined proposal), N. D. Wright et al. (2012) also collected an objective, physiological, measure of thirst (osmolarity[10]) and a subjective, questionnaire-based, measure of thirst. Thus, it was relatively straightforward to develop a model of this task within ACT-R/Φ given that the physiological system provides osmolarity measures and the SEEKING module (and the underlying affective circuit it is based on) is important for appetitive *urges* that may affect one's decision to accept or reject a potentially unfair offer.

A colleague and I developed a model that began to look at some of the processes underlying the behavior reported in this experiment by developing a process model to operate within the ACT-R/Φ architecture (Dancy & Kaulakis, 2013). With this model, we were able to simulate not only some of the processes underlying the cognitive and even affective aspect of the task, we were also able to directly model the artificial thirst induction procedure. Figure 3-9 shows a high-level diagram of the process model. As in the experiment, the model is first artificially induced to be thirsty by turning the IV switch on (in HumMod) using the *set-phys-vals* function from the physio module. The physiological model is then advanced an hour (independent of actual ACT-R time) using the *advance-phys* function from the physio module and after the model has advanced, the IV drip is switched off (again through the physio module). After this physiological change is complete, the simulation starts. During the task, the model perceives the

---

[10] See Johnson, 2007 for a useful review on mechanisms of thirst.

offer and keep, determines whether it will consider the offer as fair or not, and makes a decision to accept or reject the offer.

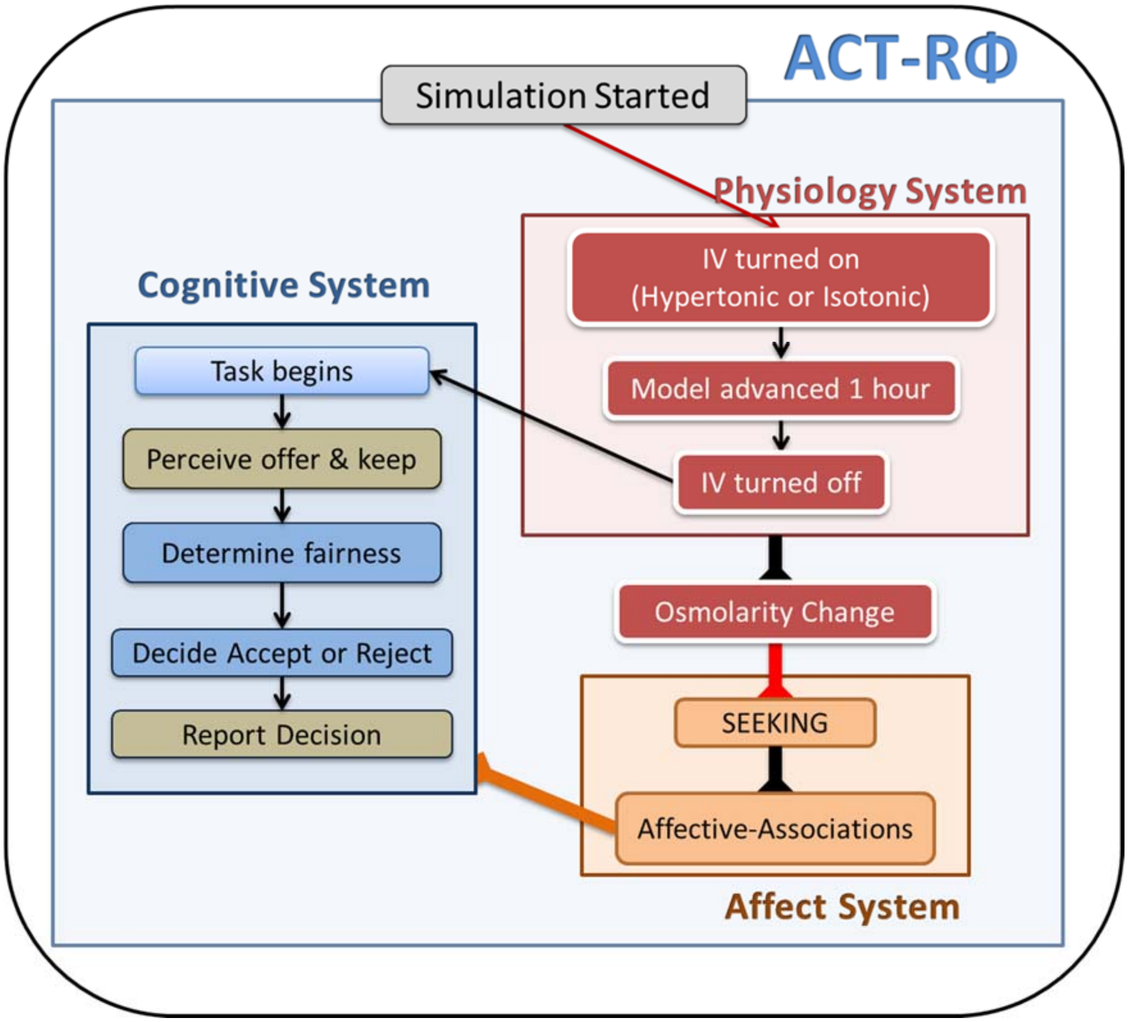

Figure 3-9: A diagram of the ACT-R/Φ thirsty model.

The model's subjective[11] thirst dynamically affects the utility values of productions rules that are marked to be related to thirsty state. The rule that makes a decision to accept the offer is associated with getting rid of the thirsty state, because accepting the offer means the model would

[11] Subjective thirst, in this case, is the objective thirst (determined by the current physiological state) with noise added that is determined using the built-in ACT-R noise function.

get to “drink” water and begin to quench its thirst. Figure 3-10 shows a histogram of the model’s decisions, split by whether the model decided to accept or reject the offer made during the game.

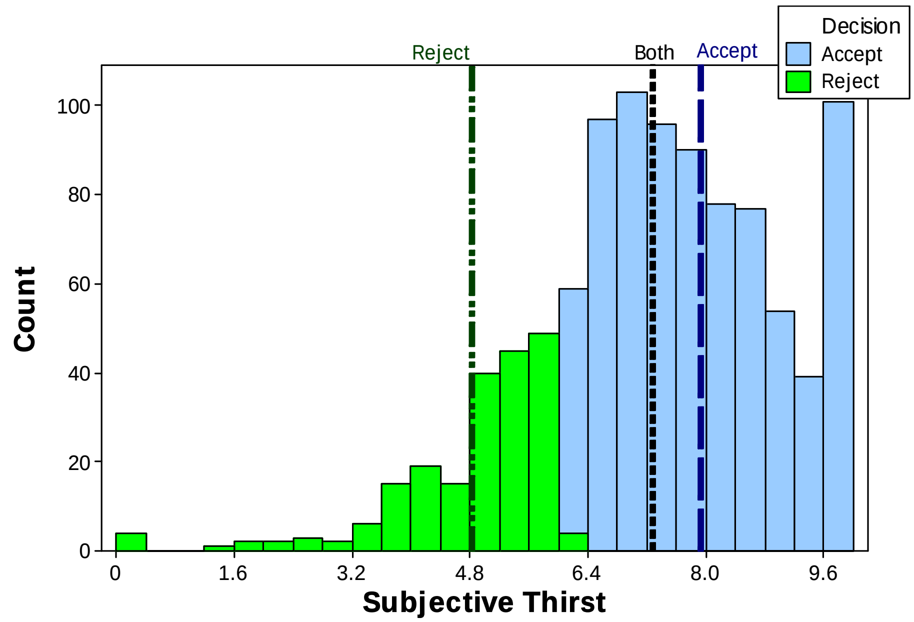

Figure 3-10. A histogram of the subjective thirst results produced by the model. The lines represent the mean for subjective thirst for models that accepted the offer, rejected the offer, and the two combined (the 2nd vertical line).

As shown in Table 3-1, this model was able to output physiological and affective (subjective thirst) results that were very similar to the original model, but the behavioral decision outcomes were not as closely matched as the other data the model had an acceptance rate of 79% while participants displayed an acceptance rate of %50. There are a few potential reasons for the mismatch in choice behavior (see Dancy & Kaulakis, 2013, for a brief discussion) one of which may be the lower number of participants in the original experiment. The low number of participants was partially due to the nature of the experimental treatment and could be augmented with an expanded version of the model presented above. With a model developed within the ACT-R/Φ architecture, one could also potentially manipulate and explore variations of the experiment to understand how the processes interaction.

Table 3-1. Thirst results from the model

Decision	Osmolarity (sd)	Subj. Thirst (sd)
$Accept_{experiment}$	Not Reported	8.90(1.7)
$Accept_{model}$	306.37(0.2)	7.94(1.2)
$Reject_{experiment}$	Not Reported	5.60(1.6)
$Reject_{model}$	305.86(3.74)	4.82(1.2)
$Both_{experiment}$	310(5.0)	7.30(1.6)
$Both_{model}$	306.27(1.7)	7.29(1.7)

Figure 3-11 shows an example change in osmolarity over time when the HumMod physiological model is made to be thirsty by hypertonic saline IV drip.

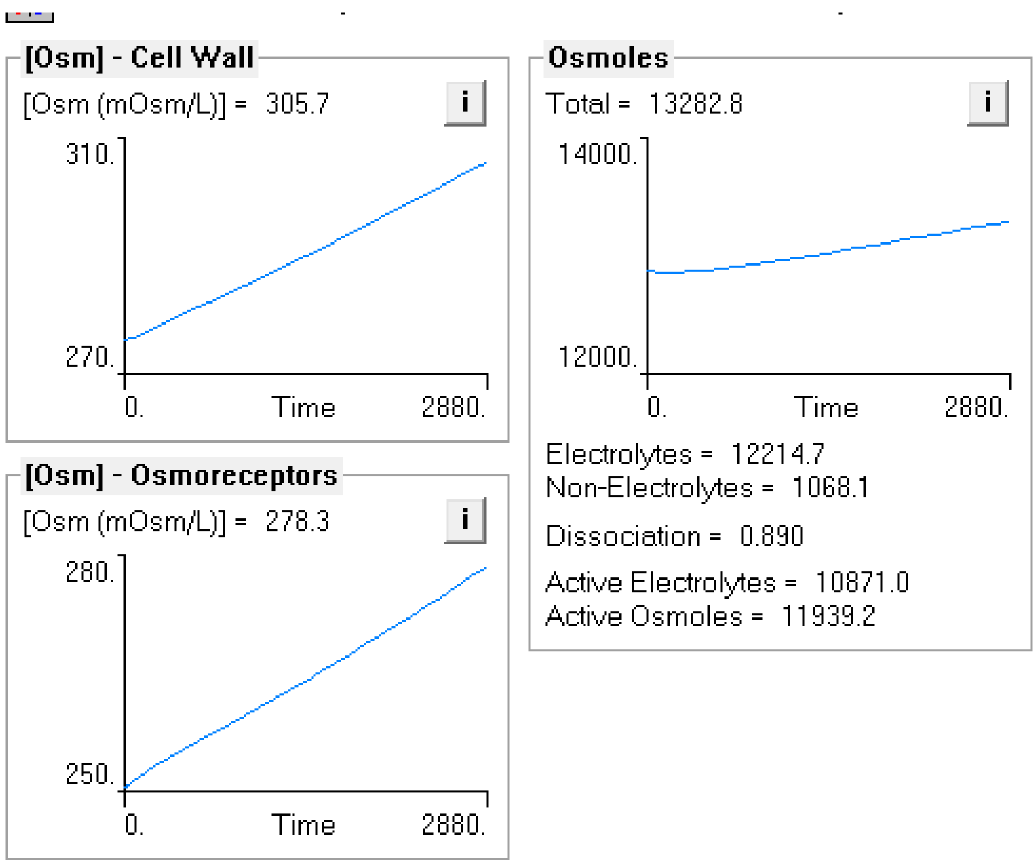

Figure 3-11: Example HumMod output that shows the change in osmolarity over time.

Because the length and volume of the IV drip can be adjusted in the model and simulation, the physiological system could be used in the future understand what subtle and overt changes to human behavior can occur leading up to the objective thirst levels observed in the original experiment, that is, one can understand how choice changes with different levels of objective thirst. If a IV is not feasible in an experiment, this model can also be used to get an estimate of how long a person may have to go without consuming any water to approach the osmolarity values originally induced in the experiment by N. D. Wright et al. (2012).

Summary

In this chapter, I've outlined the architectural model that describes and the computational system that implements the ACT-R/Φ hybrid architecture. This architecture is composed of four main systems (Figure 3-3 and Figure 3-4) that all interact and involve different (but sometimes overlapping) neural systems (Figure 3-5): (1) a physiological system that represents some of the physiology that is implemented by the two other systems and can modulate those systems; (2) an affective system that represents the current affective and emotional state at any given point of time, that is highly modulated by physiological processes, and that subsymbolically affects memory processes; (3) a cognitive system that provides representations for memory and perceptual processes and that is modulated by, but can also change, processes in the affective and physiological systems (4) a perceptual/motor system that takes input from the external environment and can also output information to the external environment.

In the computational implementation of ACT-R/Φ, the physiological system holds a current physiological state that is supplied by the HumMod integrative physiological model. Homeostatic and stress variables affect processes in the affective and cognitive systems and can also be affected by those same processes.

The affective system contains modules for the SEEKING and FEAR systems from the primary-process affect framework and an affective-associations module that associates chunks with affective values. The SEEKING module describes a form of appetitive motivation and or *wanting* urge. This module is heavily driven by homeostatic-based change in the systems physiological state and directly affects procedural rule utility. The FEAR module represents aversive affective states and is more involved in the unconditioned response to aversive stimuli than the learning of associated stimuli (though it does provide a value for learning to occur). The affective-associations module adds an independent memory and learning component to the affect system while also directly affecting the retrieval of declarative memories.

Given that both the SEEKING and FEAR modules should, in principal, have a large collection of built-in reactions to unconditioned stimuli, and the affective associations module should be biased towards certain stimuli (e.g., Öhman & Mineka, 2001) it seems future development of the architecture should include characterizing these responses. Memories that can be controlled by process-models will likely have to be *built-in* to the architecture to represent certain evolved responses and dispositions.

Future development of the architecture will also hinge on better understanding of ways to quantify the processes that are described in this chapter. Many of the models that describe affective and emotional behavior are descriptive and thus can be difficult to quantify and integrate into a systematic account of human behavior

The ACT-R cognitive architecture provides the cognitive system of the hybrid architecture. The cognitive system has several modules (*Figure 2-10*) that represent goal, memory, and perceptual processes. The memory-based modules (procedural memory, *rules*, and declarative memory, *chunks*) can be modulate by both the affective and physiological systems and both the memory and perceptual processes interact with processes from the affective and physiological systems.

Cognition interacts with affect and emotion, while both are supported and implemented by physiology. The ACT-R/Φ architecture does not describe all aspects of human behavior but combines experimental evidence and existing theories from psychology, cognitive science, and neuroscience to present a novel unified theory of human behavior that describes processes on multiple levels. The computational system can be used to predict how the physiological, affect, and cognitive levels interact to result in dynamic human behavior over time. Understanding how these levels of behavior interact is especially important for comprehending decision-making. Physiological and affective states interact with goal, memory, and perceptual processes to determine decision-making and choice behavior. Being cognizant of the ways these processes interact is important for understanding why we make the decisions we make and, more importantly, for understanding how we can make better decisions or construct our environment so that we can make better decisions.

Chapter 4 A model of decision-making and affect

Fully exploring the new extensions provided by the ACT-R/Φ architecture requires the development of process models that operate within the architecture. Though several viable candidate experiments exist that can be developed into a computational process model, I chose to model a version of the Iowa Gambling Task (IGT; Bechara, Damasio, & Damasio, 2000) that uses some of the new extensions.

In developing a useful new process model within the architecture, I followed a few informal criteria for deciding which task to model:

- The task should be considered a decision-making task
- The task should have enough complex behavior to make a computational process model especially useful and interesting
- The task should be related to affective and physiological behavior in some way
- The task should explicitly have some form of affective stimuli

Based on these criteria, I chose to model behavior during the IGT. The IGT also has an advantage of previously being used to explore decision-making changes in different clinical populations; this is an advantage from the computational process model perspective, because the model can be later modified to simulate behavior from distinct clinical populations. There was not an existing Lisp-based ACT-R process-model of the IGT, so a new computational process model had to be developed to complete the task. Though the IGT has been used to explore the role of emotional processes in decision-making, it does not contain explicit affective stimuli per se. Thus, a version of the IGT that includes subliminal affective stimuli was modeled.

I extended the Iowa Gambling task to subliminally (visible for 17ms) present picture stimuli from the International Affective Picture System (IAPS; Lang et al., 1997); these pictures are presented after the participant selects a card from a disadvantageous deck (decks a and b). These images are more explicitly related to emotion than just money rewards and punishments and IAPS is a standard picture system that is often used in emotion literature (Bradley, Miccoli, Escrig, & Lang, 2008; Keightley et al., 2003; Lang, 1995). (More information on the actual experiment and the resulting human data is available in Chapter 4).

The ACT-R/Φ IGT model (hereafter referred to as the *IGT model*) has 54 productions and 35 chunks that are initially in declarative memory; the (non-default) parameter values used in the model are in Appendix A. Reward and loss from the decks is processed by treating the reward and loss as separate entities (for a discussion on modeling the IGT in ACT-R, see Napoli, 2010). This model is similar to the one used by Napoli and Fum (2010) in that it uses the reward-loss transformation function examined by Ahn, Busemeyer, Wagenmakers, and Stout (2008), but differs in that it is a model that uses the relevant perceptual, motor, and memory components in ACT-R and uses the predominant version of ACT-R (their model uses a version of ACT-R written in another programming language). Thus, this is the first known model that uses both the perceptual and memory (both declarative and procedural) components of ACT-R. In essence, this now allows a more detailed account of the process of card-deck selections over the course of the Iowa Gambling Task. Having a model that uses modules from the main version of ACT-R also means that the model-module interaction can be more clearly designed based on work in neuropsychology and neuroscience because behavior from ACT-R modules have been related to different brain regions (Anderson, Fincham, et al., 2008). Figure 4-1 presents a mapping between brain areas, ACT-R/Φ modules, and process-model behavior.

The model communicates with the same Matlab program used by the participants in the experiments using the ACT-R JSON Network Interface (JNI) module (Hope, Schoelles, & Gray, 2013). The JNI module connects the model to the Matlab program using TCP sockets that keep the model and the game synchronized. Using the JNI module with the visual and motor modules gives one a model with more realistic behavior as a model must not only make a decision, but process visual locations on the monitor and press certain keys on the (virtual) keyboard.

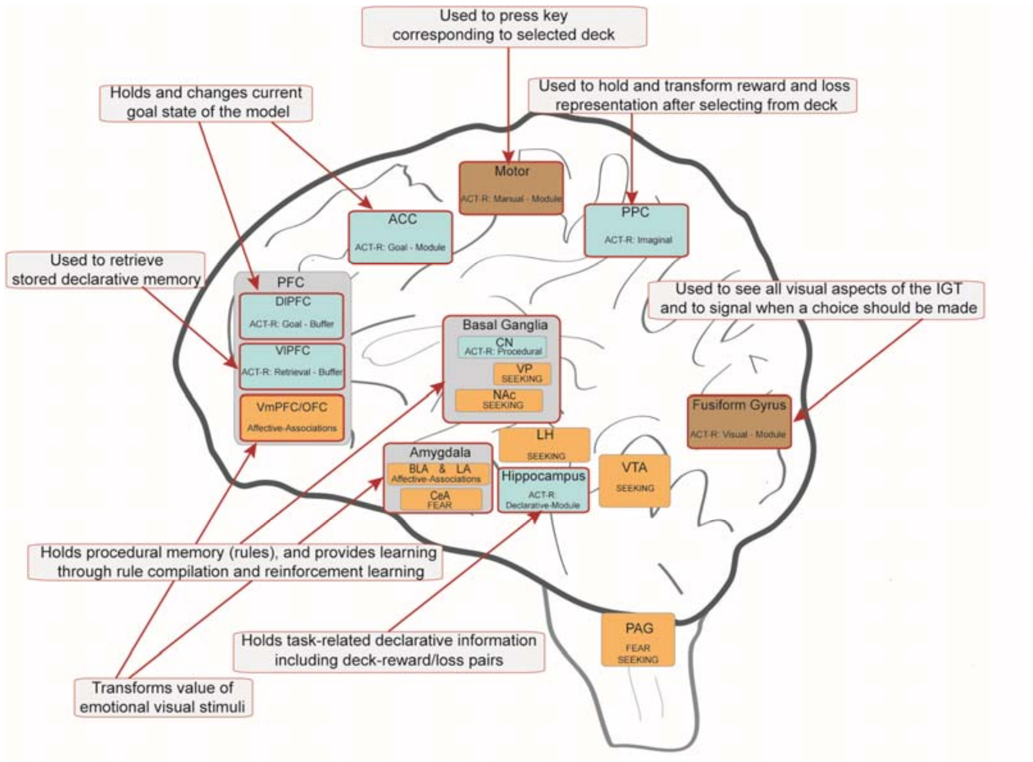

Figure 4-1. A depiction of the functional modules (and presumably brain areas) that are predicted to be used during the IGT along with an explanation for some of these task-related functions

Model description

The IGT model uses both the procedural learning and declarative learning components of the architecture. There are several chunk-types in the model (see Appendix A for a list of these chunk-types), but the most notable chunk-types are the *deck-value-pair* and *deck-key-pair* types. The deck-value-pair type stores an instance of the reward and loss values received from a deck, as well as, an instance of the normalized transformed values (as a separate chunk from the non-transformed reward/loss chunk). The transformed value that is stored in declarative memory is also used to reinforce rule choices (see the section on procedural learning in the model to get more information on this process). Figure 4-2 shows a high-level view of the IGT process model behavior while Figure 4-3 shows a more detailed diagram of the process model steps during each trial.

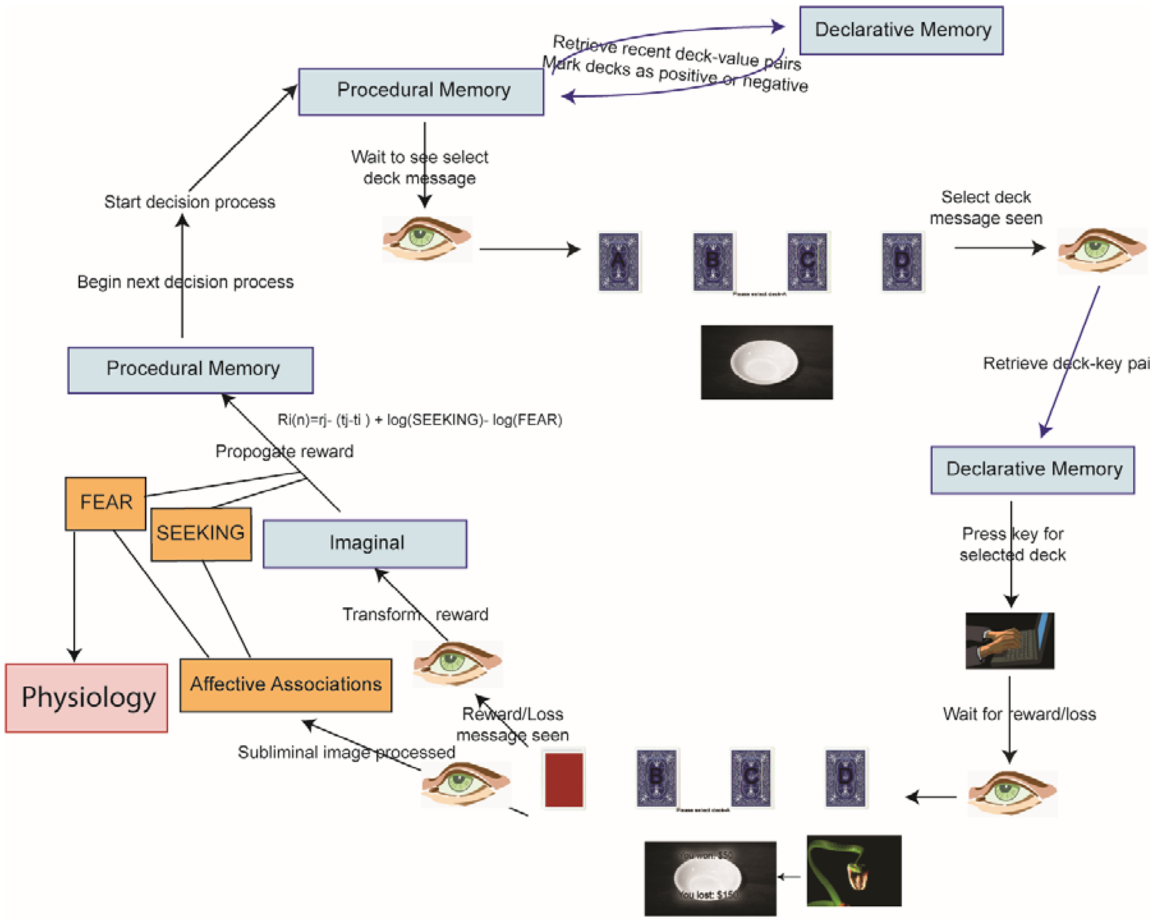

Figure 4-2. A high-level description of the IGT process model

After the simulation begins, the model retrieves a chunk for each deck that provides a value representation (a chunk of the deck-value type). After retrieval for chunks of all decks have been attempted, the model begins the *deck selection* stage of processing. Each of the decks have a rule that can be used to determine whether the deck should be considered a *good* deck or *bad* deck. Each deck rule determines this by testing whether the value in the reward slot from the chunk in declarative memory is greater than the value in the loss slot. The deck selection can be made either after at least one deck is found to be a *good* deck, or after all decks are found to be *bad* decks. The model then waits until it can make a deck selection and when the interface notifies it that the deck selection phase has begun, the model looks at the deck it intends to select and, if the deck is not exhausted, it presses the key associated with that deck. The model then waits for the reward and loss to be presented.

If the model selects from either deck A or deck B, a subliminal image is flashed (~17ms) and is placed into the visual-location buffer of the model. Because the image is only in the visual-location buffer for less than 50ms (the time it takes for a rule to fire), the image is never processed by the cognitive system and no explicit symbolic representation of the image is stored. However, the affect system does process this image (via the affective-associations module) and the image affects the current state in the SEEKING and FEAR modules, thereby having a small effect on subsymbolic properties of the memories being stored after the image presentation; no explicit representation of the image is stored in any memory systems in the affective-associations module as this is only accomplished when chunks reach the non-location buffers in the perceptual modules (i.e., the *visual* and *aural* buffers as opposed to the *visual-location* and *aural-location* buffers). The image value is determined by an equation related to its ratings in the valence, arousal, and dominance dimension (see the section on affect in the model for more information about this mechanism).

When the reward and loss are presented, it first looks at and creates an internal representation for the reward, following this, the model looks at and creates an internal representation for the loss. After creating representations of both reward and loss, the model internally transforms the reward and loss to create a single value representation of the result (see the section on procedural memory in the model for a description of the equation used to transform and collapse the reward and loss into a single value). After the reward and loss are transformed, the resulting value is propagated back to all of the rules fired during the trial. After the reward propagation is complete, the model restarts the deck selections process (i.e., the "Attempt DM retrieval of deck-value pair for all decks" stage in Figure 4-3).

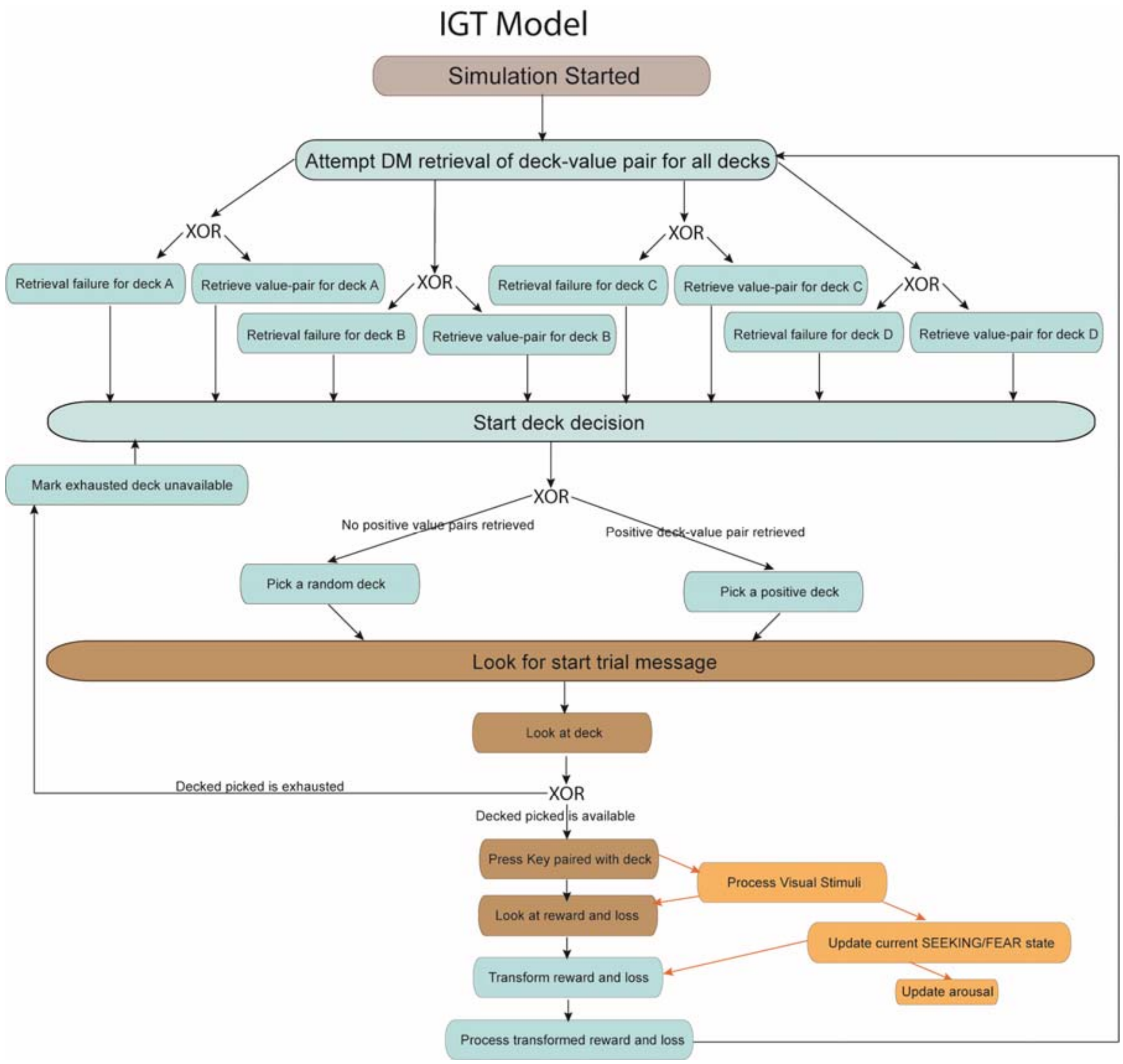

Figure 4-3. A diagram of the IGT process model

Affect in the model

The IGT model uses the affective-associations module to transform affective information in the visual buffers into a normalized value that can affect physiology (causing an increased arousal), and memory (likelihood of retrieving a memory or selecting a rule). In the experiment, subliminal visual stimuli are presented to participants after selecting from a deck A or deck B.

The model accounts for this by using the affective-associations module visual-hook to use a function that transforms the image in a visual buffer into a value.

Because each image used in the experiment was from the international affective picture system (IAPS), every image had three values attached to it that specified characteristics of the image: (1) arousal, a measurement of the amount an image excites an individual; (2) valence, a measurement of how positive (e.g., a pretty flower) or negative (e.g., a spider) an image is; (3) dominance, a measurement of how intimidating the image may be. Two similar equations were used to transform the three IAPS ratings into a normalized affective value for the FEAR and SEEKING module (Equation 4-1 and Equation 4-2, respectively).

Equation 4-1

$$FEAR_{Association} = \frac{(arousal + \varepsilon) * (10 - (valence + \varepsilon) - (dominance + \varepsilon))}{90}$$

Equation 4-2

$$SEEKING_{Association} = \frac{(arousal + \varepsilon) * ((valence + \varepsilon) + (dominance + \varepsilon) - 10)}{90}$$

The constant values (9 and 90) come from the structure of the IAPS system that allows each rating dimension to be a value from 1-9. The equations assume that a highly dominant and arousing image with a very low valence rating will result in a maximum FEAR value (with the center point between, for example, images with a completely negative valence and a completely positive valence, set at 5). This normalized FEAR value is processed in the module (as outlined in Chapter 3) and may eventually cause a change to the physiological system through connections from the affective-associations module to the FEAR module that cause changes in the physio module.

Procedural learning in the model

The IGT model includes a procedural memory learning component that uses production rule utility to determine rule selection during conflict resolution. Equation 4-3 is used to determine the probability of calling a specific production rule in ACT-R.

Equation 4-3

$$P(i) = \frac{e^{U_i / \sqrt{2s}}}{\sum_j e^{U_j / \sqrt{2s}}}$$

The summation in the denominator applies the equation in the numerator to all productions that satisfy the current conditions of the architectural state. The model uses reinforcement learning to update recent rule utilities when a reward and loss is presented in the task. The utility value after reinforcement is determined by Equation 4-4 (the utility update equation that is carried-over to ACT-R/Φ from the ACT-R architecture) and Equation 4-5 (an extended version of the ACT-R reward equation that takes into account the current value in the SEEKING and FEAR modules).

Equation 4-4

$$U_i(n) = U_i(n-1) + \alpha\big(R_i(n) - U_i(n-1)\big)$$

Equation 4-5

$$R_i(n) = r_j - \big(t_j - t_i\big) + \log(Value_{SEEKING}) - \log(Value_{FEAR})$$

In Equation 4-4, $U_i(n-1)$ represents the current utility value, α is the utility learning rate, and $R_i(n)$ is a reward that is determined by Equation 4-5. In Equation 4-5, r_j represents the reward received and $t_j - t_i$ is the temporal discount that is given to a reward so that the length of time between reward onset and a rule firing determines how much reward is applied to the utility value. $Value_{SEEKING}$ and $Value_{FEAR}$ in Equation 4-5 are reward offsets that take into account the current affective state of the model[12]. In the IGT model, the rewards and losses are transformed into a value that is reinforced according to Equation 4-6. This equation is a slightly modified version of the one that was proposed by Ahn et al. (2008) and proposed by Napoli and Fum (2010) to be used in ACT-R.

Equation 4-6

$$val(t) = W_{rew} * rew(t)^{\gamma} - W_{aversion} * loss(t)^{\gamma}$$

The reinforcement value ($val(t)$ here and r_j in Equation 4-5) is determined by the reward and loss ($rew(t)$ and $loss(t)$, respectively), an evaluation curvature function (γ), and a reward and aversion weight parameter for the reward and loss.

Partial matching is also enabled in the model. This allows a rule that only matches some of the conditions of the cognitive state to be included in a conflict resolution but with a utility that is discounted proportional to the amount of mismatch between the rule conditions and the actual architectural state. This essentially allows a rule that has been reinforced enough to fire even if a retrieval from declarative memory does not determine that that is the best action.

Declarative learning in the model

The IGT model uses the declarative memory and learning component of the ACT-R/Φ architecture to retrieve deck-keyboard location mappings. Declarative memory is also used to encode and retrieve instances (chunks) of a deck selected and the reward/loss received; this applies to both the actual reward and loss received and the transformed reward. Thus, any

[12] Higher reinforcement due to current affective state has been suggested by McClure et al. (2003) and is seen experimentally in a study by Winkielman, Berridge, and Wilbarger (2005).

declarative based determination of the best deck to choose may use the value given by the IGT or the transformed value found later to make a deck-decision (excluding the utility/reinforcement based influences discussed in the previous section). Both of these values are stored into declarative memory using the harvesting mechanism where any chunk that is cleared from the visual, aural, imaginal, goal, or retrieval buffer are automatically stored in declarative memory.

Declarative memory learning in ACT-R (and consequently ACT-R/Φ) is determined by the activation value of a chunk. Equation 4-10 shows how activation is determined in ACT-R/Φ. $Foffset_i$ and $Soffset_i$ are the two variables that differentiates the ACT-R activation equation from ACT-R/Φ. $Foffset_i$ increases the likelihood of a chunk i being retrieved given the current state of the FEAR module is similar to the state when chunk i was encoded; the $Soffset_i$ variable is used to accomplish the same functionality for the SEEKING module.

Equation 4-7, Equation 4-8, and Equation 4-9 are used to determine the variables that are used by the model to determine Equation 4-10; S_i (spreading activation) and P_i (declarative partial matching) are disabled in the model. The base-level activation of a chunk (B_i) is chunk i's base-level activation and is determined by the base-level constant value (β_i), the time since the chunk was last retrieved (t), and a decay parameter (d).

Equation 4-7

$$Foffset_i = \ln(abs\big(FEAR_{max} - abs(FEAR_{chunk} - FEAR_{current})\big) * e^{FEAR_{max}})$$

Equation 4-8

$$Soffset_i = \ln(abs\big(SEEKING_{max} - abs(SEEKING_{chunk} - SEEKING_{current})\big) * e^{SEEKING_{max}})$$

Equation 4-9

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) + \beta_i$$

Equation 4-10

$$A_i = B_i + S_i + P_i + Foffset_i + Soffset_i + \varepsilon_i$$

As discussed earlier, the affect offset parameters ($Foffset$ and $Soffset$) are determined by the current state of the affect system and the affective value attached to the chunk. The chunk that has the highest activation value (combination of the components described above) at the time of a retrieval request from the retrieval buffer will be recalled by the model as long as the activation value is above the retrieval threshold parameter.

Model results and predictions

The IGT model was run a total of 360 times, 120 times for each of the aversive, neutral, and positive image groups. Half of the model runs within each group were male and the other half were female. In the model, the distinction between male and female only comes into play with the affectively scored visual stimuli from the IAPS. Because the IGT model is a computational process model that completes a computerized version of the IGT, there are several ways the models resulting behavior can be reported. Below, I report the task-related deck selection behavior (model score, A-rate, and response time). These data are useful, because they can be directly compared to experimental data.

Model deck selections

Both the $IGT_{positive}$ and the $IGT_{negative}$ models had a higher average cumulative score than the $IGT_{neutral}$ model (Figure 4-4). The higher cumulative scores for the positive and negative IGT models are due in part to the higher arousal elicited by the positive and negative images (as compared to the neutral group). As arousal rises (up to a certain point), declarative and utility noise decreases, making it easier for the model to perform optimally given they have memories that accurately represent the state of the game.

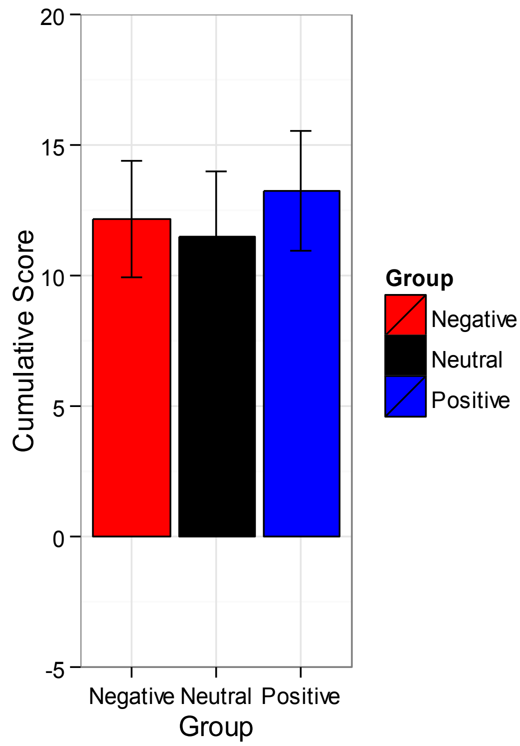

Figure 4-4. Cumulative score for models after the final block

On average the neutral image group IGT model (hereafter $IGT_{neutral}$) had the lowest average score in block 1 and $IGT_{positive}$ had the highest average score (Table 4-1) in the first block. In the last block, the $IGT_{negative}$ model had the highest score followed by the $IGT_{neutral}$ and $IGT_{positive}$ models.

Table 4-1. The score performance of the IGT model for each group

Model	Block	Score (Std Err)
Negative	1	-2.97 (0.78)
Negative	5	4.93 (1.14)
Neutral	1	-3.83 (0.88)
Neutral	5	4.21 (1.00)
Positive	1	-2.38 (0.85)
Positive	5	3.414 (1.076)

Scores for all groups increased from blocks 1-4 and showed a decrease in the final (5^{th}) block (Figure 5-3). Scores did not appear to differ in a significant way between groups; all groups exhibited similar learning over blocks and all groups still had a positive average score by the end of the task (5th block).

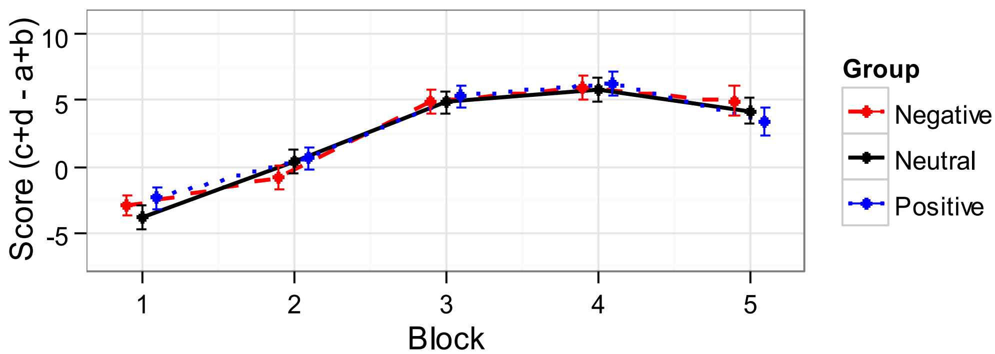

Figure 4-5. Model scores across 5 blocks

Figure 4-6 displays breakdown of model selections by deck for each block. As with the total score calculations, all model groups show a similar trend across blocks for all decks. Selections from deck B and C seem to show the biggest change from the first to the last block with deck B having the highest average selection count in the block 1 and decreasing from blocks 1 to 4 (Figure 4-6, 2^{nd} from top), and deck C showing a lower initial selection count, but increasing from blocks 1 to 4 (Figure 4-6, 3^{rd} from top).

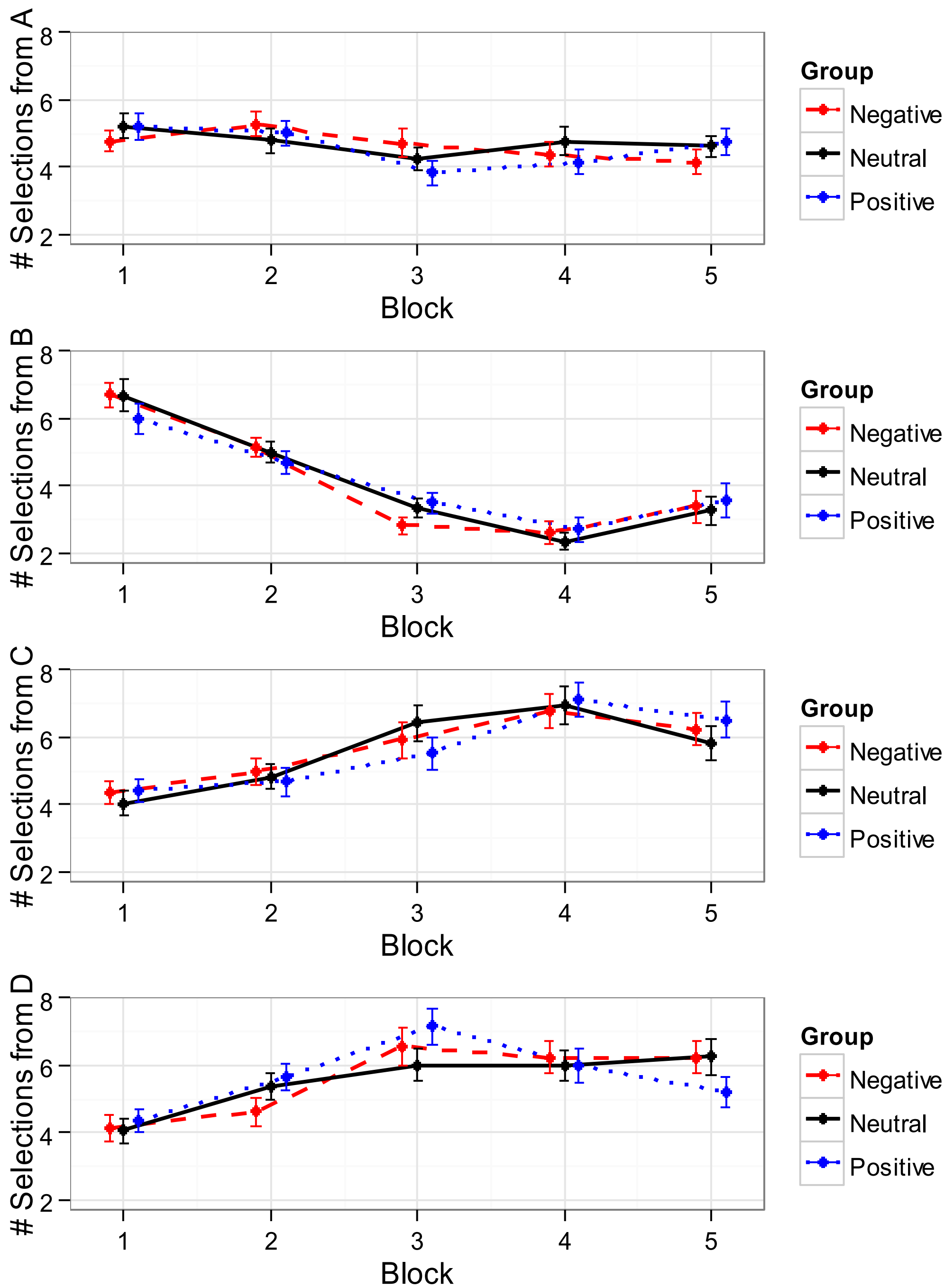

Figure 4-6. Number of selections from deck A (top), deck B (2nd from top), deck C (3rd from top), and deck D (bottom)

The rate of switching decks from trial to trial (A-Rate, Gonzalez & Dutt, 2011) did not appear to be different between model groups with all models on average switching more in the second block and then showing a general increase in repeated choices in the blocks 3-5 (Figure

4-7). The A-Rate for switching from bad decks to good decks also looked similar to the A-Rate for switching from good decks to bad decks (Figure 4-7, bottom).

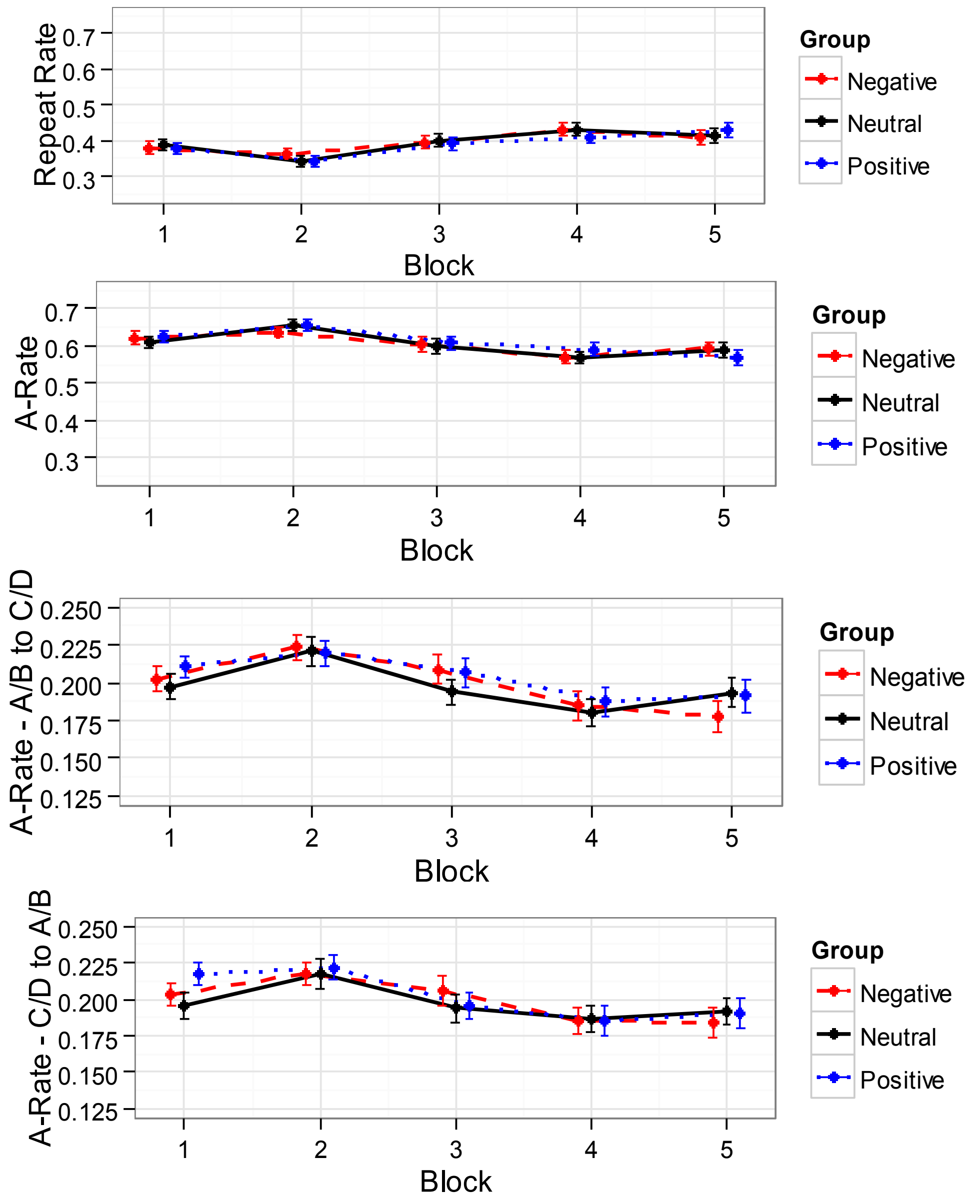

Figure 4-7: The proportion of cards selected from a deck that was also chosen in previous trial (top); the proportion of cards selected from a deck different from the previous trial (2nd from top);the proportion of cards selected from a bad deck in one trial and then from a good deck the next (3rd from top);the proportion of cards that were selected from a good deck in one trialal and then from a bad deck in the next (bottom).

Model response times

Response times are defined as the time taken during a trial to select a deck after the intertribal break message is removed and the trial has begun. Assuming a fairly consistent motor-action response time, i.e., time to select a deck after the decision has been made, differences in response time data represent differences in the time taken to decide which deck to select. Response times appear to be fairly consistent across blocks for all groups of models (Table 4-2) with all groups seeing a large decrease in response time from block 1 to block 2 (Figure 5-5).

Table 4-2. Average response time in the first and last blocks for each model.

Model	Blocks	Response Time (Std Err)
Negative	1	0.543 (0.010)
Negative	5	0.376 (0.012)
Neutral	1	0.535 (0.010)
Neutral	5	0.375 (0.014)
Positive	1	0.546 (0.012)
Positive	5	0.373 (0.012)

Response times followed the same pattern for all groups; this was an expected result because no part of the architectural extension or the model itself predict a difference due to subliminal affective visual stimuli. Average response times are slightly higher in block 1 for the good decks (c and d) than the bad decks (a and b) and slightly lower for the good decks in the final block. Response times are highest in the first block, because any rule that can be combined into one rule by the compilation mechanism will not have not yet accumulated a utility value high enough for it to be fired over the original parent rules. However, over time, more rules are compiled and their utilities reach values that have them fired instead of the original parent rules.

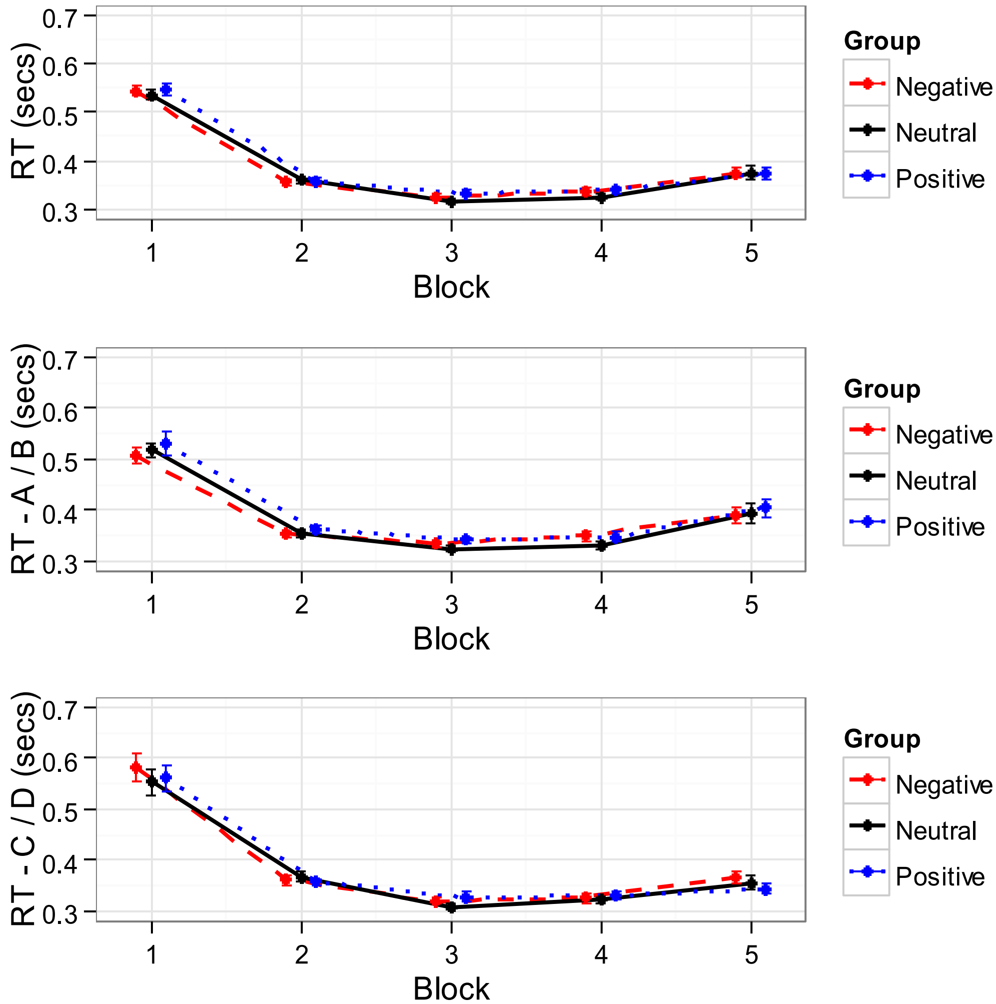

Figure 4-8. Average response time across blocks for all decks (top); average response time across blocks for bad decks (middle); average response time across blocks for good decks (bottom)

Summary

The model predicts that the cumulative score will be slightly higher for the negative and positive image groups than the neutral image groups by the end of the task. The model also does not predict a substantial difference of scores across blocks between groups despite the subsymbolic influences of visual stimuli, i.e. task-related stimuli largely control decision-making response. All groups are predicted to select from deck B the most in the first block and selection count from deck B to decrease from blocks 1 to 4. This decrease in deck B selection is the largest reason for score increasing from blocks 1 to 4 with this decrease being accompanied by a general increase in number of selections from decks C and D. The high amount of selections from deck B during the first block explains the higher repeat rate (Figure 4-7, top) as models tended to repeat selection of cards from deck B until the first large loss was received.

The higher amount of selections from deck B in block 1 (and the relatively stable number of selections from deck A) also results in a lower average response time in block 1 for bad decks (Figure 4-8, middle). The lower response time is essentially the result of production compilation as any rule resulting from the production compilation process in ACT-R has a better chance of being retrieved each time a reward is received (and the utilities of both the new and original rule are updated).

The IGT model predicts the process of selecting a deck during the IGT in addition to the actual choice. Though other IGT models have been made that use some ACT-R components (Napoli & Fum, 2010), this is the first model developed that describes the perceptual, declarative memory, and procedural memory processes involved during the task.

Initially, declarative memory is primarily used to determine which deck should be selected, with the first deck being recalled to have a positive outcome being selected. However, as the game progresses, reward-based procedural memory begins to take over and the model uses more of the utility and reinforcement mechanisms to make a deck selection. This means that if the last card selected from an advantageous deck resulted in a negative outcome (a higher reward than loss) that deck is more likely to be selected despite the declarative memory retrieved towards the end of the task, because of a higher utility for selecting a card from that deck. Thus, the model also predicts that punishment schedule should make a fairly large difference in the IGT, for example, the first two deck B punishments (which occur after making selections 9 and 14 from deck B) make a difference in the way the deck is selected because these punishments, when given in close proximity, affect rule noise (as generally predicted by Aston-Jones & Cohen, 2005), procedural memory reinforcement, and declarative memory of the deck reward-loss outcome.

This model uses some of the new extensions of the ACT-R/Φ architecture that were not used by previously developed ACT-R/Φ models. This model is useful because it completes the same computerized version of the IGT that is completed by participants during an experiment, making it more straightforward to compare model data to participant data and more straightforward to understand the ways that the model predicted the data well and the areas where the model can be improved to better fit the current data and predict future data. The model will be valuable to have as we continue to understand the ways cognitive processes can result from and cause shifts in strategies over time and what these shifts may mean for the way we make decisions in different environments.

In the next chapter, I present experimental results from a study that had participants complete the same modified IGT as the model. This data is then compared with model predictions in Chapter 6.

Chapter 5 An empirical study on decision-making and affect

As briefly discussed in the previous chapter, I ran an experiment to observe physiological and behavioral effects of subliminal affective visual stimuli during the Iowa Gambling Task. I also collected questionnaire data to ascertain personality metrics and current affective state metrics for each participant. Below I report the experimental method and results from the experiment.

Method

Participants

110 undergraduate students were recruited as participants for this study (58 male and 52 female). Both eye-tracking and behavioral data were collected for all 110 of the participants, but Electrodermal Activity (EDA) data was only collected for the final 77 participants because the measurement device was acquired after the study began. All participants were given course extra credit for participation.

Materials

I used a 19 inch Dell monitor with 1680x1050 native resolution for task presentation. To run the underlying software, either a Fujitsu t901 running Windows 7 or Dell Optiplex 755 running Ubuntu Linux 12.04LTS were used. The Fujitsu used Matlab (R2009a) to run the software, while the Dell used Octave (3.2.4).

I developed a computerized version of the Iowa Gambling Task (Bechara et al., 1999) to run in Matlab/Octave that uses the Psychtoolbox matlab/octave extensions (Brainard, 1997). Psychtoolbox extensions were used due to their high timing accuracy, community support, and cross-platform availability.

The Iowa Gambling Task (IGT) is a decision-making task that consists of four decks (A, B, C, and D), one of which can be selected during each trial. Participants complete 100 trials (without knowing the amount of trials they will complete when completing the task) and can make a maximum of 40 card selections from any deck. In this version of the IGT, participants had a maximum of 3.5 seconds to select a deck and if they failed to make a selection in the allotted time on any trial, a random deck was selected for them. After a deck selection was made, a simulated card was flipped over revealing either a blank red or blank black card. The card was shown for 4 seconds followed by the display of the reward and loss which were displayed for 3.5 seconds. If participants selected from deck A or deck B, an image from a predetermined image-set was flashed in place of the background image (the bowl in Figure 5-1) for 17ms and then the reward and loss were displayed to the participant; however, if participants selected from deck C or deck D, a plain gray background was shown in place of the background image for 17ms. Each intertrial break lasted 3.5 seconds, except for after the 20th, 40th, 60th, and 80th trials, which had a longer intertrial break (see the next section for more details). Participants used a version of the Iowa Gambling Task that included a fixed reward and punishment schedule that is the same as the schedule used for the *original* IGT by Bechara, Tranel, and Damasio (2000).

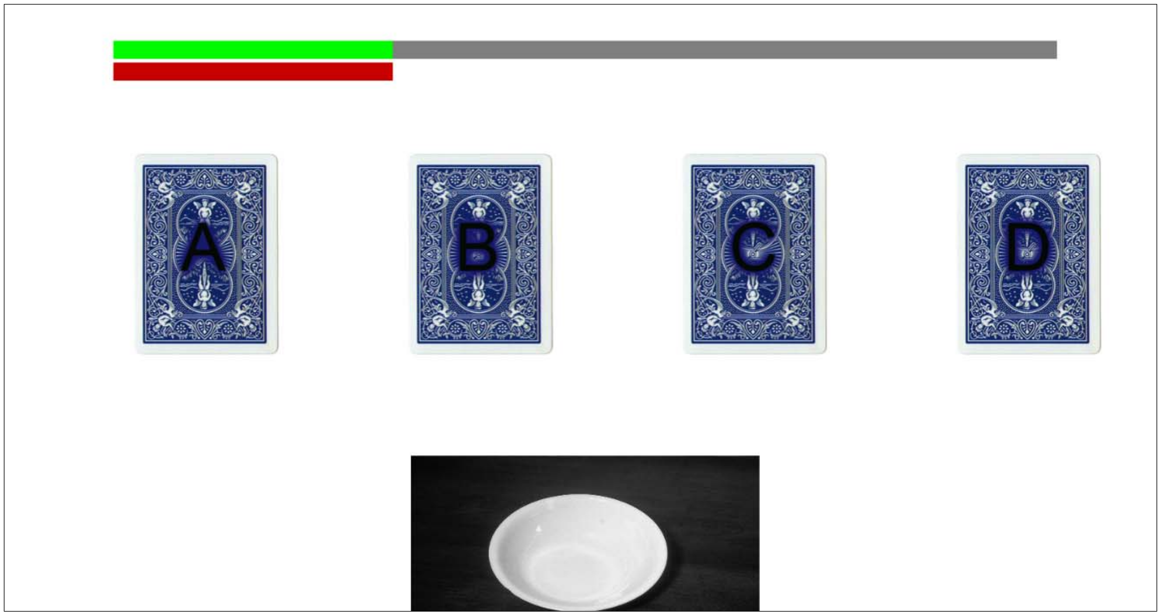

Figure 5-1. A screenshot of the computerized Iowa Gambling Task completed by participants

The visual stimuli presented during the IGT were obtained from the International Affective Picture System (IAPS; Lang et al., 1997). The IAPS is a set of pictures (commonly used in studies using affective pictures) rated on a scale (from 1 to 9) for arousal, valence, and dominance; these pictures are available online through a license.

While completing the task, participants' eye behavior was recorded using an ISCAN ETL-500 eye tracker. The eye-tracker delivered real-time data to the task-presentation computer using a serial connection with a 57600 baud rate.

EDA data was recorded for some participants using the Affectiva Q Sensor (http://www.affectiva.com/q-sensor).This wireless sensor was strapped to the participants palm. The data were recorded with a 32hz sampling rate

Participants completed the positive and negative affect schedule (PANAS; Watson, Clark, & Tellegen, 1988). This scale captured the participants affective state (positive or negative) after they arrived and signed the consent form and after they completed the IGT; a copy of the PANAS was given to participants in available in Appendix B.

Participants also completed the Affective Neuroscience Personality Scales (ANPS). This questionnaire can be used to gather personality ratings in dimensions related to primary-process affect theory (Davis & Panksepp, 2011). For this study, participants' ratings in the SEEKING, FEAR, and PLAY dimensions were analyzed. The version of the PANAS used for the study is available in Appendix B.

Design and procedure

The experiment was designed to understand how different sets of subliminal images (not integral to the task) may affect decision-making and physiological behavior exhibited during the task. Each participant was assigned to one of three groups that determined which set of images they were shown: (a) a *negative* image group that consisted of images with low a rated valence, a high arousal, and were either a snake, spider, or an attacking dog; (b) a *neutral* image group that consisted of images with a medium rated valence and a low arousal (c) a *positive* image group that consisted of images with a high rated valence and a high arousal. Table 5-1 lists the images used in image sets used by the different groups. The set of images were different depending on the sex of the participant, because IAPS provided ratings that showed a difference in some images depending on the rater's sex (this was most noticeable in the *positive* images).

An image from a set (depending on the group the participant was placed into) was shown in place of the background image of the box where the reward and losses were shown for 17ms when the participant selected from deck A or deck B. If the participant made a selection from deck C or deck D, a plain gray background image was shown for 17ms. Directly after this 17ms, the original background image was restored and the reward and loss that the participant received in response to their deck selection was presented in the same box. All images used from the image set, as well as the background image used throughout the task, were converted to grayscale.

Table 5-1. The IAPS images (and the accompanying average valence, arousal, and dominance rating) used in the experiment.

IAPS Picture Number	Picture-Set	Valence (std)	Arousal (std)	Dominance (std)
1050	$\text{Negative}_{\text{Male}}$	3.90 (2.28)	6.84 (1.55)	3.14 (1.78)
1202	$\text{Negative}_{\text{Male}}$	4.03 (1.81)	6.20 (1.45)	4.80 (2.32)
1220	$\text{Negative}_{\text{Male}}$	3.88 (1.76)	5.40 (2.50)	5.02 (2.60)
1304	$\text{Negative}_{\text{Male}}$	3.89 (1.60)	6.39 (1.90)	3.18 (1.47)
1525	$\text{Negative}_{\text{Male}}$	3.55 (1.59)	6.14 (2.31)	3.79 (2.24)
1670	$\text{Neutral}_{\text{Male}}$	6.51 (1.76)	2.95 (1.67)	5.90 (187)
7006	$\text{Neutral}_{\text{Male}}$	4.65 (1.10)	2.08 (1.58)	5.83 (2.03)
7010	$\text{Neutral}_{\text{Male}}$	4.95 (1.43)	1.55 (1.36)	7.25 (2.31)
7080	$\text{Neutral}_{\text{Male}}$	5.43 (1.26)	1.98 (1.63)	7.47 (2.1)
7175	$\text{Neutral}_{\text{Male}}$	4.78 (1.18)	1.55 (0.96)	6.51 (2.08)
4180	$\text{Positive}_{\text{Male}}$	8.21 (1.34)	7.43 (1.97)	6.41 (2.54)
4210	$\text{Positive}_{\text{Male}}$	8.25 (1.30)	7.80 (1.90)	6.97 (2.42)
4232	$\text{Positive}_{\text{Male}}$	7.88 (1.10)	7.52 (1.51)	6.81 (2.02)
4664	$\text{Positive}_{\text{Male}}$	7.99 (1.25)	7.72 (1.45)	6.54 (1.93)
8501	$\text{Positive}_{\text{Male}}$	8.14 (1.24)	6.86 (2.16)	5.61 (2.67)
1050	$\text{Negative}_{\text{Female}}$	3.02 (1.93)	6.90 (1.82)	3.02 (2.08)
1120	$\text{Negative}_{\text{Female}}$	3.03 (1.74)	7.20 (1.86)	3.22 (2.42)
1201	$\text{Negative}_{\text{Female}}$	2.93 (1.81)	6.87 (2.09)	3.82 (2.26)
1202	$\text{Negative}_{\text{Female}}$	2.98 (1.65)	5.80 (2.47)	3.92 (2.34)
1525	$\text{Negative}_{\text{Female}}$	2.67 (1.74)	6.86 (2.16)	2.55 (1.99)
1670	$\text{Neutral}_{\text{Female}}$	5.88 (1.84)	3.52 (2.05)	5.40 (1.74)
7004	$\text{Neutral}_{\text{Female}}$	5.14 (0.59)	1.94 (1.60)	6.65 (2.10)
7010	$\text{Neutral}_{\text{Female}}$	4.92 (0.48)	1.97 (1.58)	6.13 (2.29)
7012	$\text{Neutral}_{\text{Female}}$	4.97 (0.87)	2.88 (1.90)	6.17 (1.68)
7175	$\text{Neutral}_{\text{Female}}$	4.95 (0.80)	1.87 (1.48)	6.44 (2.01)
4505	$\text{Positive}_{\text{Female}}$	7.20 (1.16)	6.46 (1.77)	5.85 (2.00)
4525	$\text{Positive}_{\text{Female}}$	7.67 (1.24)	6.70 (1.69)	5.49 (1.63)
4660	$\text{Positive}_{\text{Female}}$	7.22 (1.40)	6.31 (1.95)	5.39 (1.66)
8001	$\text{Positive}_{\text{Female}}$	7.46 (1.28)	6.62 (1.83)	5.80 (1.92)
8501	$\text{Positive}_{\text{Female}}$	7.67 (1.97)	6.02 (2.50)	6.49 (2.31)

Before participating in the study, all participants read and signed a consent form approved by the Office of Research Protections (ORP) at Penn State. After consenting to the form, all participants filled out a PANAS questionnaire. All participants who had there EDA recorded were then fitted with the Q Sensor EDA device.

The participant was then told they were going to be completing a “gambling style task” and fitted with eye-tracking equipment. After calibrating the eye-tracker, the participant was presented with instructions on the screen and asked to read the instructions and notify the experimenter when they had read all of the instructions. They then completed an introductory task in which they practiced selecting different decks using the keyboard to allow them to get acclimated to the card selection procedure. Once the participant finished this portion of the task, they completed the actual IGT itself.

Every 20 blocks, participants were able to take a short break. During this break, they remained seated and were asked one question to get an idea of what explicit knowledge of they may have of the task: “*Tell me all you know about the game related to winning or losing. Keep in mind there is no wrong answer.*” After answering the question, participants were instructed to hit a certain button on the keyboard when they were ready to begin to next block of trials.

After the IGT was completed, participants filled out another PANAS questionnaire and then had the EDA device removed from their palm. Participants were then asked “Did you discover anything new by the end of the game?” and were partially debriefed on the task itself. Participants then completed the ANPS questionnaire and were fully debriefed before ending the study session.

Experiment results and discussion

The first six of the 110 participants were treated as pilot subjects for the study. Thus, data from 104 participants were used for analysis. The results for task performance, electrodermal activity (EDA), eye-behavior (not reported here), and questionnaire data were found.

Task behavior and performance

Task-related performance data was collected from 104 participants during the Iowa Gambling Task. Matlab scripts were written to preprocess the raw data before analysis. Any trials that resulted in an automatic response due to the participant not making a deck selection within the allotted 3.5s were excluded from analysis; this resulted in 4% of the trials being excluded from analysis. Additionally, if more than 19% of a participant's trials (20 or more) resulted in an automatic deck selection, all of that participant's data were excluded. This filter process removed 11 participants' data from further analysis, so data from 93 total participants were analyzed. The *negative*, *neutral*, and *positive* (image) groups each had 31 participants. Data were analyzed using the R statistical software environment. ANOVAs were run in R using the *nlme* package (Pinheiro, Bates, DebRoy, & Sarkar, 2013).

Participant deck selections

Participants selected 100 total cards from a deck and deck selection analysis was split into five blocks (1-20, 21-40, etc.) to match scoring analysis in previous studies that have explored behavior during the IGT (e.g., Bechara et al., 1999; Stocco et al., 2009; Turnbull et al., 2005). Score in the IGT is typically calculated by subtracting the total number of selections of the *bad* decks (decks A and B) from number of selections of the *good* decks (decks C and D).

The positive image group recorded the highest average cumulative score (6.1) with the neutral image group having a smaller score (1.9) and the negative image group having a negative average cumulative score (-0.4). An ANOVA failed to reveal a significant difference between groups ($F(2,91) = 0.942, p = .39$).

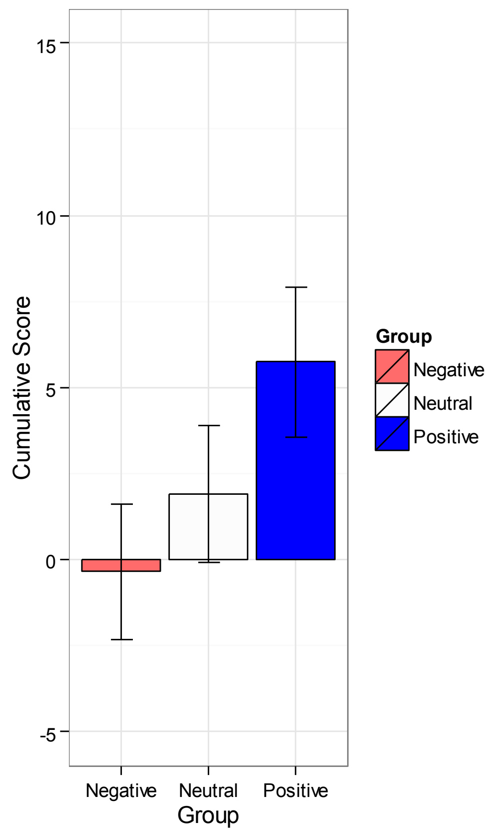

Figure 5-2. Cumulative score ($\pm SEM$) for all participants after the final block

By the 5th block, the negative, neutral, and positive groups had a positive score (Table 5-2, Figure 5-3) indicating that, on average, the participants in each group were selecting from *good* decks by the end of the task.

Table 5-2: Mean score of participants in all of the blocks by group. Standard errors are in presented in the parenthesis

Group	Block 1	Block 2	Block 3	Block 4	Block 5
Negative	-4.2 (1.4)	0.2 (1.3)	1.1 (1.5)	-0.2 (1.3)	2.7 (1.4)
Neutral	-3.5 (1.1)	-0.3 (0.9)	2.4 (0.8)	1.5 (1.0)	1.8 (1.1)
Positive	-3.3 (1.3)	-0.5 (1.2)	3.0 (1.2)	2.6 (1.1)	4.0 (1.0)

Scores increased for all groups from blocks 1-3 and blocks 4-5, but decreased from blocks 3-4 (Figure 5-3). Scores did not appear to differ in a significant way between groups; all groups exhibited similar learning over blocks and all groups had a positive average score by the end of the task (5th block).

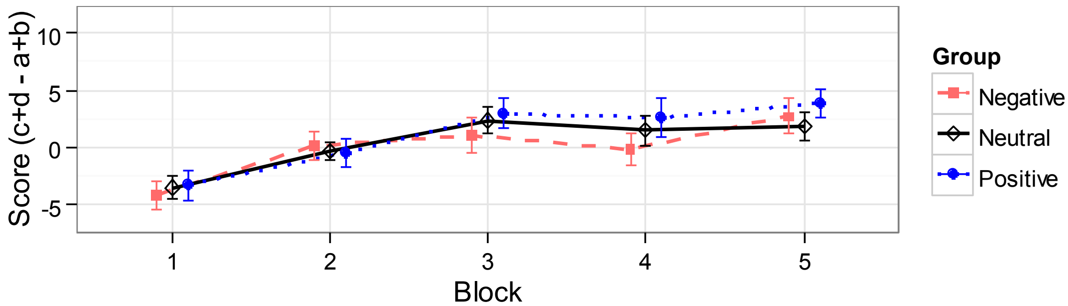

Figure 5-3: IGT score ($\pm SEM$) across blocks for each group. The vertical bars at each point represent the standard error.

A $3X5$ (group by block) mixed factor ANOVA of participant score revealed a highly significant effect of block ($F(4,\ 360)\ =\ 13.22, p\ <\ .0001$) on score, but did not reveal a significant group ($F(2,\ 90)\ =\ 0.81,\ p\ =\ .4$) or a group:block interaction ($F(8,\ 360)\ =\ 0.40,\ p\ =\ .9$) effect.

Participant A-rate

Given that this task may be considered a type of repeated choice paradigm (e.g., see Gonzalez & Dutt, 2011, for an example repeated choice paradigm), the proportion of deck selections that were alternating (e.g., select a card from deck A the 1st trial and deck B the next) are also an important indicator of participant behavior over the course of the task; hereafter this metric is referred to as A-rate, following Gonzalez and Dutt (2011). There was a decrease in overall A-rate (switching between any of the decks) for the negative and neutral groups. (Table 5-3; Figure 5-4, top). The positive group had a slight decline in A-rate in blocks 1-4, but showed an increase in switching between decks in the last block (Figure 5-4, top).

Table 5-3. Mean A-rate across blocks by participant group. Standard error are presented in the parenthesis.

Group	Block 1	Block 2	Block 3	Block 4	Block 5
Negative	0.653 (0.029)	0.647 (0.028)	0.644 (0.032)	0.567 (0.030)	0.545 (0.031)
Neutral	0.679 (0.037)	0.700 (0.031)	0.640 (0.026)	0.579 (0.036)	0.615 (0.038)
Positive	0.650 (0.022)	0.669 (0.027)	0.604 (0.034)	0.596 (0.025)	0.650 (0.026)

A-rate for switching from good to bad decks, and vice versa, decreased between blocks 1-4 for the neutral and positive groups and decreased between blocks 2-5 for the negative group (Figure 5-4, bottom).

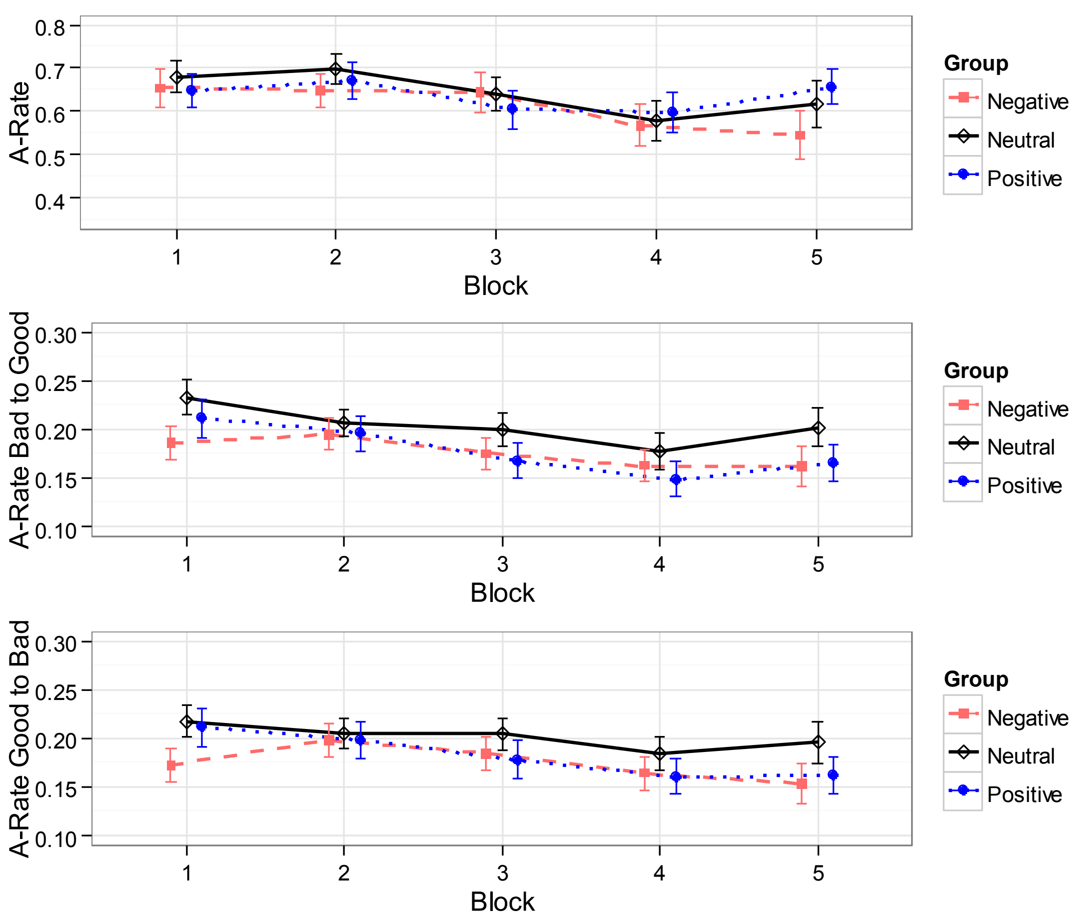

Figure 5-4. The proportion of cards selected from a deck (±*SEM*) that is different than the previous trial (top); proportion of cards selected from a bad deck in one trial and then from a good deck the next (middle); proportion of cards that were selected from a good deck in one trial and then from a bad deck in the next (bottom).

A $3X5$ mixed-factor ANOVA of overall A-rate revealed a highly significant block effect ($F(4, 364) = 4.554, p < .005$). An ANOVA of A-rate from bad decks to good decks revealed a highly significant block effect ($F(4, 364) = 5.702, p < .0005$), while an ANOVA of A-rate from good decks to bad decks also revealed a significant block effect

($F(4, 364) = 3.765, p < .01$)).

Table 5-4. ANOVA results for A-rate for all participants. Bold values signify significant factors

Response-Variable	Group	p_{Group}	Block	p_{Block}	Group:Block	$p_{Group:Block}$
A-rate	$F(2,90) = 0.22$	$p = .8$	$F(4,360) = 4.53$	$\mathbf{< .005}$	$F(8,360) = 0.94$	$p = .5$
A-rate Bad to Good	$F(2,90) = 1.46$	$p = .2$	$F(4,360) = 5.61$	$\mathbf{< .0005}$	$F(8,360) = 0.49$	$p = .9$
A-rate Good to Bad	$F(2,90) = 1.09$	$p = .3$	$F(4,360) = 3.64$	$\mathbf{< .01}$	$F(8,360) = 0.67$	$p = .7$
A-rate Within Good or Bad Decks	$F(2,90) = 1.36$	$p = .3$	$F(4,360) = 5.10$	$\mathbf{= .0005}$	$F(8,360) = 0.67$	$p = .7$

Participant response times

Response times are defined as the time taken during a trial to select a deck after the intertrial break message is removed and the trial has begun. Assuming a fairly consistent motor-action response time, i.e., time to select a deck after the decision has been made, differences in response time data represent differences in the time taken to decide which deck to select.

Response times appear to be fairly consistent across blocks for all groups with the negative image group having the most noticeable change from block 1 to block 5 (Figure 5-5); the negative group had the highest average response time in block 1 and the lowest response time in block 5.

Table 5-5. Average response time in seconds across blocks for each group. Standard errors are in parenthesis

Group	Block 1	Block 2	Block 3	Block 4	Block 5
Negative	0.90 (0.04)	0.81 (0.03)	0.83 (0.04)	0.73 (0.03)	0.76 (0.05)
Neutral	0.84 (0.04)	0.80 (0.03)	0.83 (0.03)	0.83 (0.4)	0.81 (0.04)
Positive	0.82 (0.05)	0.79 (0.03)	0.85 (0.03)	0.81 (0.04)	0.79 (0.05)

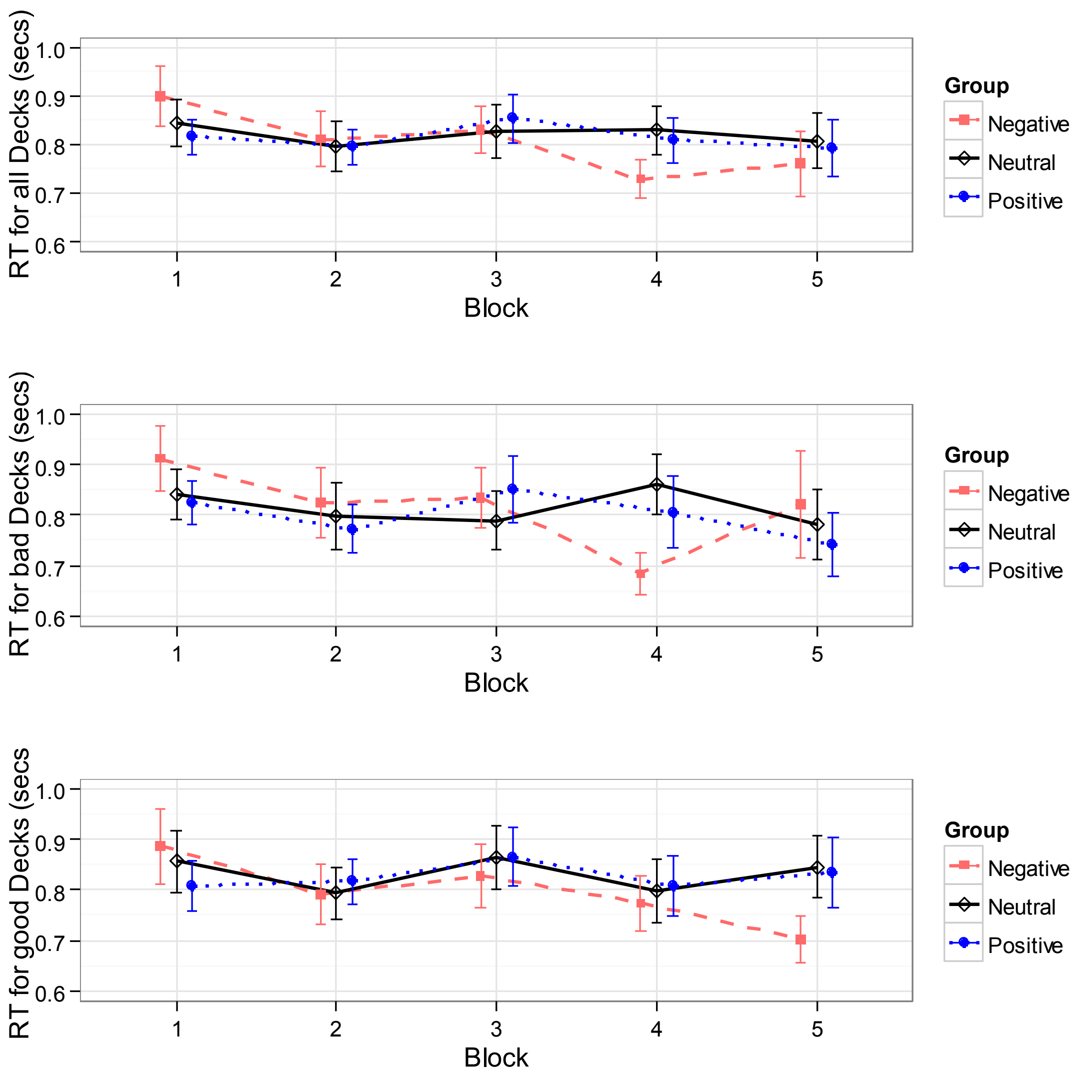

Figure 5-5. Average response time ($\pm SEM$) across blocks for all decks (top); average response time across blocks for bad decks (middle), and average response time across blocks for good decks (bottom)

Response times were log-transformed, $\log(RT)$, for all ANOVA analyses. A $3X5$ mixed factor ANOVA of response times for all decks revealed a significant block effect for overall ($F(4,360) = 2.95, p < .05$) and bad deck ($F(4,347) = 2.80, p < .05$) response times (Table 5-6).

Table 5-6. ANOVA results for participants' response time. Bold values signify significant factors

Response-Variable	Group	p_{Group}	Block	p_{Block}	Group:Block	$p_{Group:Block}$
$RT_{All,\ All\ decks}$	$F(2,90) = 0.10$	$p = .9$	$F(4,360) = 2.95$	$\boldsymbol{p < .05}$	$F(8,360) = 0.88$	$p = .5$
$RT_{All,\ Bad\ decks}$	$F(2,90) = 0.02$	$p = 1$	$F(4,347) = 2.80$	$\boldsymbol{p < .05}$	$F(8,347) = 1.25$	$p = .3$
$RT_{All,\ Good\ decks}$	$F(2,90) = 0.28$	$p = .8$	$F(4,358) = 1.41$	$p = .2$	$F(8,358) = 0.67$	$p = .7$

Electrodermal activity

Electrodermal Activity (EDA) Data was collected to assess peripheral nervous system response to different events during each trial. Data from 10 participants had to be discarded due to hardware setup malfunction. The data were also filtered to exclude any participants who were excluded from task-related behavioral analysis. This process left 54 total participants for data analysis, 34 male (11 in the negative group, 11 in the neutral group, and 12 in the positive group) participants and 20 female (7 in the negative group, 6 in the neutral group, and 7 in the positive group) participants.

Each trial was separated into three 4-second epochs for analysis: (a) an *anticipatory* phase that starts at the beginning of the intertrial break; (b) a *deck* phase that starts after the participant selects a deck; (c) a *reward/loss* phase that starts after the rewards and losses are presented on the screen. A batch function that uses continuous decomposition analysis (CDA; Benedek & Kaernbach, 2010) was run in from the ledalab data analysis suite within Matlab to preprocess the raw EDA data; the script parameters used in the function call are presented in Appendix B. The script was used to extract mean skin conductance response (SCR) for responses above 0.05 microsiemens (μS) for all phases of each trial[13].The skin conductance response (SCR) data were then standardized using the formula $y = \log(x)$ for ANOVA-based data analysis.

Figure 5-6 shows the mean anticipatory SCR (μS/sec) found for each group for bad decks and good decks. Anticipatory SCR appears to be similar between groups for *good* decks. However the average SCR for exhibited by the negative group for *bad* decks is higher than in the neutral and positive groups.

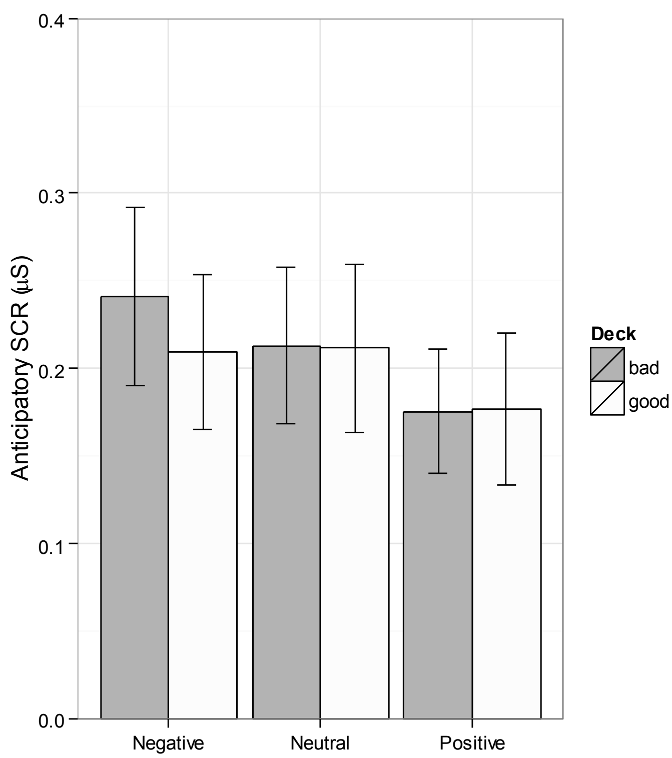

Figure 5-6. The mean ($\pm SEM$) anticipatory SCR (μS/sec) of each group over the entire IGT.

[13] There were other data measures that were found using the script, but those are not reported in this dissertation.

Despite the visual difference between groups in Figure 5-6, a $3X1$ ANOVA for anticipatory SCR for bad decks between groups failed to reveal group as a significant factor ($F(2,51) = 0.80$, $p = .5$). A $3X1$ ANOVA of anticipatory SCR for good decks between groups also failed to reveal group as a significant factor ($F(2,51) = 0.22$, $p = .8$).

Figure 5-7 shows the response to different outcomes for bad decks and good decks (e.g., trials where the outcome after selecting from a bad deck resulted in an overall loss in money, or a *loss*). The figure shows that the negative group showed the greatest response to loss when selecting from a bad deck and showed the greatest response in for all reward/loss-deck combinations except in response to losses after selecting from a good deck. The positive group showed the opposite trend, with all responses to reward/loss-deck combinations being smaller than the negative and neutral groups. The positive group did, however, show a slightly higher response to losses after selecting from a good deck.

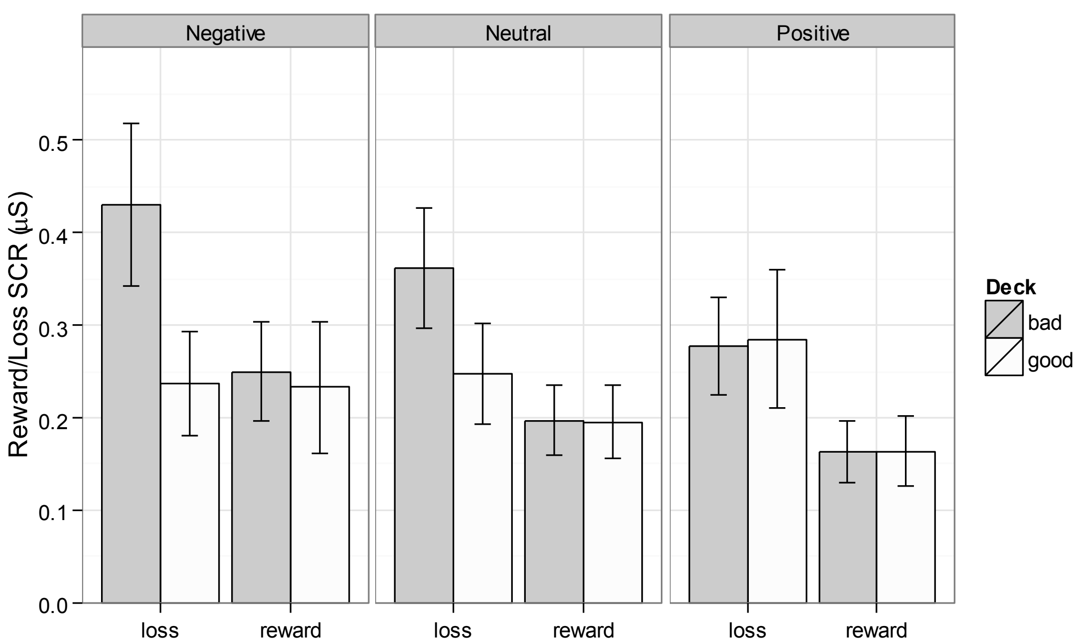

Figure 5-7. The mean (±SEM) reward/loss SCR (µS/sec) in response to trials where outcome was a loss was greater than the reward (*loss*) and in response to trials where the outcome was a reward was greater than the loss (*reward*) for each group over the entire IGT.

A $3X1$ ANOVA of SCR response to losses when selecting from a bad deck did not reveal group as a significant factor ($F(2,51) = 1.43, p = .25$). Group also failed to be a significant factor for SCR response to rewards when selecting from a bad deck ($F(2,51) = 0.99, p = .38$). A $3X1$ ANOVA of the difference between the response to a loss when selecting from a bad deck and the response to a loss when selecting from a good deck revealed group as a marginally significant factor ($F(2,51) = 2.53, p = .09$).

A plot of SCR during the *anticipatory* phase for all decks (Figure 5-8, top) across 5 blocks shows a similar pattern of SCR for all groups through blocks 1 and 2. In the 3rd block, the negative group shows a spike in SCR while the neutral and positive groups only show a slight increase in SCR. Figure 5-8 (bottom) shows the anticipatory SCR divided by deck selected. The negative group showed a similar trend across blocks 1-4 for bad and good decks (albeit with a higher response to bad decks). The negative group also had the highest SCR for bad decks in all blocks but block 1. Interestingly, the neutral group exhibited is a spike in mean SCR for bad decks in block 3, but does not show a similar increase for good decks. The positive image showed a similar SCR pattern across blocks for both bad and good decks.

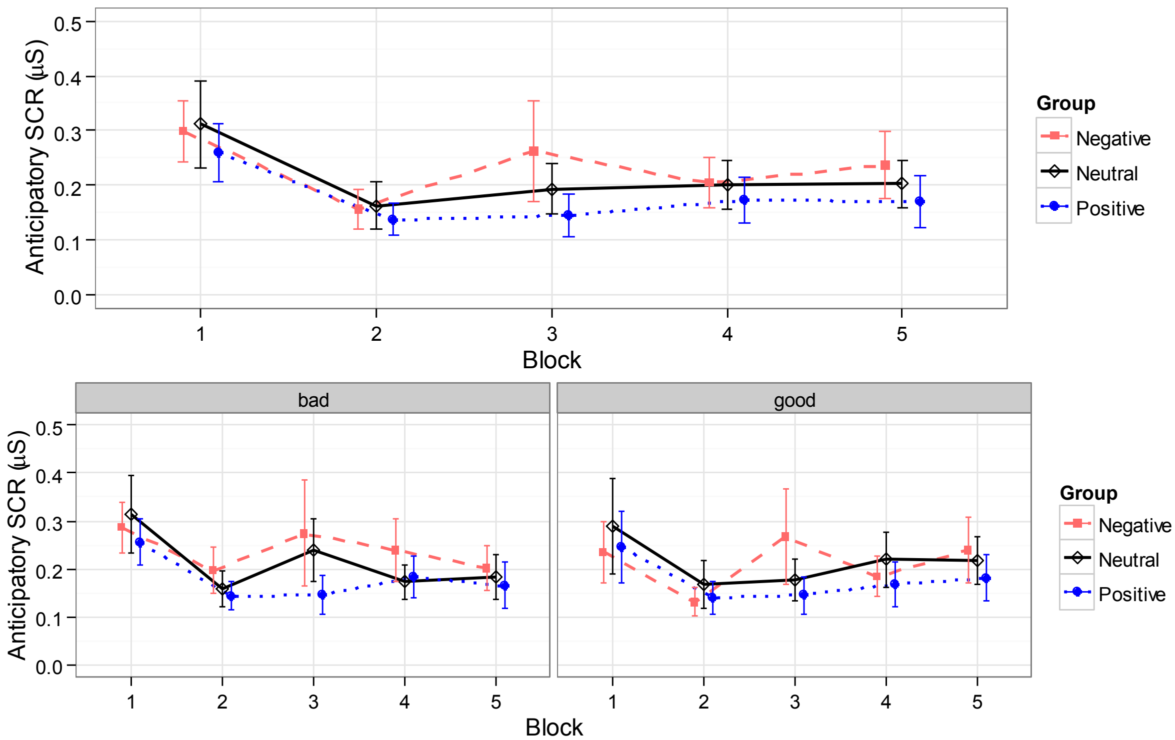

Figure 5-8. Anticipatory SCR results (with SEM error bars) for all participants for all decks (top) and bad and good decks (bottom).

Table 5-7 presents the results found from running a $3X5$ mixed factor ANOVA with SCR as the response variable with all three phases (anticipatory, deck, reward/loss) of data. The table also gives an account of the data subset by type of deck selected, bad (decks a and b) or good (decks c and d). Anticipatory SCR was not significantly different between groups ($F(2,51) = 0.35, p = .70$) or within the group*block interaction ($F(8,203) = 0.60, p = .78$), but was significantly different within blocks ($F(4,203) = 14.19, p < .0001$). All phases of the trials followed the same general pattern with blocks being the only significant factor; running an ANOVA on the data based on deck selection (good or bad) also revealed the same pattern of results.

Table 5-7. ANOVA results for SCR (log-transformed) of participants. Bold values signify significant factors

Response-Variable	Group		Block		Group:Block	
$SCR_{Anticipatory}$	$F(2,51) = 0.35$	$p = .7$	$F(4,203) = 14.19$	$\boldsymbol{p < .0001}$	$F(8,203) = 0.60$	$p = .8$
SCR_{Deck}	$F(2,51) = 0.41$	$p = .7$	$F(4,203) = 4.02$	$\boldsymbol{p < .005}$	$F(8,203) = 0.41$	$p = .9$
$SCR_{Reward/Loss}$	$F(2,51) = 0.40$	$p = .7$	$F(4,203) = 5.09$	$\boldsymbol{p < .001}$	$F(8,203) = 0.15$	$p = 1$
$SCR_{Anticipatory,\ Bad}$	$F(2,51) = 0.58$	$p = .6$	$F(4,193) = 5.81$	$\boldsymbol{p < .0005}$	$F(8,193) = 0.93$	$p = .6$
$SCR_{Deck,\ Bad}$	$F(2,51) = 0.53$	$p = .6$	$F(4,193) = 2.03$	$\boldsymbol{p = .09}$	$F(8,193) = 0.63$	$p = .8$
$SCR_{Reward/Loss,\ Bad}$	$F(2,51) = 0.45$	$p = .6$	$F(4,193) = 3.41$	$\boldsymbol{p < .05}$	$F(8,193) = 0.19$	$p = 1$
$SCR_{Anticipatory,\ Good}$	$F(2,51) = 0.13$	$p = .9$	$F(4,201) = 3.68$	$\boldsymbol{p < .01}$	$F(8,201) = 0.32$	$p = 1$
$SCR_{Deck,\ Good}$	$F(2,51) = 0.15$	$p = .9$	$F(4,201) = 4.32$	$\boldsymbol{p < .005}$	$F(8,201) = 0.44$	$p = .9$
$SCR_{Reward/Loss,\ Good}$	$F(2,51) = 0.24$	$p = .8$	$F(4,201) = 4.52$	$\boldsymbol{p < .005}$	$F(8,201) = 0.46$	$p = .9$

Questionnaires

A brief questionnaire, the positive and negative affect scale (PANAS), was given to the participant before and after the IGT was completed, to get an affective state measurement. The PANAS provides a metric of participants' current positive and negative affective state. After the task was completed participants were also asked to complete a personality-assessment questionnaire, the affective neuroscience personality scales (ANPS).

PANAS

Table 5-8 presents descriptive statistics of participants PANAS positive affect scores before and after completing the IGT. Overall, the positive image group displayed the highest score both before and after the task, though none of the scores were significantly different.

Table 5-8. Average positive PANAS scores before and after the IGT

Group	Valence	Time	Score (Std Err)
Negative (N=31)	Positive	Before	27.5 (1.3)
Negative (N=31)	Positive	After	24.3 (1.3)
Neutral (N=30)	Positive	Before	28.1 (1.0)
Neutral (N=30)	Positive	After	26.0 (1.5)
Positive (N=30)	Positive	Before	28.5 (1.0)
Positive (N=30)	Positive	After	27.2 (1.2)

A bar chart of the positive score (Figure 5-9, top) shows a slight difference in positive scores between groups before the IGT and a more pronounced difference (with the positive image group having the highest average score) after the IGT. A bar chart of the average difference between scores before and after the IGT for participants shows a general decrease in score after the IGT with the negative group having the highest decline in score and the positive group having the lowest decline.

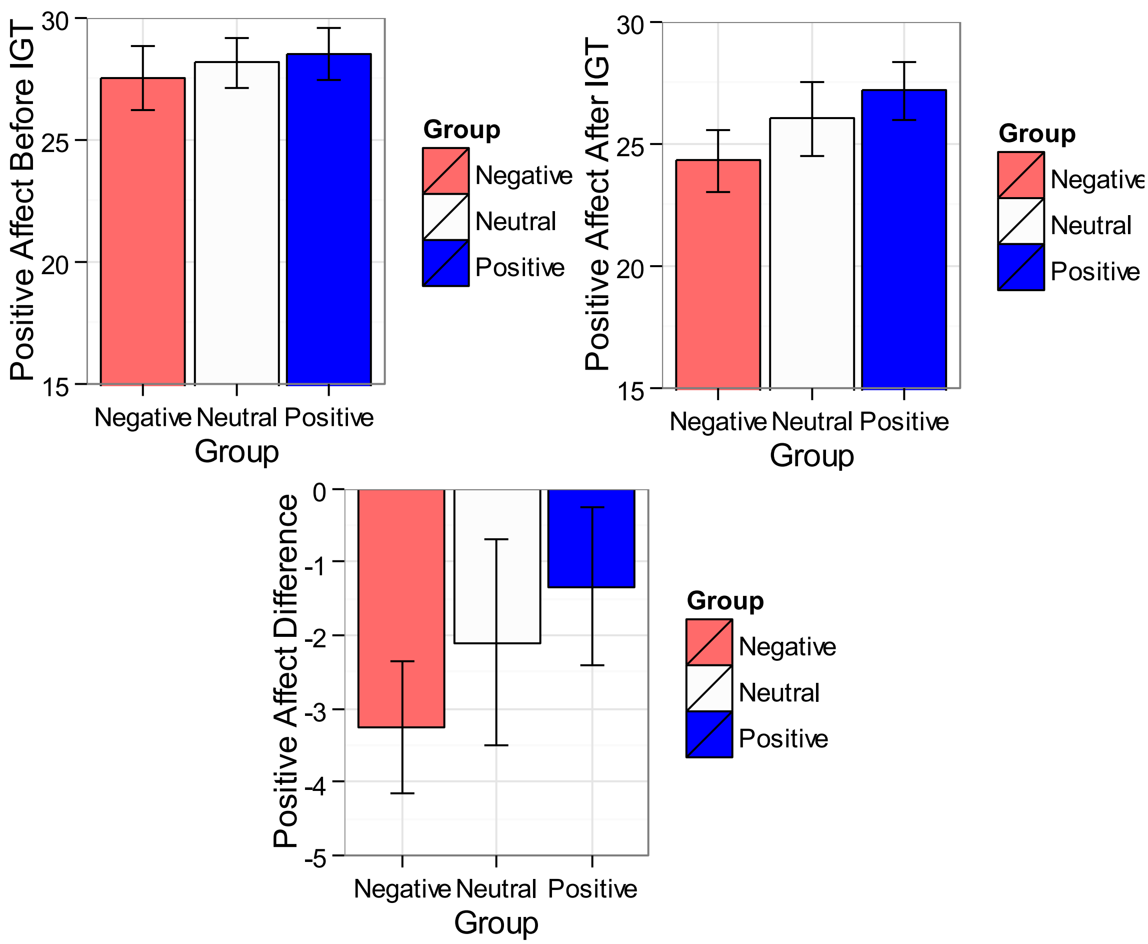

Figure 5-9. Mean PANAS positive affect scores (± SEM) before (top-left) and after (top-right) the IGT, and the average difference between those scores for each participant (bottom)

An ANOVA of the average difference of positive affect scores did not revealed group as a significant factor ($F(2,88) = 0.72,\ p = .49$). Individual paired t-tests comparing the scores of participants before and after the IGT were run for each group. A two-tailed paired t-test revealed a highly significant difference between the scores for the negative image group ($t(30) = 3.60,\ p = .001$), but the difference was not significant for the neutral group

($t(29) = 1.49$, $p = .2$) or the positive group ($t(29) = 1.23$, $p = .2$). A one-sided paired t-test of scores showed a highly significant decrease in PANAS positive scores from before the IGT to after for the negative image group.

Table 5-9 presents descriptive statistics of participants PANAS negative affect scores before and after completing the IGT. As with the positive scores, the positive image group (group 3) displayed the highest score before the task, however the negative group showed a higher score after the IGT was completed, though none of the scores were drastically different.

Table 5-9. Average negative PANAS scores before and after the IGT

Group	Valence	Time	Score (Std Err)
1 (N=31)	Negative	Before	13.3 (0.54)
1 (N=31)	Negative	After	14.3 (1.01)
2 (N=30)	Negative	Before	13.6 (0.50)
2 (N=30)	Negative	After	14.1 (0.63)
3 (N=30)	Negative	Before	14.1 (0.91)
3 (N=30)	Negative	After	13.9 (0.84)

A bar chart of the negative score (Figure 5-10, top) shows a slight difference in negative scores between groups before the IGT and a similar (albeit flipped) difference with the negative image group having the highest average score after the IGT. A bar chart of the average difference between scores before and after the IGT (Figure 5-10, bottom) for participants shows an increase in score for the negative and neutral groups after the IGT and a slight decrease for the positive group.

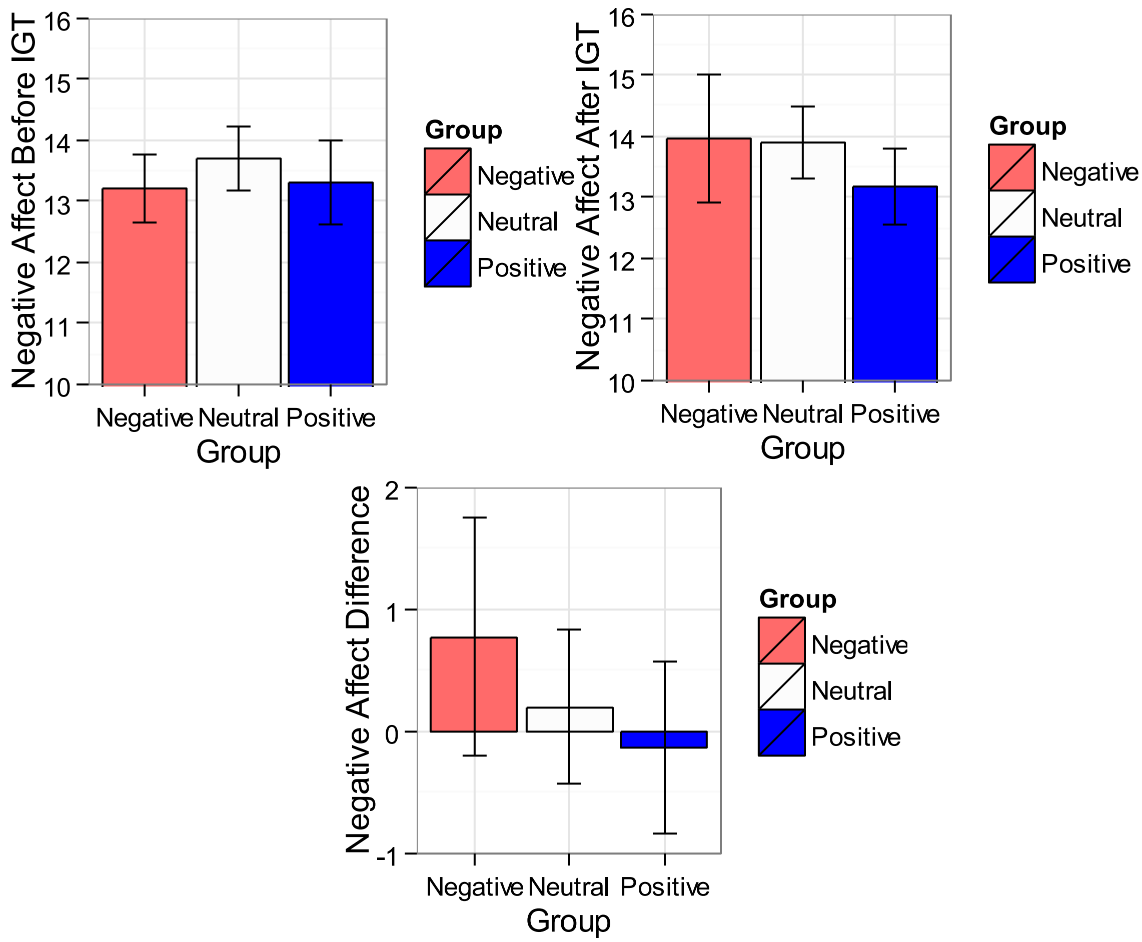

Figure 5-10. Mean PANAS negative affect scores (± SEM) before (top-left) and after (top-right) the IGT, and the average difference between those scores for each participant (bottom)

As with positive affect scores, a one-way ANOVA of negative scores failed to reveal group as a significant factor before the task ($F(2,88) = 0.20$, $p = .8$) or after the task ($F(2,88) = 0.32$, $p = .7$). An ANOVA of the average difference of scores for participants also failed to reveal group as a significant factor ($F(2,88) = 0.85$, $p = .4$). Individual paired t-tests comparing the negative scores of participants before and after the IGT were run for each group. A two-tailed paired t-test did not reveal a significant difference between the scores for the negative image group ($t(30) = -0.79, p = .4$), the neutral group ($t(29) =$ -0.32, $p = .8$), or the positive group ($t(29) = 0.19$, $p = .9$).

ANPS

After completing the IGT and the PANAS questionnaire, participants completed the Affective Neuroscience Personality Scales (ANPS) questionnaire (Davis & Panksepp, 2011). These scales give personality measures based on the primitive-affect framework proposed by Panksepp and Biven (2012). ANPS provides a measure for the SEEKING, FEAR, RAGE, PLAY, CARE, and GRIEF dimensions of the framework. This section only considers scores from the SEEKING, FEAR, and PLAY dimensions as they are likely the most related to the task.

On average, all groups had similar score on the ANPS along the SEEKING, FEAR, and PLAY dimensions (Table 5-10). A one way ANOVA did not reveal group to be a significant factor for SEEKING ($F(2,89) = 0.14$, $p = .87$), FEAR ($F(2,89) = 0.19$, $p = .82$), or PLAY ($F(2,89) = 1.28$, $p = .28$) scores.

Table 5-10. Average SEEKING, FEAR, and PLAY scores

Group	Dimension	Score (Std Err)
Negative (N=31)	SEEKING	29.5 (1.0)
Negative (N=31)	FEAR	24.0 (1.2)
Negative (N = 31)	PLAY	31.2 (1.1)
Neutral (N=31)	SEEKING	29.0 (1.1)
Neutral (N=31)	FEAR	25.0 (1.0)
Neutral (N=31)	PLAY	30.5 (1.1)
Positive (N=30)	SEEKING	28.8 (0.7)
Positive (N=30)	FEAR	24.5 (1.0)
Positive (N=30)	PLAY	28.8 (1.1)

Correlations between positive affect change, cumulative score, and ANPS ratings were computed to find any relationships between personality scores, affective change, and task performance (Tables 13-21). Correlations of performance and questionnaires among participants in the negative group show a marginally significant positive relation between score and PLAY rating, as well as a significant relation between change in positive affect (PANAS Positive) and SEEKING disposition (Table 5-11).

Table 5-11. Correlations between overall IGT score and questionnaire scores for participants in the negative image group. (** = p < .01, * = p ≤.05, #= p < .1)

All	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.08					
PANAS Negative	.15	-.16				
ANPS SEEKING	.21	.38*	-.07			
ANPS FEAR	.15	-.08	.12	-.25		
ANPS PLAY	.34#	.12	.16	.61**	-.28	

Table 5-12. Correlations between overall IGT score and questionnaire scores for participants in the neutral image group. (** = p < .01, * = p ≤.05, #= p < .1)

All	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.23					
PANAS Negative	.08	-.19				
ANPS SEEKING	.16	.17	.10			
ANPS FEAR	.03	.37	.07	.1		
ANPS PLAY	.06	.04	0	.68**	.01	

Table 5-13. Correlations between overall IGT score and questionnaire scores for participants in the positive image group. (** = p < .01, * = p ≤.05, #= p < .1)

All	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.08					
PANAS Negative	-.17	.15				
ANPS SEEKING	.01	.52**	-.09			
ANPS FEAR	.17	.02	.20	-.16		
ANPS PLAY	.12	.22	-.08	.53**	-.02	

Figure 5-11 presents the significant correlations shown in Tables 13-15. The figure shows that few significant correlations were found for all participants.

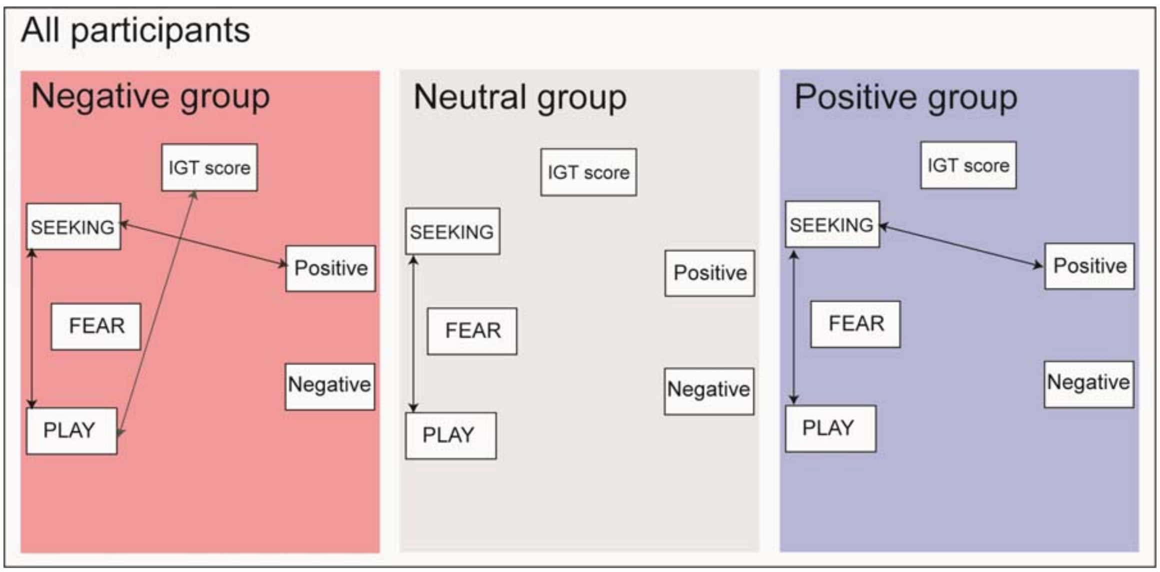

Figure 5-11. A representation of the significant correlations between several variables values recorded during the task for all participants.

Each participant was categorized as being *high* or *low* in an ANPS dimension depending on if they had an ANPS score above or below the mean in their group. Table 5-14 presents the marginally to highly significant results (i.e., $0 \leq p \leq .1$) from running mixed factor ANOVAs with score, A-Rate, response time, SCR, and PANAS score differences ($Score_{After} - Score_{Before}$) as responses and an ANPS dimension as one of the predictors.

Table 5-14. Significant ANOVA results for notable response variables with ANPS scale categories as a factor

Response-variable	Factor/Interaction	$F(nf,df) = F$	$p - value$
Score	Block:SEEKING	$F(4,344) = 2.29$	$p = .06$
Score	PLAY	$F(1,86) = 5.75$	$p < .05$
A-Rate	Group:SEEKING	$F(2,86) = 3.45$	$p < .05$
A-Rate	FEAR	$F(1,86) = 3.87$	$p = .05$
A-Rate$_{\text{Bad Deck to Good Deck}}$	FEAR	$F(1,86) = 10.03$	$p < .005$
A-Rate$_{\text{Good Deck to Bad Deck}}$	FEAR	$F(1,86) = 7.98$	$p < .01$
A-Rate$_{\text{Within Good or Bad Decks}}$	FEAR	$F(1,86) = 8.99$	$p < .005$
Response Time	Group:SEEKING	$F(2,86) = 3.45$	$p < .5$
Response Time	Group:Block:SEEKING	$F(8,344) = 2.89$	$p < .005$
Response Time	Group:Block:PLAY	$F(8,344) = 2.11$	$p < .05$
Response Time$_{\text{Bad deck}}$	Group:SEEKING	$F(2,86) = 4.73$	$p = .01$
Response Time$_{\text{Bad deck}}$	Group:Block:SEEKING	$F(8,331) = 3.13$	$p < .005$
SCR$_{\text{Anticpatory}}$	Group:FEAR	$F(2,47) = 2.81$	$p = .07$
SCR$_{\text{Anticpatory}}$	Group:PLAY	$F(2,47) = 3.22$	$p < .05$
SCR$_{\text{Anticpatory, Bad}}$	Group:FEAR	$F(2,47) = 3.41$	$p < .05$
SCR$_{\text{Anticpatory, Bad}}$	Group:PLAY	$F(2,47) = 2.61$	$p = .08$
SCR$_{\text{Anticpatory, Good}}$	Group:PLAY	$F(2,47) = 3.39$	$p < .05$
SCR$_{\text{R/L}}$	Group:FEAR	$F(2,47) = 2.77$	$p = .07$
SCR$_{\text{R/L}}$	Group:PLAY	$F(2,47) = 3.10$	$p = .05$
SCR$_{\text{R/L, Bad}}$	Group:FEAR	$F(2,47) = 2.88$	$p = .07$
SCR$_{\text{R/L, Bad}}$	Group:PLAY	$F(2,47) = 2.54$	$p = .09$
SCR$_{\text{R/L, Bad}}$	Block:PLAY	$F(4,177) = 2.39$	$p = .05$
SCR$_{\text{R/L, Good}}$	Group:PLAY	$F(2,47) = 3.18$	$p = .05$
PANAS Positive$_{\text{Change}}$	SEEKING	$F(1,85) = 12.09$	$p < .01$
PANAS Positive$_{\text{Change}}$	Group*FEAR	$F(2,85) = 3.10$	$p = .05$
PANAS Positive$_{\text{Change}}$	Group*PLAY	$F(2,85) = 12.79$	$p < .0001$
PANAS Negative$_{\text{Change}}$	FEAR	$F(1,85) = 4.49$	$p < .05$
PANAS Negative$_{\text{Change}}$	Group*PLAY	$F(2,85) = 2.75$	$p = .07$

The first row of Table 5-14 suggests that participant score across blocks may have a different trend depending on whether the participant fell into the *high* or *low* category in the ANPS SEEKING dimension. Figure 5-12 shows that participants with a high SEEKING rating seemed to select more cards from advantageous decks earlier in the IGT and continued to select from these decks more often as the task progressed. The SEEKING effect is especially evident in the neutral and negative groups as participants in these groups with a low score in the SEEKING dimension showing a tendency to select from the bad decks more often even into the 4th block.

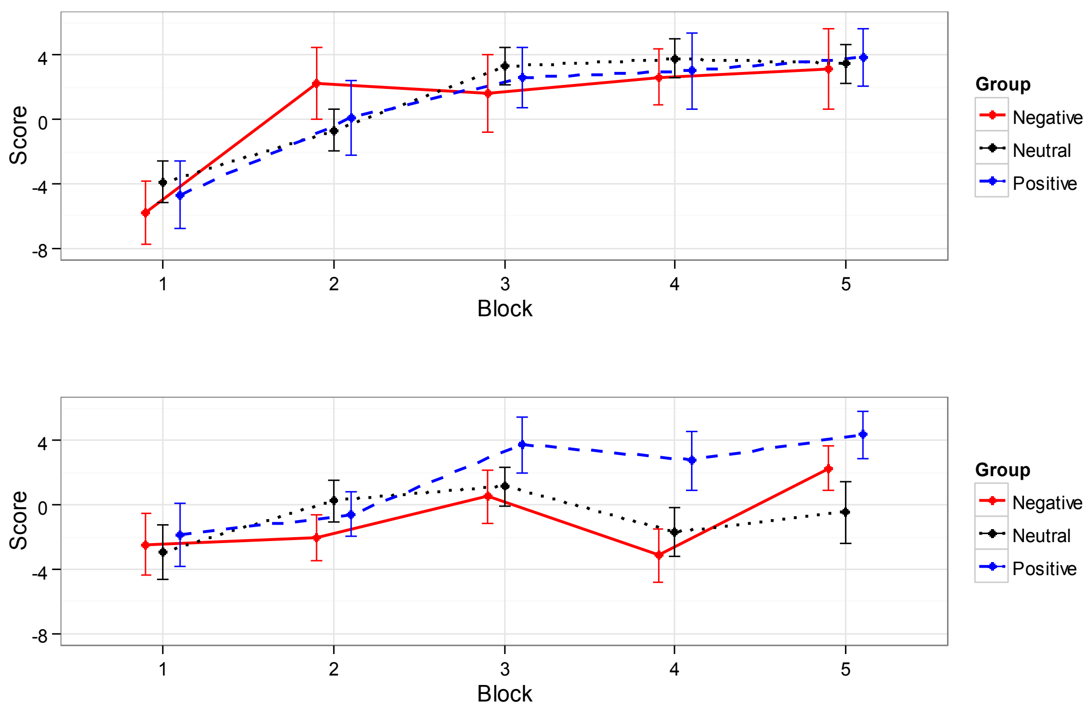

Figure 5-12. Score across blocks for participants who had a high SEEKING rating (top) and a low SEEKING rating (bottom)

Table 5-14 also indicated that categorization in the FEAR dimension of the ANPS scale may influence A-Rate (rate of switching between decks). Figure 5-13 shows that participants with a high score in the FEAR dimension showed a trend of less switching from good decks to bad decks, especially as the task progressed.

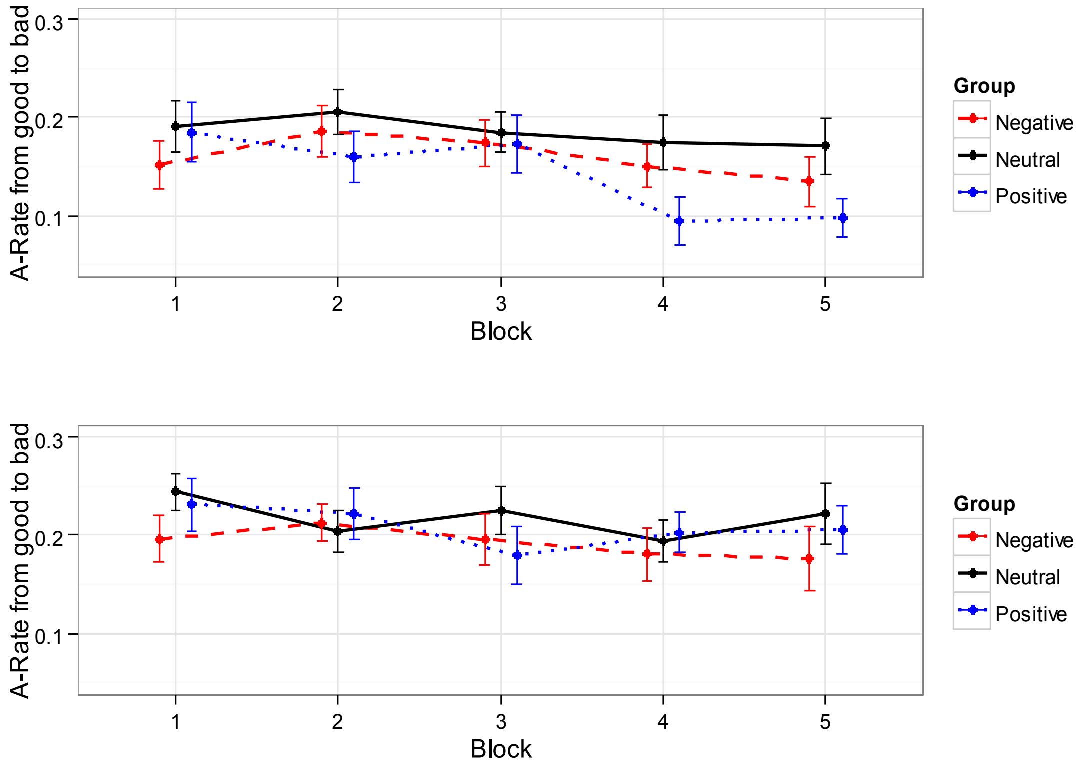

Figure 5-13. A-Rate from good decks to bad decks for participants with a high FEAR classification (top) and with a low FEAR classification (bottom)

Results also indicate that participant SEEKING score may have played a role in response times. In trials where participants selected a bad deck, participants with a high SEEKING rating in the negative group had the lowest response times among all the groups for blocks 1-4 (Figure 5-14, top). However, participants who had a low SEEKING score displayed a nearly opposite pattern with the negative group having the highest response times for blocks 1-3 (Figure 5-14, bottom).

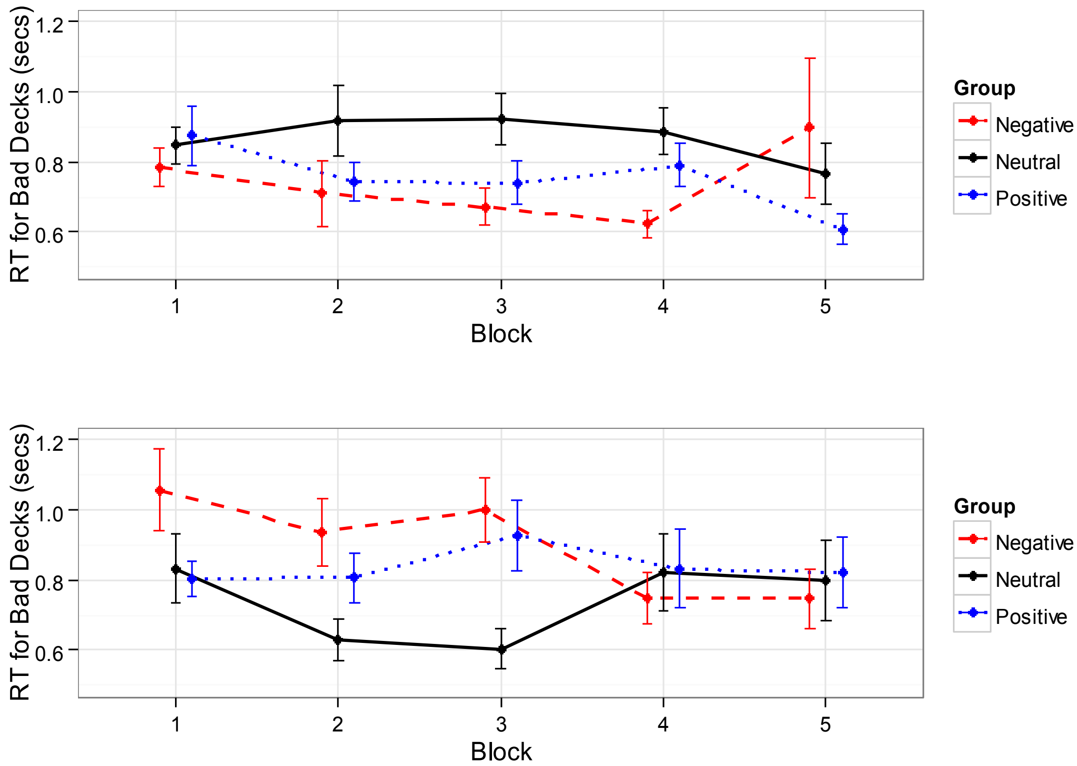

Figure 5-14. Response time for selecting bad decks for participants with a high SEEKING score (top) and participants with a low SEEKING score (bottom)

The section of Table 5-14 that presents ANOVA results with PANAS-measured affective change as a response variable indicates that the intuition that where a participant falls on the FEAR ANPS scale may be important for how their affective state changes as a result of completing the IGT given that the FEAR system is theorized to mediate defensive/aversive-based feelings and behavior. Figure 5-15 shows participants who were categorized as being *high* on the FEAR scale had a higher negative affect score as compared to participants in the *low* category. Furthermore, post-IGT affective state, on average, seems to shift towards a less negative state as subliminal image valence becomes more positive (as indicated by the groups).

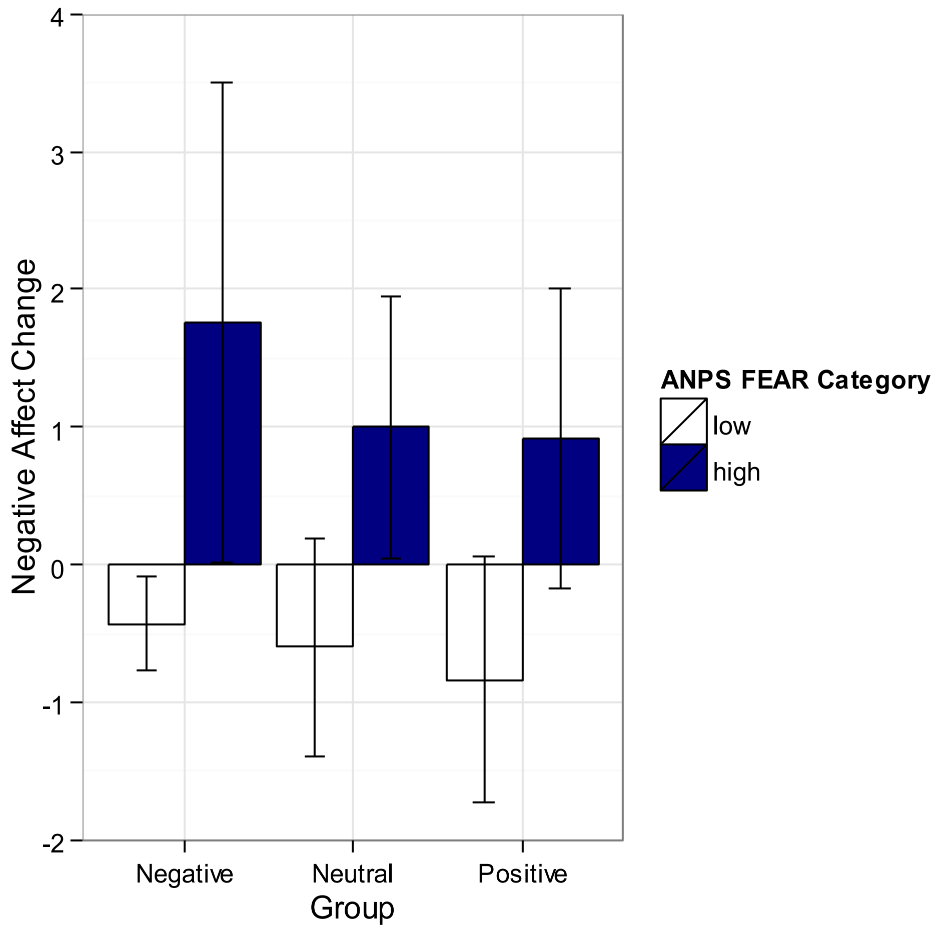

Figure 5-15. Negative affect change after completing the IGT

As suggested by Table 5-14, personality rating in the SEEKING dimensions also appeared to have an effect on participant positive affective change with those in the high SEEKING category, on average, reporting a more positive affective state after the IGT. Similar to Figure 5-15, group also appeared to affect positive affective change with those in the positive image group reporting a higher change in both SEEKING categories and the only positive change among all of the SEEKING-group combinations.

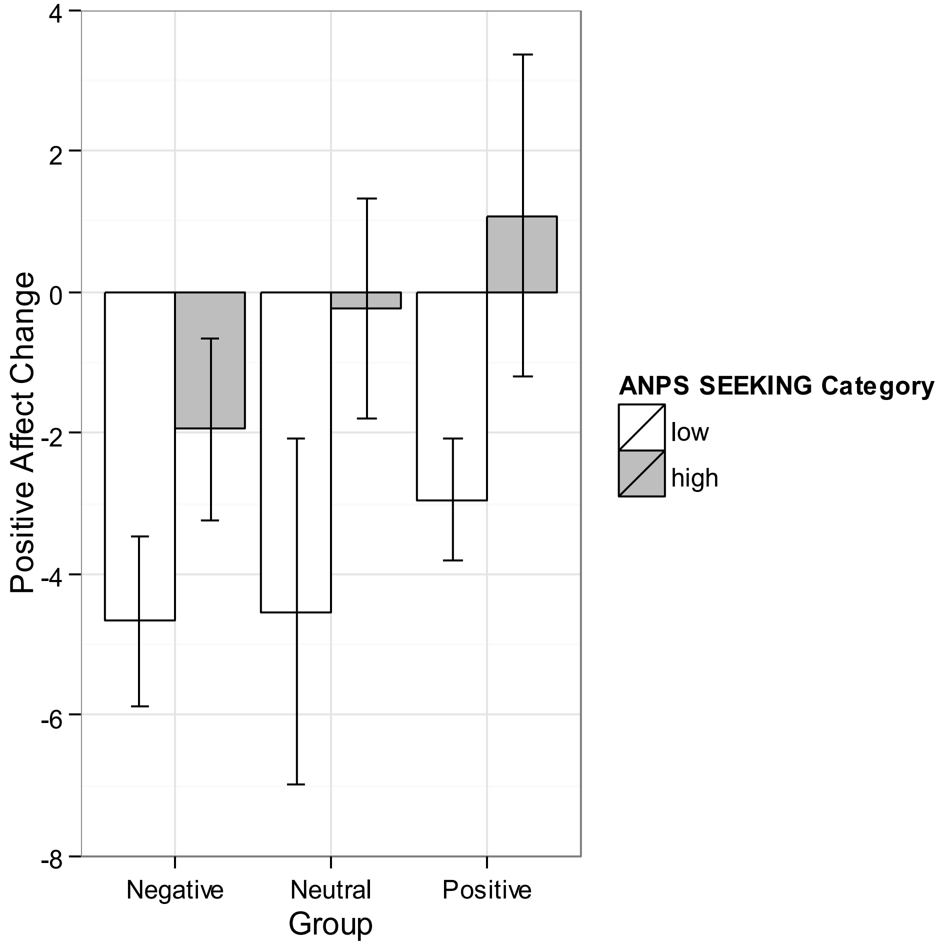

Figure 5-16. Positive affect change after completing the IGT

Post-hoc analysis with data subsets

Recently, it has been reported that there may be a difference between males and females in the IGT (Van den Bos et al., 2013). In turn, the data were analyzed *post-hoc* to test for any differences among male and female participants. Task-related and questionnaire data from 53 males (19 in the *negative*, 16 in the *neutral*, and 18 in the *positive* groups) and 40 females (12 in the *negative*, 15 in the *neutral*, and 13 in the *positive* groups) were analyzed.

Male and female deck selection

As an initial test to see if sex did factor into overall performance on the IGT, an ANOVA for cumulative score with group and sex as factors revealed a marginally significant group:sex interaction ($F(2,88) = 2.920$, $p = .06$). If the data are split into subsets based on sex the order of average cumulative score of the groups flips (Figure 5-17). The cumulative score among males was highest in the positive image group while the average cumulative score among females was highest in the negative image group. The negative group in males and positive group in females have negative cumulative scores.

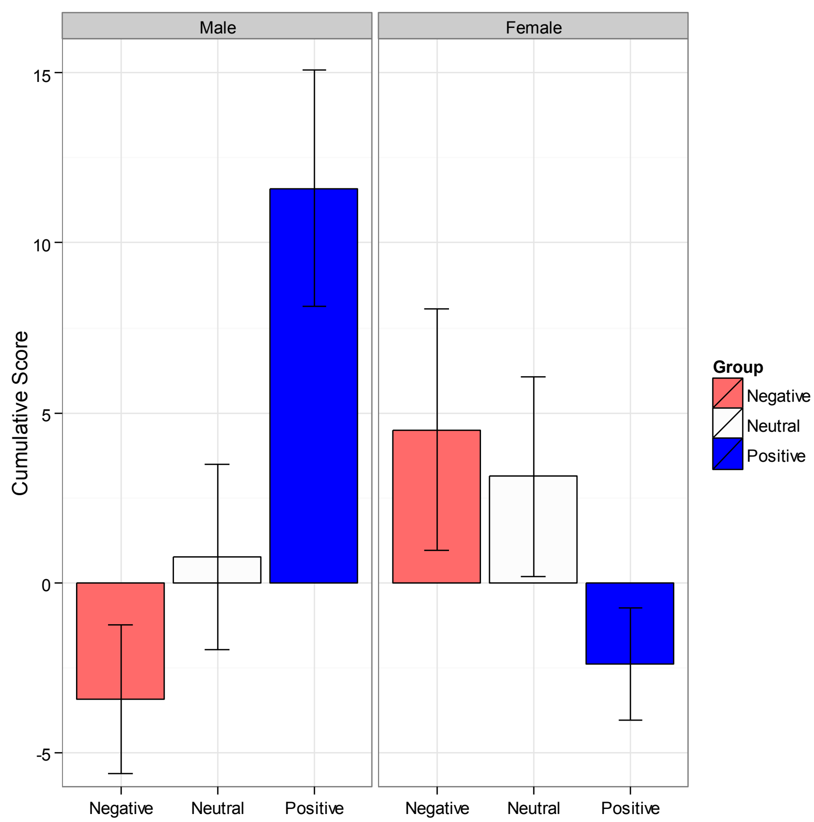

Figure 5-17. Cumulative score for male (left) and female (right) participants after the final block.

The positive image group appears to have the same general pattern between sexes (albeit with different score magnitudes) with score increasing between blocks 1-3 and 4-5 (**Error! Reference source not found.**). Groups 1 and 2 appear to be different with males in the negative image group failing to consistently decide advantageously by the final block. Males in the neutral group show a consistent upward trend in score across blocks.

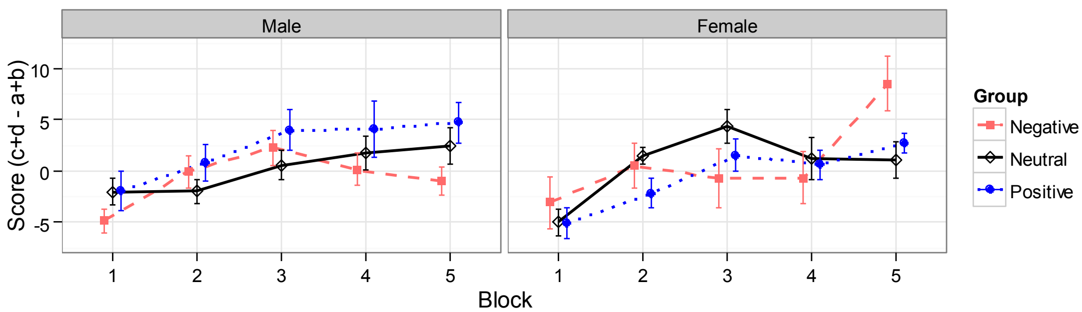

Figure 5-18. IGT score for male (top) and female (bottom) participants

Running a mixed factor ANOVA for score that included participant sex as a potential factor revealed a marginally significant group:sex interaction ($F(2,\ 87)\ =\ 2.77, p\ =\ .07$) and a significant group:block:sex interaction ($F(8,\ 348)\ =\ 2.12,\ p\ <\ .05$). A separate ANOVA of the male and female subsets of data revealed a marginally significant group effect for males ($F(2, 50)\ =\ 2.81,\ p\ =\ .07$) and a marginally significant group:block interaction for females ($F(8, 148)\ =\ 1.93, p\ =\ .06$).

Table 5-15. ANOVA results for score for all, male, and female participants. Bold values signify significant factors

Response Var.	Group		Block		Group:Block	
Score	$F(2,90)\ =\ 0.81$	$p = .5$	$F(4,360)\ =\ 13.22$	$\boldsymbol{<.0001}$	$F(8,360)\ =\ 0.40$	$p\ =\ .9$
$\text{Score}_{\text{Male}}$	$F(2,50)\ =\ 2.81$	$\boldsymbol{p = .07}$	$F(4,200)\ =\ 5.91$	$\boldsymbol{<.0005}$	$F(8,200)\ =\ 0.72$	$p\ =\ .7$
$\text{Score}_{\text{Female}}$	$F(2,37)\ =\ 0.58$	$p = .6$	$F(4,148)\ =\ 8.87$	$\boldsymbol{<.0001}$	$F(8,148)\ =\ 1.93$	$\boldsymbol{p = .06}$

Male and female A-rate

Given that score was found to show different patterns across blocks between sexes, one may expect the deck switching behavior to differ between sexes. As with participant score, males and females did show different deck alternating rates (A-rates) across blocks.

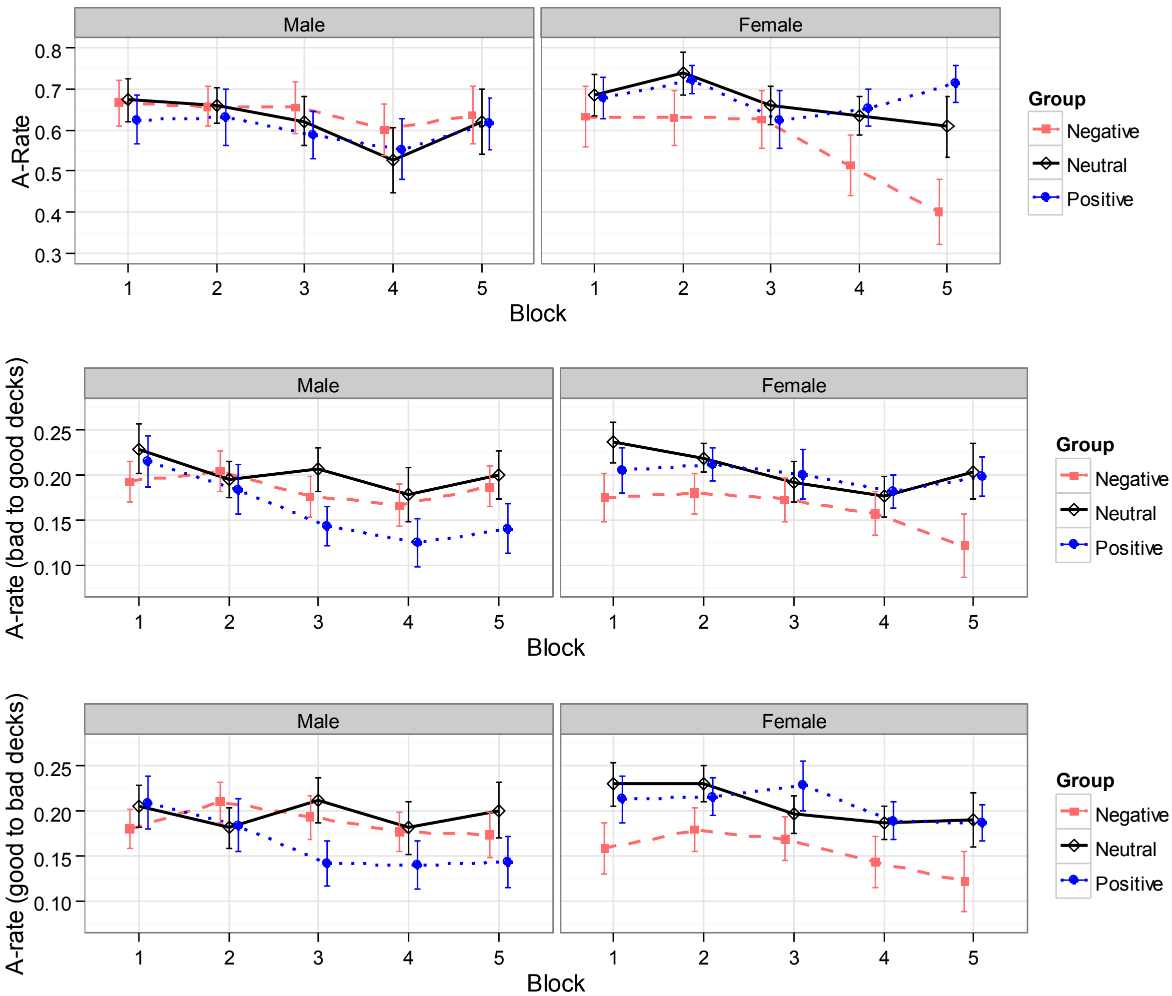

Figure 5-19. The proportion of cards selected from a deck different from the previous trial (top);the proportion of cards selected from a bad deck in one trial and then from a good deck the next (middle);the proportion of cards that were selected from a good deck in one trial and then from a bad deck in the next (bottom). The error bars represent standard error

Male participants in all groups show a decrease in overall A-rate from blocks 1 to 4 followed by an increase from blocks 4 to 5 (Figure 5-19, top left). Female participants in the negative group had a decreasing overall A-rate across all blocks, in the neutral group had an overall decreasing overall A-rate from blocks 2 to 5, and in the positive image group only had a decreasing overall A-rate from blocks 2-3 (Figure 5-19, top right). Male participants in the neutral group tended to have less consistent A-rates for bad decks to good decks (Figure 5-19, middle left) and for good decks to bad decks (Figure 5-19, bottom left). A-rate for bad decks to good decks in female participants was lower across all blocks in the negative image group than the neutral and positive groups.

A $3X5$ mixed-factor ANOVA of overall A-rate for male participants (Table 5-4) revealed a significant block effect ($F(4, 200) = 2.42, p < .05$). An ANOVA of A-rate from bad decks to good decks for male participants revealed a highly significant block effect ($F(4, 200) = 3.94,\ p < .005$). An ANOVA of A-rate from good decks to bad decks for male participants did not reveal any significant factors.

A $3X5$ mixed-factor ANOVA of overall A-rate for female participants revealed a significant block effect ($F(4, 148) = 3.22, p < .05$). An ANOVA of A-rate from bad decks to good decks for female participants did not reveal any significant factors. An ANOVA of A-rate from good decks to bad decks for female participants revealed a marginally significant block effect ($F(4, 148) = 2.41, p = .05$) and a marginally significant group effect ($F(2,37) = 3.07, p = .06$).

Table 5-16. ANOVA results for A-rate for male and female participants. Bold values signify significant factors

Response-Variable	Group	$\boldsymbol{p_{Group}}$	Block	$\boldsymbol{p_{Block}}$	Group:Block	$\boldsymbol{p_{Group:Block}}$
A-rate$_{\text{Male}}$	$F(2,50) = 0.16$	$p = .9$	$F(4,200) = 2.42$	$< \mathbf{.05}$	$F(8,200) = 0.22$	$p = 1$
A-Rate$_{\text{Male, Bad to Good}}$	$F(2,50) = 1.11$	$p = .3$	$F(4,200) = 3.94$	$< \mathbf{.005}$	$F(8,200) = 0.84$	$p = .6$
A-Rate$_{\text{Male, Good to Bad}}$	$F(2,50) = 0.72$	$p = .5$	$F(4,200) = 1.56$	$p = .2$	$F(8,200) = 1.26$	$p = .3$
A-Rate$_{\text{Male, Within Good or Bad Decks}}$	$F(2,50) = 1.03$	$p = .4$	$F(4,200) = 3.18$	$< \mathbf{.05}$	$F(8,200) = 1.08$	$p = .4$
A-rate$_{\text{Female}}$	$F(2,37) = 2.19$	$p = .1$	$F(4,148) = 3.22$	$< \mathbf{.05}$	$F(8,148) = 1.60$	$p = .1$
A-Rate$_{\text{Female, Bad to Good}}$	$F(2,37) = 2.13$	$p = .1$	$F(4,148) = 1.84$	$p = .1$	$F(8,148) = 0.57$	$p = .8$
A-Rate$_{\text{Female, Good to Bad}}$	$F(2,37) = 3.07$	$\boldsymbol{p = .06}$	$F(4,148) = 2.41$	$\boldsymbol{p = .05}$	$F(8,148) = 0.35$	$p = .9$
A-Rate$_{\text{Female, Within Good or Bad Decks}}$	$F(2,37) = 2.66$	$\boldsymbol{p = .08}$	$F(4,148) = 2.16$	$\boldsymbol{p = .08}$	$F(8,148) = 0.42$	$p = .9$

Male and female response times

To see if the difference in deck selection behavior between sexes was accompanied by a difference in the time it took to make a chocie, response time data were examined with sex as a factor and split into male and female groups. A $3X5$ ANOVA for response time revealed a significant group:block:sex interaction for response time for all decks ($F(8{,}348) = 2.04$, $p < .05$) and a significant block:sex interaction for response time for good decks ($F(4{,}346) = 2.74$, $p < .05$).

The response time of males for all decks was highest in the negative group in block 1 and also lowest in the negative group in block 5 (Figure 5-20, top left). This same trend continued for response times for selections of bad decks and good decks (Figure 5-20, middle and bottom left) with male participants in the positive group having the same response time in block 5 for selections of bad decks.

Among female participants, the average response time for all decks for the negative group was the opposite of males with the female negative group having the lowest response time in block 1 and the highest response time in block 5 (Figure 5-20, top right); group 2 had the highest and lowest response times in blocks 1 and 5 (respectively). Response time for bad decks for the negative group was lowest in the block 1 and highest in block 5, but response time for selecting good decks was lowest in the negative group for all blocks, but block 4 (Figure 5-20, bottom right).

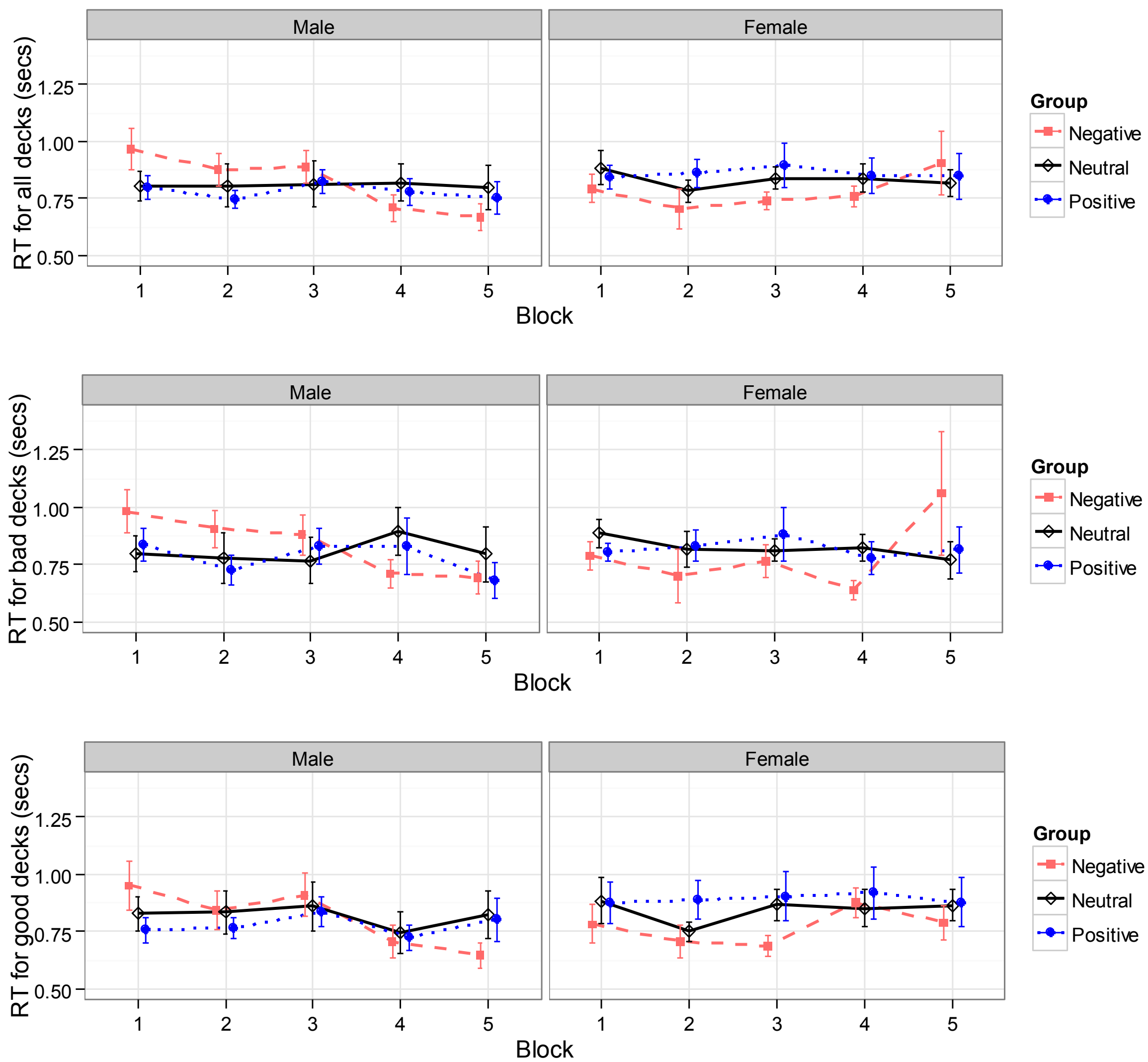

Figure 5-20: Average response time ($\pm SEM$) of males and females across blocks for all decks (top); average response time across blocks for bad decks (middle); average response time across blocks for good decks (bottom)

An ANOVA of response times for male participants (Table 5-6) for all decks revealed a highly significant block effect ($F(4, 200) = 4.66, p < 0.005$) and a significant group/block interaction effect ($F(8, 204) = 2.08, p < 0.05$); An ANOVA of response time for bad decks revealed a significant block effect ($F(4,190) = 2.89, p < .05$) and an ANOVA of response time for good decks revealed a significant block effect ($F(4,199) = 3.69, p < .01$).

An ANOVA of response time for female participants for all decks did not reveal any siginificant factors.

Table 5-17. ANOVA results for response time of male and female participants. Bold values signify significant factors

Response-Variable	Group	p_{Group}	Block	p_{Block}	Group:Block	$p_{Group:Block}$
$RT_{\text{Males, All decks}}$	$F(2,50) = 0.04$	$p = 1$	$F(4,200) = 4.66$	$\boldsymbol{p < .005}$	$F(8,200) = 2.08$	$\boldsymbol{p < .05}$
$RT_{\text{Males, Bad decks}}$	$F(2,50) = 0.20$	$p = .8$	$F(4,190) = 2.89$	$\boldsymbol{p < .05}$	$F(8,190) = 1.62$	$p = .1$
$RT_{\text{Males, Good decks}}$	$F(2,50) = 0.03$	$p = 1$	$F(4,199) = 3.69$	$\boldsymbol{p < .01}$	$F(8,199) = 1.15$	$p = .3$
$RT_{\text{Females, All decks}}$	$F(2,37) = 0.81$	$p = .5$	$F(4,148) = 0.59$	$p = .7$	$F(8,148) = 0.72$	$p = .7$
$RT_{\text{Females, Bad decks}}$	$F(2,37) = 0.82$	$p = .4$	$F(4,145) = 0.93$	$p = .4$	$F(8,145) = 1.39$	$p = .2$
$RT_{\text{Females, Good decks}}$	$F(2,37) = 0.89$	$p = .4$	$F(4,147) = 0.59$	$p = .7$	$F(8,147) = 0.42$	$p = .9$

Male and female electrodermal activity

The group:block:sex interactions shown in previous sections suggest that the visual stimuli may also have affected EDA activity differently depending on sex. EDA data that are divided by participant sex are analyzed in this section. For these EDA analyses there were 34 male (11 in the negative group, 11 in the neutral group, and 12 in the positive group) participants and 20 female (7 in the negative group, 6 in the neutral group, and 7 in the positive group) participants.

Figure 5-21 shows the mean anticipatory SCR for males (left) and females (right) for each group. Males and females displayed differences between groups with males in the negative group having a noticeably higher anticipatory SCR after selecting from a bad deck and females in the neutral condition having a similar high anticipatory SCR. The neutral and positive groups for the male, and the negative and positive groups from the female, display similar anticipatory SCRs when selecting from either a good deck or bad deck.

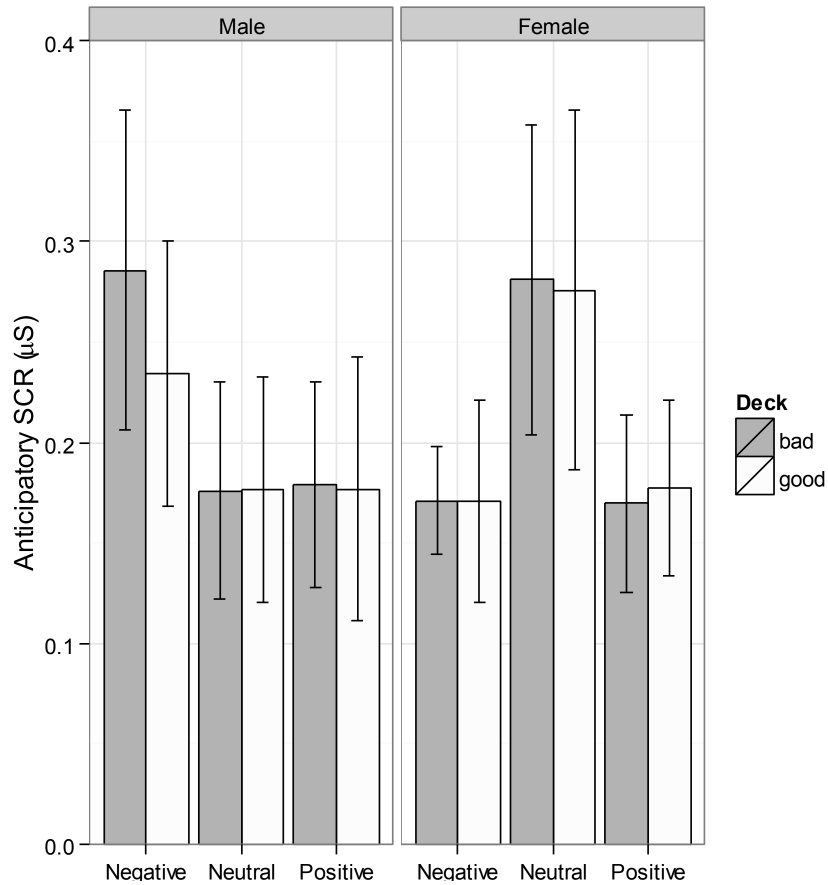

Figure 5-21. The mean (±SEM) anticipatory SCR (µS/sec) of males and females group over the entire IGT.

A $3X1$ ANOVA of anticipatory SCR for participants when selecting from a bad deck failed to reveal a significant group:sex interaction ($F(2,48) = 0.85$, $p = .4$). A $3X1$ ANOVA of anticipatory SCR for participants when selecting from a good deck also failed to reveal a significant group:sex interaction ($F(2,48) = 0.64$, $p = .5$).

Similar to responses found during the anticipatory phase, males in the negative group and females in the neutral group had the highest response after selecting from a bad deck and receiving a loss during the reward/loss phase, as shown in Figure 5-22.

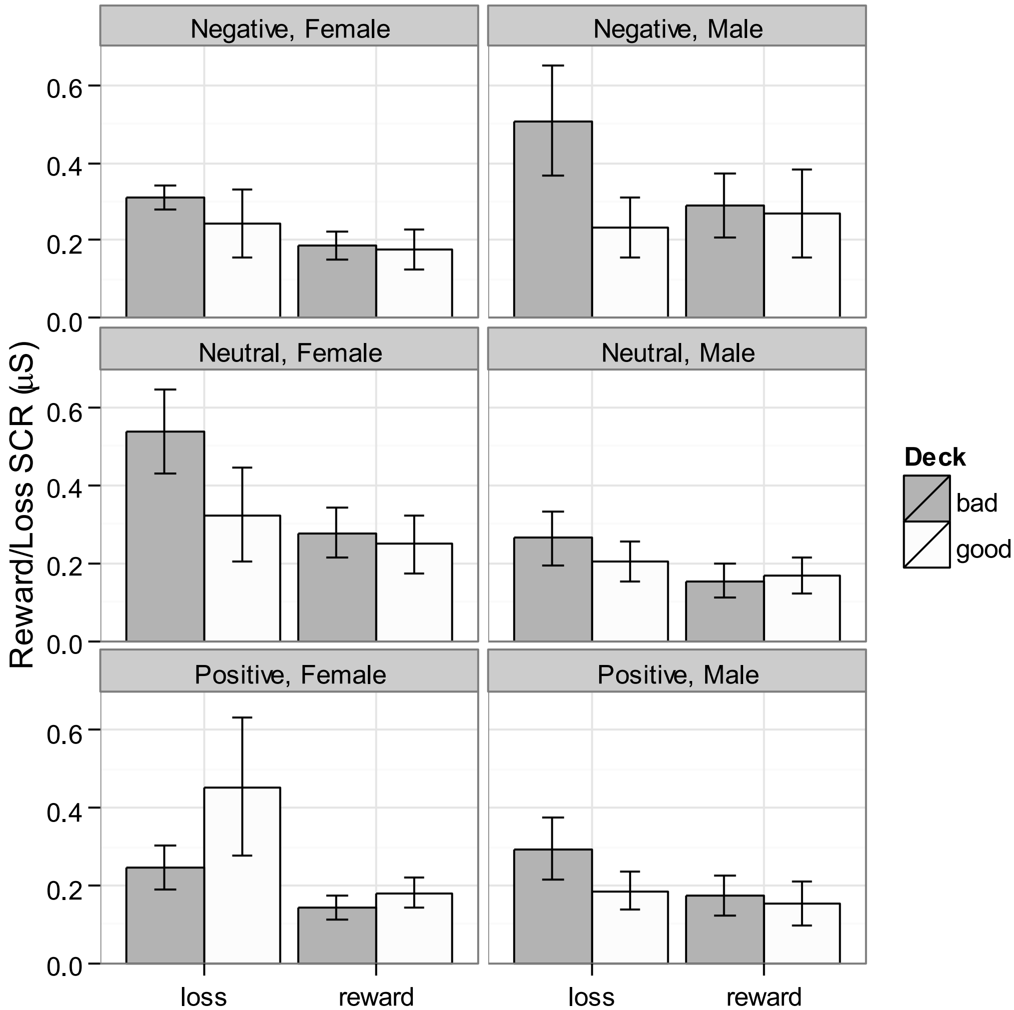

Figure 5-22. The mean (±SEM) reward/loss SCR (µS/sec) in response to trials where outcome was a loss was greater than the reward (loss) and in response to trials where the outcome was a reward was greater than the loss (reward) for each group over the entire IGT.

An ANOVA of SCR response to a loss after selecting from a bad deck revealed a marginally significant group:sex interaction ($F(2,48) = 2.83,\ p = .07$). An ANOVA of SCR response to a reward after selecting from a bad deck did not reveal sex as a significant factor ($F(1,48) = 1.09,\ p = .3$) or a significant group:sex interaction ($F(2,48) = 1.16,\ p = .3$). Sex was also found to be a marginally significant factor with an ANOVA of the SCR response to loss when selecting from a good deck ($F(1,48) = 3.35,\ p = .07$).

Anticipatory SCRs for males in the positive and negative image groups spike from blocks 1-2 and steadily decline in blocks 2-5 (Figure 5-23). This change was higher before selections of one of the bad decks, as male Anticipatory SCRs recorded before the good decks did not show as drastic of a difference. The SEM of male participant SCR tended to be higher in positive and negative groups as compared to neutral groups and SEM was noticeably higher in the negative image groups in blocks 2 and 3.

Anticipatory SCR elicited by female participants seemed to show a different pattern across blocks than the male participants (Figure 5-23). Female participants in the neutral image group, on average showed a spike in SCR during the second block that quickly decreased in the third block to levels more similar to the positive and negative image group. Furthermore, this pattern of SCR change is present before selecting both bad and good decks for the neutral group.

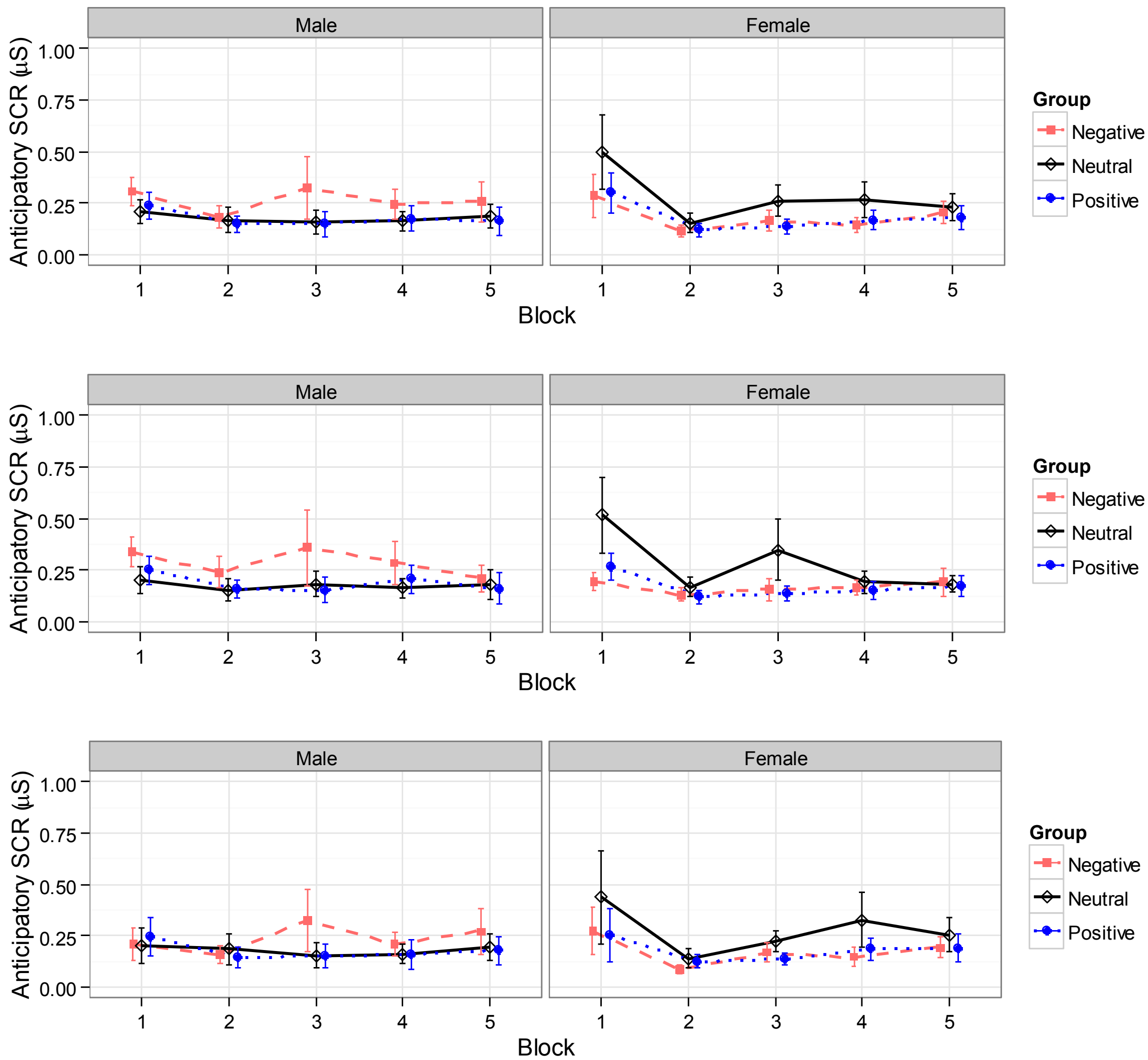

Figure 5-23. Anticipatory SCR results (±*SEM*) for male and female participants for all (top), bad (middle), and good (bottom) decks

Male and female PANAS ratings

PANAS scores were divided by participant sex to see if males and females had similar changes in affect after the IGT. The differences in IGT score and EDA activity between males and females suggests that there may have been differences between males and females in overall change in affective state after the IGT is complete.

Figure 5-24 displays positive affect changes in males and females in the negative, neutral, and positive image groups. Males in the negative image group displayed the greatest reduction in positive affect scores after the IGT. Similar to male participants, female participants in the negative image group also had a higher decrease in positive score after the task than the positive group, however, unlike the males, females in the neutral image group had a slightly greater reduction in score than the negative group after the IGT.

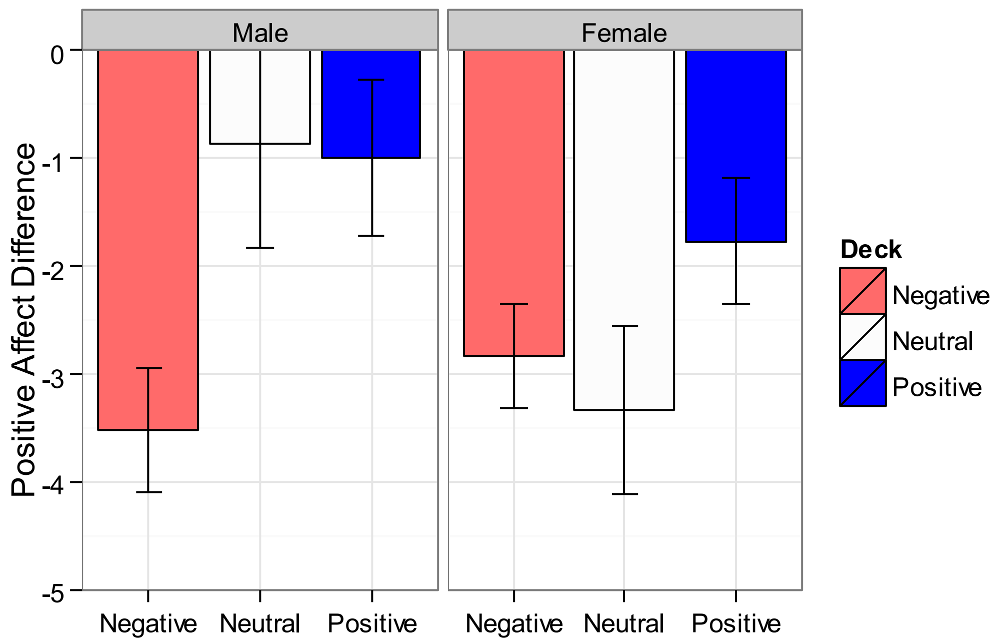

Figure 5-24. Mean PANAS positive affect change (± SEM) in males and females

Figure 5-25 shows negative affect changes in males and females in the negative, neutral, and positive image groups. Males in the negative image group displayed the greatest increase in negative affect scores after the IGT. Females in the negative image group showed a slight decrease in negative affect score while those in the neutral and positive groups showed a slight increase in negative affect.

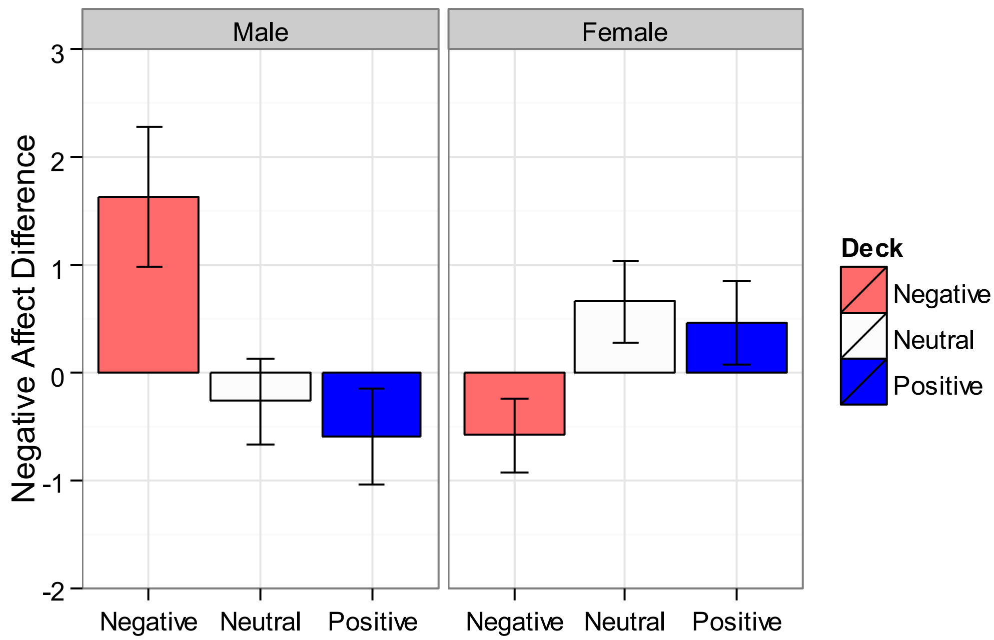

Figure 5-25. Mean PANAS negative affect change (± SEM) in males (left) and females (right)

Though the differences appear to be fairly substantial between groups, group was not found to be a significant factor in males ($F(2,48) = 1.01, p = .4$) or in females ($F(2,37) = 0.58, p = .6$).

To see if there were any significant relations between response variables recorded during the study (within the male and females subsets of data), correlations between several response variables were found for each group within both male and female. Tables 17-22 give the results from running correlations on the subsets of data.

Male participants in the negative image group showed a marginally significant positive correlation between cumulative score and both the change in negative affect and their rating in the FEAR personality dimension (Table 5-18); male participants in this group also showed a significant positive correlation between change in positive affect and their rating in the SEEKING personality dimension. Male participants in the neutral group displayed a highly significant positive correlation between change in positive affect and their rating in the FEAR personality dimension (Table 5-19). Finally, male participants in the positive image group showed a significant positive correlation between change in positive affect and their rating in the SEEKING personality dimension; male participants in this group also showed a marginally significant correlation between change in negative affect and their rating in the PLAY personality dimension (Table 5-20).

Table 5-18. Correlations between overall IGT score and questionnaire scores for male participants in the negative image group. (** = p < .01, * = p ≤.05, #= p < .1)

Male	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.24					
PANAS Negative	.41#	-.19				
ANPS SEEKING	.14	.49*	-.24			
ANPS FEAR	.43#	.10	.36	.01		
ANPS PLAY	.11	.09	.24	.69**	.26	

Table 5-19. Correlations between overall IGT score and questionnaire scores for male participants in the neutral image group. (** = p < .01, * = p ≤.05, #= p < .1)

Male	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.34					
PANAS Negative	.22	.01				
ANPS SEEKING	-.13	-.05	.27			
ANPS FEAR	.37	.72**	.23	-.01		
ANPS PLAY	-.07	-.26	.09	.64**	-.22	

Table 5-20. Correlations between overall IGT score and questionnaire scores for male participants in the positive image group. (** = p < .01, * = p ≤.05, #= p < .1)

Male	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.21					
PANAS Negative	-.17	.08				
ANPS SEEKING	.01	.53*	-.01			
ANPS FEAR	.16	.02	.26	-.33		
ANPS PLAY	.23	-.04	-.43#	.62**	-.36	

Female participants in the negative image group showed a highly significant positive correlation between change in negative affect and their rating in the SEEKING personality dimension (Table 5-21). Female participants in the neutral image group exhibited a marginally significant positive correlation between cumulative score and their rating in the rating in the SEEKING personality dimension; female participants in this group also exhibited a significant positive correlation between change in positive affect and their rating in the ANPS SEEKING personality dimension and the group exhibited a high significant positive correlation between change in positive affect and their rating in PLAY personality dimension (Table 5-22). Finally, female participants in the positive image condition showed a marginally significant positive correlation between cumulative score and change in positive affect; these participants also displayed marginally significant positive correlations between the change in positive affect, and their ratings in the SEEKING and PLAY personality dimensions.

Table 5-21. Correlations between overall IGT score and questionnaire scores for female participants in the negative image group. (** = p < .01, * = p ≤.05, #= p < .1)

Female	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	-.22					
PANAS Negative	-.25	.08				
ANPS SEEKING	.20	.11	.78**			
ANPS FEAR	-.06	-.37	-.37	-.56#		
ANPS PLAY	.48	.14	.28	.47	-.72**	

Table 5-22. Correlations between overall IGT score and questionnaire scores for female participants in the neutral image group. (** = p < .01, * = p ≤.05, #= p < .1)

Female	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.12					
PANAS Negative	-.07	-.40				
ANPS SEEKING	.48#	.60*	-.18			
ANPS FEAR	-.30	-.02	-.12	.21		
ANPS PLAY	.2	.65**	-.27	.69**	.21	

Table 5-23. Correlations between overall IGT score and questionnaire scores for female participants in the positive image group. (** = p < .01, * = p ≤.05, #= p < .1)

Female	IGT Score	PANAS Positive	PANAS Negative	ANPS SEEKING	ANPS FEAR	ANPS PLAY
PANAS Positive	.53#					
PANAS Negative	.06	.34				
ANPS SEEKING	-.41	.50#	-.28			
ANPS FEAR	.22	.00	.17	.02		
ANPS PLAY	-.21	.55#	.35	.49#	.17	

Tables 18-23 give results from running the same correlation analysis with the data divided by sex. Figure 5-26 provides a pictorial representation of the correlation analysis for each sex. The figure shows that more significant correlations were found when the data were divided by sex.

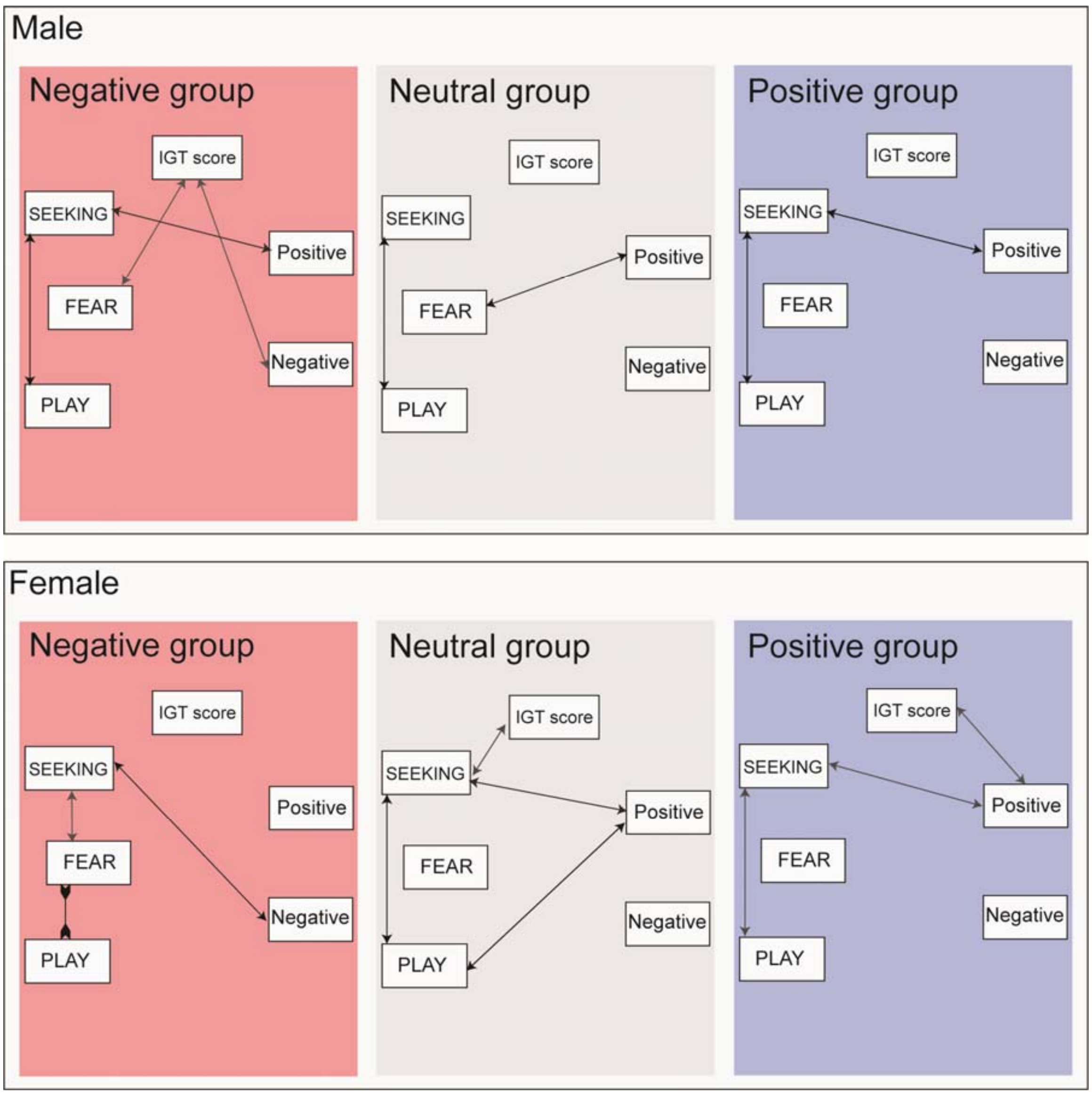

Figure 5-26. A representation of the significant correlations between several variables values recorded during the task for male participants (top) and female participants (bottom). The reversed arrowhead (female negative group) represents a significant negative correlation.

Summary

Participants provided electrodermal activity data and behavioral data, both of which were recorded while they completed a modified version of the Iowa Gambling Task that added the presentation of subliminal visual stimuli after the selection of cards from *bad* decks. The visual stimuli were taken from the International Affective Picture System and were separated into negative, neutral, and positive categories to be presented to participants assigned to the negative, neutral, or positive groups (respectively). Participants completed the Positive and Negative Affect Scales (PANAS) before and after the task to get a measurement of affective state that may be related to their assigned group (and thus the group of images subliminally presented). Participants also completed the Affective Neuroscience Personality Scales to see if any personality differences were related to significant differences in EDA, task-related, or PANAS behavior.

Participant data did not indicate any significant difference between negative, neutral, or positive groups for score across the blocks of the IGT. Participants' average cumulative score (total score at the end of the task) was highest in the positive image group, followed by the neutral and negative groups.

There did appear to be sex differences in task-score as in males cumulative scores followed the same group ordering as the male and female combined data, but female participants had the opposite ordering between groups with the negative image group having the highest cumulative score followed by the neutral and positive groups. Van den Bos et al. (2013) have discussed results that indicate sex-related behavioral differences can occur in the IGT; these data thus extend those results by indicating that this sex-related factor in decision-making may also interact with stimuli that are not consciously attended. Group also generally was not a significant factor for A-Rate (rate of switching between decks) or response times.

EDA skin conductance response (SCR) showed a decreasing trend across blocks for all groups indicating a habituation to the task as it progressed. The average response to a loss when selecting from a bad deck was also much higher in the negative group than the neutral and positive group, providing an early indication that subliminal images affect some non-conscious processes, despite not having any reported awareness of the images. There also appeared to be sex related differences in the response to subliminal affective images when presented with a loss and a general sex related difference between the SCR responses of males and females to any loss. In addition to their responses to losses in bad decks, participants' responses to losses after selecting from good decks differed depending on sex. There also appears to be a sex-related difference in SCR across time, but more participants from each sex likely should be run through the experiment for any definitive conclusions regarding this issue.

Questionnaire data obtained using the PANAS and ANPS also had interesting results. There were group-related trends in changes in PANAS scores with the negative groups showing the highest decrease positive-related PANAS rating after the task (significant) and the highest increase in negative related PANAS rating after the task (not significant); the positive image group also showed a lower decrease in positive-related PANAS scores after the task and slight decrease in negative-related PANAS scores after the task.

Scores in the SEEKING category of the ANPS were related with changes in positive affect (PANAS) for the positive and negative image groups. When categorized as either *high* or *low* in ANPS personality dimensions (in particular the SEEKING, and FEAR dimensions) participants in different groups generally displayed different trends of behavior. Those who were *high* in SEEKING generally began deciding advantageously more quickly and did not decide disadvantageously in any blocks after first registering an average score above 0. Those who were in the negative and neutral images groups and *low* in SEEKING, however, took longer to begin to decide advantageously and were selecting more cards from the bad decks than the good into the

4th block. Interestingly those in the positive image group who were low in SEEKING displayed a similar general pattern to those high in SEEKING.

Participants categorized as high in the FEAR dimension generally displayed a lower switching rate from good to bad decks across blocks than those low in FEAR. The lower switch rate indicates that reaction to explicit negative events and stimuli (as measured by the FEAR dimension of the ANPS) was greater in some way in participants with a higher FEAR disposition.

In summary, the results indicate that the treatment used in the study (the sets of subliminal images) did work in that they caused noticeable and significant effects on different aspects of participant physiological, affective, and decision behavior. Even more interesting, this treatment unexpectedly interacted with participant sex, indicating that not only did the non-conscious images affect processing, but the images may have affected processing in different ways in males and females. Personality characteristics also affected the decision process resulting in different decision-making performance on the task.

Chapter 6 Comparisons of experimental results and model

The ACT-R model discussed in Chapter 4 completed a computerized version of the Iowa Gambling Task. This is the same Matlab-based IGT as participants that were recorded for the data that were presented in Chapter 4. The model uses (simulated) visual and motor mechanisms to interact with the software environment. Before comparing the model to any data, it is important to understand what criteria drove it's development, that is, how it was judged as it was been developed and improved over time.

Judging a model

When examining a computational model, there are several criteria one can use to judge whether it is a *good* of a model. There is simple question that I assume most would ask and seek to answer before figuring out any equation or functions that would go into a computational model: does the model already exist? This determines how much ground-work must be done before developing the model and how much work will have to be put into the model for it to be considered useful. It turn, once the model has been developed, the usefulness of the model is much greater if another one does not exist.

Another important criteria for a computational model is whether or not it can complete the task. It is a simple criteria, but one that anyone who wishes to write software that implements a computational model will likely appreciate. Also important is the model's ability to communicate with a task that is very similar to whatever task one may eventually be using to gather experimental data. The more similar the tasks are, the higher the fidelity of the model, and the easier it is to adjust the model to any changes in an experimental task due to research questions (i.e., the model becomes more useful as it can be expanded to do more versions of the same task later in a more straightforward manner).

Table 6-1. Some criteria for determining how *good* a computational model is.

Criteria
Does a similar model exist?
Does the model complete some version of the task?
Can the model complete the same task (including software) that was used to collect experimental data?
How long does it take for the model to run?
What level of behavior does the model describe?
Does the model simulate processes involved or is it solely focused on output?
How well do model predictions match experimental data? And what data are matched to the experiment
If there have been previous versions of the experiment, can the model give you predictions that have not been found in previous versions of the experiment, but that can be recorded in the future?

A criteria commonly used to judge a computational model is how well it predicts and matches data. This is an especially useful criteria, because it can allow a reader to understand how well *good* your model is without necessarily understanding all of the processes involved in the model; of course, model that matches data very closely can still not be generalizable to similar tasks.

In the next section, I review some comparisons between the model presented in Chapter 4 and the experimental data presented in Chapter 5. Score is the most common metric that is used to examine behavior during the IGT, so I compare the models predicted score and general deck selection behavior to the data reported in this dissertation.

Model predictions and experimental data

Using the same task, the model produced task performance that was comparable to experimental participants. For all groups, the model predicts a similar scoring trend from blocks 1 to 3 (Figure 6-1); the $IGT_{negative}$ model and negative image groups show the biggest difference in score predicted and actual score output out of the three groups. The 4th block is a critical block for model and participant data comparison as, during this block, the model continues to get a higher score on the task (similar to previous blocks), but experimental data shows that participants tended to select more disadvantageously in block 4 than block 3.

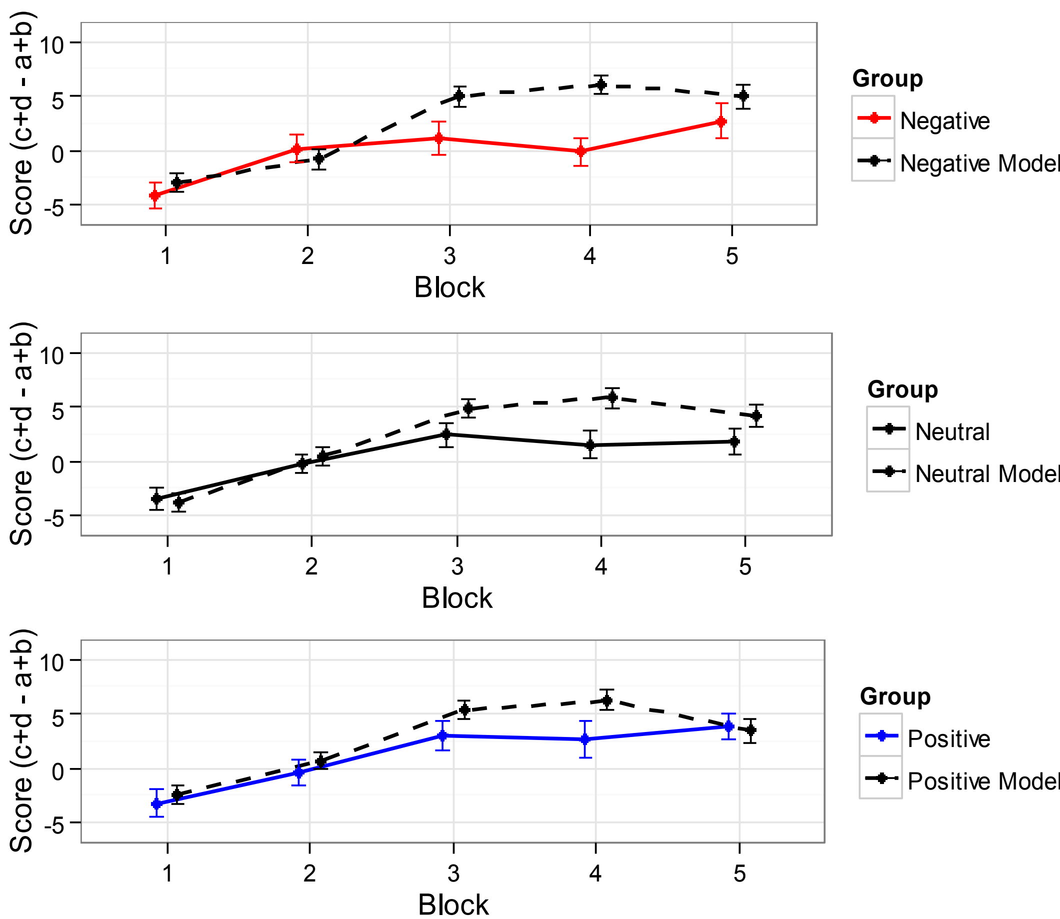

Figure 6-1. Score performance comparison between participants and mode for negative (top), neutral (middle), and positive (bottom) groups.

Table 6-2 presents results from comparing model score predictions across five blocks to actual participant scores. The $IGT_{Neutral}$ model predicted scores for that particular group the best ($r = .97$), followed by the $IGT_{Positive}$ model ($r = .90$) and the $IGT_{Negative}$ model ($r = .75$).

Table 6-2. Comparison measures between score performance predicted by the model and actual score performance exhibited by participants

Model/Group	r	$RMSD$
$IGT_{Negative}$/Negative	.75	3.48
$IGT_{Neutral}$/Neutral	.97	2.49
$IGT_{positive}$/Positive	.90	2.05
All	0.85	2.74

Figure 6-2 shows comparisons between the A-rate predicted by the negative, neutral, and positive versions of the IGT model and the data presented in Chapter 5. The negative and neutral IGT models showed a similar A-rate trend as the task was completed. These two models also both exhibit trends that are similar to the experimental data found in the negative and neutral groups. The positive IGT model showed an A-rate trend from blocks 1-4 that was similar to the experimental data, but failed to predict the increase in A-rate in the final block of the task.

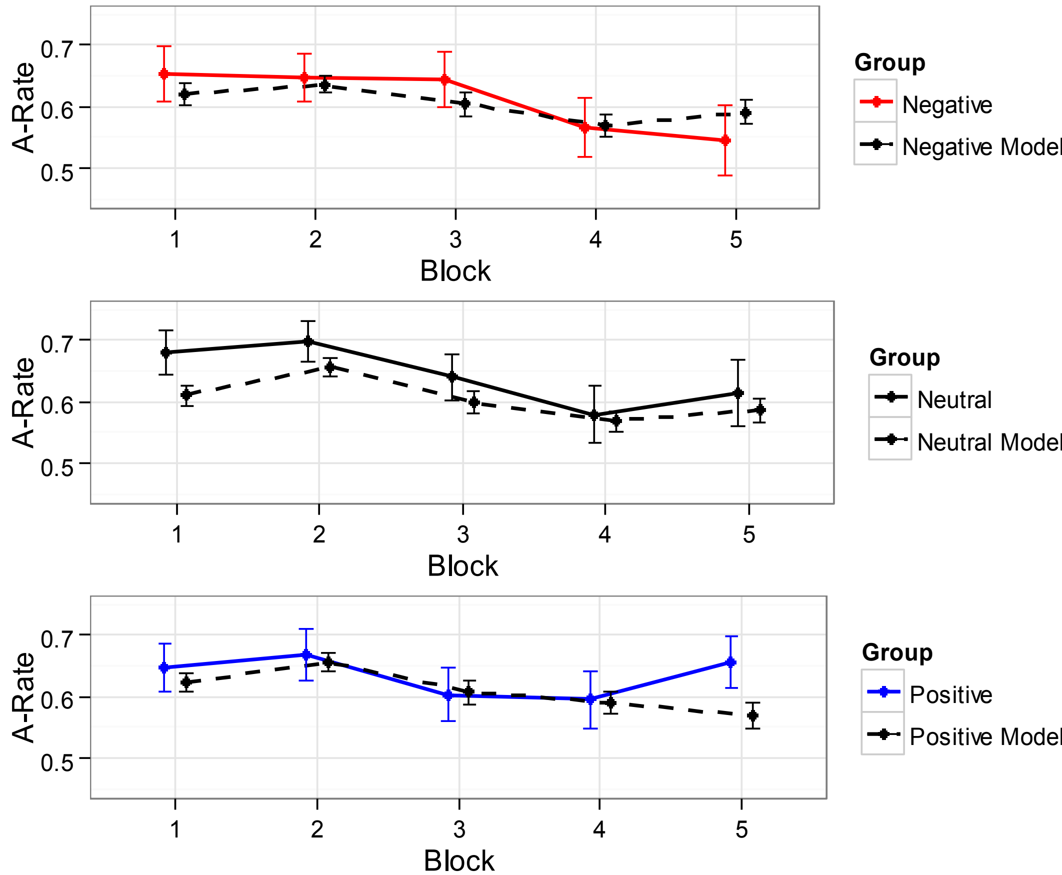

Figure 6-2. A-Rate comparison between model and participants for negative (top), neutral (middle), and positive (bottom) groups.

The IGT model best predicts the A-Rate for the neutral ($r = .93$) and negative ($r = .8$) participant groups, however there was a very low correlation between the models predicted A-Rate and participants in the positive image group. The low correlation is likely due to the large deviation in the 5th block in the positive group where A-Rate is predicted to decrease from block 4 in the model, but participants actually show an increase (Figure 6-2) indicating participants in the positive image group repeated deck selections less frequently in the final block.

Table 6-3. Comparison measures between A-Rate that is predicted by the model and actual A-Rate exhibited by participants

Model/Group	r_{A-Rate}	$RMSD_{A-Rate}$
$IGT_{Negative}$/Negative	.8	0.03
$IGT_{Neutral}$/Neutral	.93	0.04
$IGT_{positive}$/Positive	.39	0.04
All	.68	0.04

Individual deck analysis of the model and participant data indicates that the cause for the disparity between the predicted and actual data differs depending on group (Table 6-4). On average, the model predicts less of the selection data for decks A and D for the negative and positive groups, while it predicts less of the selection data for decks A and B for the neutral group.

Table 6-4. Comparison measures between number of cards selected from individual decks that are predicted by the model and actual number of deck selections made by participants

Model/Group	r_a	$RMSD_a$	r_b	$RMSD_b$	r_c	$RMSD_c$	r_d	$RMSD_d$
$IGT_{Negative}$/Negative	.29	0.69	.77	1.71	.84	0.86	.44	1.33
$IGT_{Neutral}$/Neutral	.47	1.06	.55	1.95	.91	1.07	.95	0.62
$IGT_{positive}$/Positive	.60	1.20	.93	1.61	.77	0.98	.30	1.15
All	.43	1.01	.71	1.76	.80	0.97	.51	1.07

There was a considerable disparity between predicted model response times and participant response times with participants having a higher average response time across all of the blocks than the model prediction (Figure 6-3).

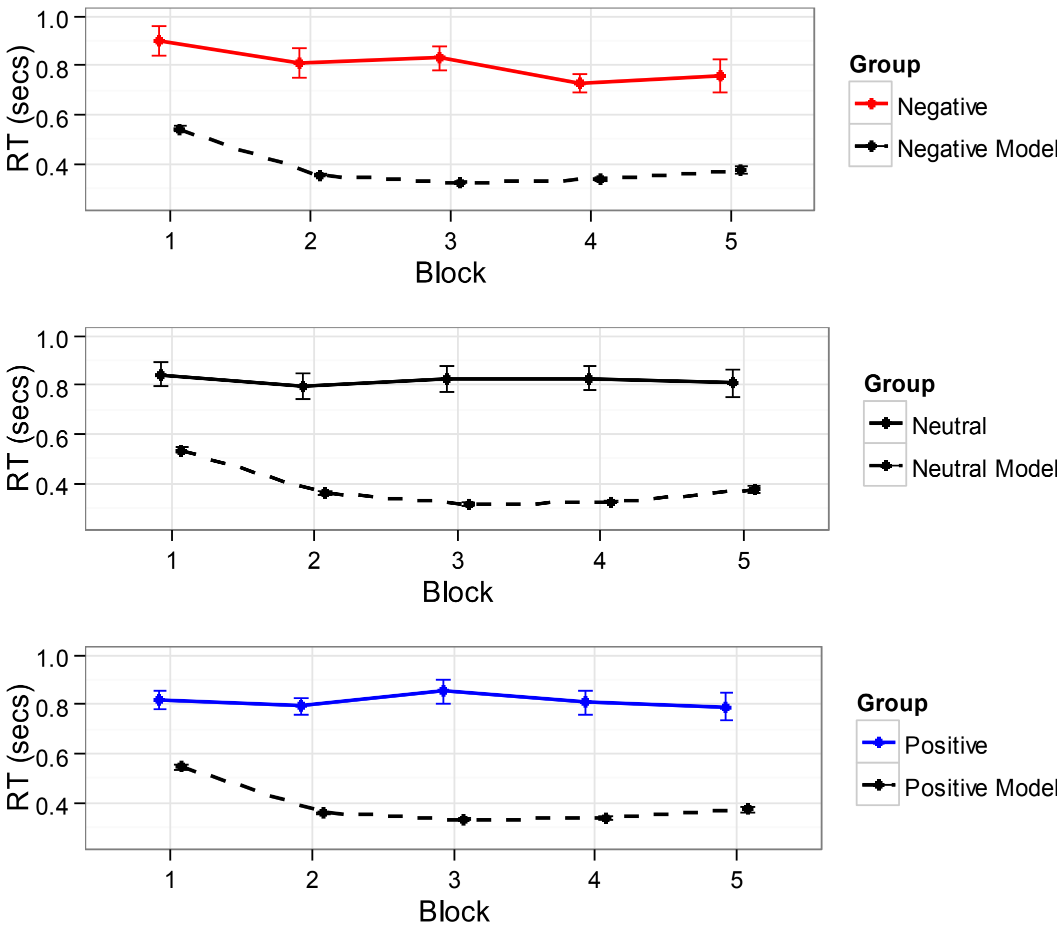

Figure 6-3. Response time comparison between model and participants for negative (top), neutral (middle), and positive (bottom) groups.

As expected from Figure 6-3, the model did not predict average participant response time well with the Negative model/group comparison having the only r above .74 (Table 6-5). Because the data show very little change in response time across blocks, RMSD is likely the only reliable measure for comparison between those data and the neutral and positive models.

Table 6-5. Comparison measures between response time that is predicted by the model and actual response time exhibited by participants

Model/Group	$r_{Response\ Time}$	$RMSD_{Response\ Time}$
$IGT_{Negative}$/Negative	.74	0.42
$IGT_{Neutral}$/Neutral	.49	0.44
$IGT_{positive}$/Positive	.00	0.43
All	.44	0.43

Thus, the model predicts participant score in the IGT fairly well (the measurement typically used and reported across studies), but does not fare as well when it comes to A-Rate or response times. The model predicts some groups' individual deck selections well, but fails to provide a good prediction for other decks selections in the same group across blocks.

The model seems to predict task-related behavior in the IGT for neutral groups the best, potentially indicating two things about this version of the IGT: (1) the mechanisms used by the model to represent the affective visual stimuli likely do not represent the processing exhibited by the actual model; (2) personality and sex differences in visual stimuli processing may mean more participants are needed to get more clear results.

Summary

The model presented in Chapter 4 predicted participant score results fairly well in all groups and predicted results from participants in the neutral group the best. Analysis of individual deck selections predicted by the model revealed prediction deficiencies typically in two of the deck selection counts (e.g., number of card selected from deck A and deck D). The model did not predict participant A-Rate or response time as well as score. A-rate in the neutral group was predicted best by the model. The difference in model predicted response time and actual found response time indicates two things: (a) that the processes used once a participant could select a deck were fairly stable and little learning related to deck-key pairs occurred; (b) another process is likely involved in making the decision is missing for the model (e.g., a last minute appraisal of the deck that is intended to be selected).

The criteria from Table 6-1 indicate that some aspects of the model are useful and worth take seriously, but it still has room for improvement. No other instance of a LISP-based ACT-R model could be found, so a new model had to be developed. What's more, the only other computational process model that could be found does not run in the main LISP version of ACT-R (it runs in ACT-R Python), and did not include any representation of perceptual or declarative memory processes, which this model predicts should factor into the behavior of the task[14]. This model does actually complete the task, and it communicates with the same computerized version of the IGT that participants communicated with during the study.

The model gave useful predictions, but still can be improved in terms of fit to experimental data. This can be accomplished by adding additional strategies and perhaps changing parameter values. The changes to the model should come with good theoretical reasoning as the model can be changed in many ways to *fit the data*.

Because the model uses processes within the ACT-R/Φ, the behavior can also be related to behavior and activity in the brain (e.g., Figure 4-1). This gives one the opportunity to judge the processes involved in behavior in the model based on studies that have associated the same processes with brain activity (e.g., Fukui et al., 2005; X. Li et al., 2010). Thus, the model's use of original ACT-R components is also beneficial, because it can be used to predict and understand systematic brain responses during the task.

---

[14] It seems any model that uses any form of decision-making that includes explicit memory should have a representation for declarative memory processes and how these processes can affect behavior.

Though the model can be improved in several ways, it is a useful starting point for developing more interesting decision-making models in the ACT-R/Φ architecture. It should be judged as a *good* model, or at the least *useful*, depending on the criteria one prefers to use when judging a model. Irrespective of the criteria one uses to judge a model, it is important that as these computational process models are developed, they are reported in some way and in enough detail to be reproducible given the complexity they usually contain. Because variations of the IGT have been used to examine different populations (e.g., Kobayakawa, Koyama, Mimura, & Kawamura, 2008; Shurman, Horan, & Nuechterlein, 2005; Van den Bos, Harteveld, & Stoop, 2009; Verdejo-Garcia et al., 2007), and to examine different interesting questions, this model is especially useful as it can be used as a base-model to be expanded to help improve the understanding of the processes that mediate human behavior in different populations.

Chapter 7 Conclusions

In this chapter, I conclude the dissertation with some general insights gained from completing the work (see Table 7-1 and the subsequent discussion) and a discussion of the work presented in chapters 3-6. I also review some contributions of this work, and some useful ways the work might be used and expanded upon. I conclude the chapter with some limitations of the work and final remarks.

Table 7-1. Some general insights that have arisen from completing work in chapters 3-5

Related Chapter	Insight
Chapter 3	Some behavioral actions, and tendencies for action, should be built into a computational architecture that describes human behavior to account for seemingly evolved behavior (e.g., automatic behavior exhibited when hungry or thirsty).
Chapter 3	Process redundancy (e.g., procedural and declarative memory processes both being used to complete a task) is important to represent in an architecture that constrains and implements human behavior and having representations on the physiological, affective, and cognitive level in a computational architecture makes this redundancy more tractable.
Chapter 3	Though several descriptive models exist for the architectural processes that implement and result-in behavior, few of them are already quantified in a manner suitable for representation in a computational system; quantifying the models will be important if we wish to have a more intricate understanding of the processes that underlie human behavior. The lack of quantitative models is perhaps most apparent in studies of emotion.

Chapter 3	The memory systems that come to dominate overtly affective behavior are not widely described from a computational point of view. In understanding how processes that unfold over-time to mediate human behavior, we should also be considering how these affective memory processes are interacting and adding value to the more commonly studied *declarative* and *procedural* memory processes.
Chapter 3	What is typically described as *noise* in memory processes may be more accurately described and represented as changes in underlying physiology.
Chapter 3	Though mood, emotion, and affect may be described as separate constructs, they likely represent an interaction of several processes that have already been described in the literature. Affect and emotion represent an interaction between lower-level (i.e., mostly subcortical) affective systems and higher-level (i.e., mostly cortical) cognitive systems that result in affective states and sometimes conscious recollection of affective states. Mood, which is typically described as having a longer temporal effect on behavior, can in fact be represented by the similar processes with the addition of global physiological changes (e.g., changes in epinephrine levels) that continuously modulate affective and cognitive states and that change overtime nonlinearly.
Chapter 4	Understanding and modeling how strategies shift over time in making a decision is important when trying to understand what processes may have been used to come to a final choice or decision. Computational process models provide a way to understand these shifts with more detail because they allow explicit quantitative modeling of the processes and how they may interact with different strategies.
Chapter 4	Small changes in reward representation can have huge effects on the process of decision-making even in a task where the decisions are more simple and constrained. Furthermore, the process of deck selection and processing makes a difference because the time in which a decision is made internally may be separated temporally enough for a reward or loss to be significantly attenuated.

Chapter 5	Sex and personality differences (and their interactions) should be reported (or at the least considered more) in studies involving decision-making as there appear to be small sex differences in task behavior.
Chapter 5	When using affective stimuli and recording affect-based data it is important to understand how both sex and personality differences may interact to affect the data. This is especially important if one wishes to generalize from an experiment. In many studies adding a personality measurement may be adding just 10-15 minutes to the study.
Chapter 6	Though useful, standardized subjective ratings may not be the most straightforward to quantify stimuli used for simulations that focus on behavior on physiological and affective levels, that is, below the tertiary level of behavior (an example of behavior on the tertiary level would be the active interpretation of stimuli through memory and self-appraisal mechanisms).

ACT-R/Φ

This dissertation presents a novel hybrid computational cognitive architecture, ACT-R/Φ, which begins to describe how physiology, affect, and cognition interact over-time to implement and constrain human decision -making. ACT-R/Φ combines the ACT-R cognitive architecture with representations of affect (using systems from primary-process affect theory) and physiology (using the HumMod integrative model of human physiology and simulation system). The architecture describes how changes in physiology and affective states modulate memory and action processes, thereby affecting decision-making and choice. It integrates knowledge of how stress systems, affect (or emotion) systems, and homeostatic systems can change cognition into a coherent architectural model that constrains and mediates human behavior.

In moving past descriptive models of how the physiological, affective, and cognitive processes interact the difficulty of developing a computational architecture becomes apparent:

despite our continually improving understanding of the human body and how it implements and is affected by the mind, theories and models of the processes that mediate behavior largely remain strictly descriptive. Even theories that offer an equation of some form to describe behavior typically do so without (explicitly) considering how the formula interacts with other systematic components of behavior.

Therefore, though ACT-R/Φ represents an advancement in unifying the understanding of how physiological, affective, and cognitive processes interact-with and modulate one another, concessions had to be made the implementing the architecture and some connections had to be quantified that will need to be explored in continual development of the architecture. The persistent movement towards more systematic and unified theories of human behavior (for which Newell, 1973, and Newell, 1990, explicitly advocated) should shed more light on the plausible mathematical descriptions of some of the processes in the architecture.

The IGT model

In addition to the existing process models that have already been developed to use the physiological and/or affective components in the ACT-R/Φ architecture, this dissertation describes a process model that completes the Iowa Gambling Task, a popular task used to study the interactions between affective behavior and decision-making. The model completes and communicates with the same software version of the IGT that was used for a related study.

The process model describes how declarative learning, procedural learning, and perceptual motor processes interact over the course of the task. It predicts that in the beginning of the task, choices will be made based on explicit memory of deck-reward/loss pairs (declarative memory retrieval), and over-time, reinforcement-based learning (utility-based procedural memory selection) will begin to dominate the task. The model also predicts that subliminal images may have a small effect on the task-behavior, but will not have a large effect on task-behavior or physiological changes.

This is one of the few process models built to run in ACT-R that use its declarative-learning, procedural learning (reinforcement learning and production compilation), and perceptual/motor components. The model also communicated with the same software as participants in a related experiment. Though running the model through the task took less time as it could be run in *simulation-time* as opposed to real-time, a certain level of programming knowledge was required to implement this type of a model. This required extra knowledge may be a reason fewer complex models are developed that run in the same environment as experiment participants, despite the higher fidelity achieved with such a model.

The IGT study

This dissertation also presented the results from having human participants complete a version of the IGT that included subliminal affective visual stimuli (with the same software that communicates with the IGT process model). Participants were shown either negative images, neutral images, or positive images (as determined by standardized ratings) after selecting cards from a bad deck. Contrary to initial hypotheses, image group was only a significant factor in a few of the task performance measurements and physiological response.

Participant sex, however, was a factor with males and females showing significantly different performance trends. A recent review indicates that men and women tend to differ in decision-making performance in the IGT (Van den Bos et al., 2013). In particular, Van den Bos et al. (2013) have shown that women tend to select more cards from deck B (a bad deck) than men while men tend to select more cards from decks C and D (good decks) than women. The authors posit that women participants may have been more responsive to short-term losses than the long-term winning strategy (i.e., they alternate deck selection after losses from bad *and* good decks). Indeed, the experimental data presented in this dissertation indicate that females displayed a different A-rate (switching-rate) trend across blocks than males; females in the neutral and positive groups had a higher average A-rate than males in the neutral and positive groups in all blocks.

This sex difference also significantly interacted with the participants group for several of the measurements. The participant sex and group affected how they reacted to losses in the bad decks (even without conscious awareness that an image was flashed before the loss) and their overall decision behavior throughout the task. The differences in the effects of subliminal images on physiological and cognitive behavior could be due to differences in amygdala based processing of the stimuli (e.g., Cahill, 2006; Gillath & Canterberry, 2012; Hamann & Canli, 2004; Hamann, Herman, Nolan, & Wallen, 2004).

Affective state change after the task was affected by assigned group. Differences in affective personality also interacted with groups for some task performance measures. Thus, the data indicate that focused task-related choice-behavior is more likely be affected by affective stimuli that are not consciously attended given certain personality characteristics. The data also indicate that this subliminal stimuli effect on decision-making may be mediated by the sex of the decision-maker. When recorded in the same experiment, the questionnaire and EDA data make it clear that making level of measurement clear is important when making any conclusions about the interactions between affective and cognitive processes; depending on the type of affective data, one may come to different conclusions.

The model predicted task-performance (score) well, but did not predict other aspects of behavior (e.g., response-time and rate of switching between decks) as consistently. The model provided stronger predictions for the neutral group, potentially indicating that the model does not represent all of the processes underlying affective changes correctly. Though the model was initially developed with the intuition that the affective pictures would affect choice in the task, task-behavior on all metrics were very similar and the subliminal stimuli essentially only affected background affective state in any noticeable manner; subsymbolic changes to memory in the model due the affective visual stimuli were essentially overpowered by the task-related reinforcement. Though, the model did not predict many aspects of participant behavior well, results were promising given the complexity of developing such a process-model and the costs of simply getting it to complete the task using many of the relevant learning mechanisms.

Some open research questions

Upon reflection of some of the work completed over the course of this dissertation, some questions have arisen that would be interesting and useful to explore in the future.

Table 7-2. Some questions that one may have after understanding the work in this dissertation

Questions
What is the most useful level to represent and simulate sex differences in a hybrid computational architecture that includes physiological, affective, and cognitive representations?
What are the most useful ways to quantify affective memory processes that integrate with declarative memory and procedural memory processes?
How independent are these affective memory processes?
How does the interaction of physiological, affective, and cognitive processes change social behavior and consequently affect social networks?
How do small changes in the IGT task that are not overtly related to the actual deck and reward/loss affect behavior? Does, for example, using a mouse instead of a keyboard, lead to any interesting effects?

How would changes in the complexity in IGT modulate strategy selection in making deck-selection decisions? Does, say, a more strict reinforcement-learning strategy become more dominant as complexity increases or does a more declarative-based strategy become more dominant? Furthermore, how may the increase in complexity modulate the effect of non-integral stimuli subliminal stimuli on decisions?

In addition to the questions posed above, there likely are several other questions that one can pose from the perspective of this dissertation work. Below, I highlight the contributions of the work described in this dissertation.

Contributions of this work

In line with the interdisciplinary nature of the College of Information Sciences and Technology, this work integrated research from and contributes to multiple disciplines. The new hybrid cognitive architecture, the decision-making model, and the experimental results positively contribute to work in emotion, cognitive science, computational physiology, neuroscience, and decision-making.

A cognitive architecture with physiology and affect

ACT-R/Φ is a novel hybrid computational architecture that extends a widely used cognitive architecture (ACT-R) with representations of physiology and affect. It combines a cognitive architecture, an integrative model of human physiology (HumMod), and a major theory in affect and emotion (primary-process affect theory). This is the first and only cognitive architecture to use an integrative model of human physiology to dynamically change cognitive behavior. The affect representation in ACT-R/Φ provides one of the few accounts of affect/emotion in a cognitive architecture and is the first to connect cognitive, affective, and physiological representations into one system. The architectural model that is implemented in the ACT-R/Φ computational architecture is also the first to suggest that the ACT-R architectural model and primary-process affect model can be combined into a coherent theoretical model based on work in neuroscience, cognitive science, and psychology. The ACT-R/Φ hybrid architecture can be used to model how the interactions between physiology, affect, and cognition may result in differences in human behavior. Because ACT-R/Φ is an extended version of the ACT-R cognitive architecture, ACT-R process models can be used within the ACT-R/Φ architecture with little to no modification; thus, previous ACT-R models can be directly compared to similar models that use the new components of ACT-R/Φ. This architecture is freely available, so that the ACT-R community, or anyone else wishing to develop a cognitive model, can use the architecture to develop interesting models and further extend the architecture.

A computational process model of the Iowa Gambling Task

I've developed a computational process model that completes the Iowa Gambling Task (IGT). This process model communicates with a Matlab version of the IGT. This was the same version of the IGT that was completed by participants in the related study. This process model runs within the ACT-R/Φ architecture and is the first known computational process model that uses perceptual, memory, and affective processes to complete the IGT. The IGT model can be expanded to examine how changes in the IGT (e.g., different reward/loss deck assignments) and changes in physiology (e.g., a rise in cortisol during the task) may change behavior during the task. The model code is currently available by request so that it can be used and extended by other researchers.

A study examining the effects of subliminal images on decision-making

A modified version of the Iowa Gambling Task was run to examine how subliminally presented images may affect decision-making related behavior. This is the first experiment to explore the effects of subliminally presented images on decision-making behavior during the IGT; the images were obtained from the International Affective Picture System (IAPS). This is also the first study to explore associations between ratings on the Affective Neuroscience Personality Scales (ANPS; personality scales based on primary-process affect theory), ratings on the positive and negative affect scales (PANAS), and decision-making behavior with or without subliminal image presentation. Thus, both the experimental treatment (a set of subliminal images), and using affect measurement instruments (PANAS and ANPS) are contributions to the literature.

The software used to conduct the study was developed using the Psychtoolbox library (Brainard, 1997) in Matlab and is freely available. Making this software also represents a contribution as no existing Matlab-based software version is freely available and the scripts used to run the study presented in this dissertation is particularly useful because it can be used to conduct versions of the IGT that require timing accuracy in the milliseconds.

Potential future directions of this research

There are several ways the work presented in this dissertation could be expanded upon, because the dissertation focuses on understanding and simulating the behavior arising from systematic interactions between physiology, affect, and cognition. I discuss three particular useful ways this research can be used and expanded upon below: simulating drug effects on human behavior on multiple levels to better understand how a drug may affect behavior in the short and long term; better understanding of the interaction of physiological, affective, and cognitive processes in addiction and compulsion behaviors; developing more realistic and tractable agents in military simulations.

Simulating drug effects on physiology, affect, and cognition

Given that drugs often have certain physiological, emotional, and cognitive effects, it would be useful to use ACT-R/Φ to simulate and predict the effects of different drugs on human behavior and what these effects mean for human performance. The physiological system in ACT-R/Φ (HumMod) provide a large integrative physiological model that provide several access points and can be modified to replicate these effects. The most intriguing drug to model in ACT-R/Φ is caffeine given that it is the most commonly consumed psychostimulant in the world.

Caffeine is known to cause an increase in ACTH and cortisol levels (Lovallo, Al'absi, Blick, Whitsett, & Wilson, 1996) at varying levels depending on daily consumption (Lovallo et al., 2005). Cortisol is one of the main hormones in the HPA stress axis and can have both slow gene-mediated and fast (e.g., within minutes) non-genomic effects on the brain (Groeneweg, Karst, de Kloet, & Joëls, 2011; Joëls et al., 2011). It affects both peripheral and neural structures as cortisol is released peripherally and passes through the blood-brain barrier, allowing it to affect several physiological systems. Also, cortisol not only modulates the PVN of the HPA-axis to control future release of cortisol, but also modulates hippocampus and BLA activity, thereby affecting emotional declarative memory systems. The effect of cortisol on the PVN may also be important given its efferent connections to the LC that could hold implications for both stress, liking, and wanting related behaviors (e.g., Van Bockstaele, Reyes, & Valentino, 2010).These widespread systematic changes, in combination with the fact that cortisol secretion due to caffeine intake or stress can affect the body over the course of hours, make it clear a systematic perspective is needed to understand how caffeine's effects on cortisol may change human behavior and being able to understand this effect over time is equally as important.

Caffeine also affects neural catecholamines (e.g., norepinephrine and dopamine) through its antagonistic effects on adenosine, which is a catecholamine antagonist. Norepinephrine (via the LC and NTS) acts a global arousal modulator, affecting several neural structures including the BLA and the hippocampus, two structures especially important for affective declarative memory. As discussed earlier, dopamine is a major neurotransmitter implicated in SEEKING or wanting affective behavior and is implicated consequently implicated in reinforcement learning. Thus, caffeine can affect multiple memory systems through effects on global modulatory hormones and catecholamines. What's more, these effects range from the minute to hour time-scale, pointing to the need to understand and predict when different amounts of caffeine at different times, and in combination with different stimuli, may affect cognitive performance and decision-making in a positive or negative way.

Though experimental work has shed a good amount of light on the mechanisms of caffeine's effects on human behavior, computational simulations that use existing evidence from experimental and theoretical work can be used to systematically explore how these drug effects without running into ethical or experimental control concerns. A more recent cognitive model simulation of the effects of caffeine and stress on cognition (Ritter et al., 2009) was useful examination on how one can represent some effects of caffeine in a computational cognitive architecture, but did not have the ability to represent the effects just discussed on multiple levels and time-scales. ACT-R/Φ could be used to model caffeine's effects on cognition as it gives a modeler representations of physiological and affective components (in addition to cognitive) on an architectural level. Furthermore, recent work on modeling cognitive effects on a social level (Zhao et al., 2012) could allow the systematic exploration of how caffeine effects on the physiological, affective, and cognitive levels affect memory and decision-making and how this results in changes in social networks and dynamics overtime.

Understanding the interaction of processes during addiction and compulsion

Several recent theories on addiction and compulsion have been proposed from an emotion and affective neuroscience point of view (e.g., Alcaro & Panksepp, 2011; Robinson & Berridge, 2008; Verdejo-García & Bechara, 2009). These theories often posit an overpowering appetitive urge (SEEKING or wanting) that guides the addiction or compulsion and an involvement in dopaminergic areas. Two especially compelling related theories are those involving SEEKING or wanting/incentive salience (e.g., see Alcaro & Panksepp, 2011; Robinson & Berridge, 2008, for reviews). Incentive salience has also been implicated from an anthropological point of view, giving some validation to the idea that these appetitive urges, and the dissociation between liking the addictive or compulsory stimulus and wanting it (e.g., K. S. Smith et al., 2011). Perceptual cues, environmental contexts, and stress can all cause process changes related to an addiction or compulsion on the physiological, affective, and cognitive levels.

Given the primitive dopaminergic SEEKING system is implicated in being largely responsible for the basic appetitive motivation underlying the strong compulsory urge to SEEK out addictive substances, it is important to understand how the physiological, affective, and cognitive process can unfold overtime. Computational simulations can be used to predict how combinations of environmental stimuli will affect decision-making in those suffering from compulsion and addiction. More specifically, ACT-R/Φ has accounts for affectively enforced memories (including those partially driven by SEEKING or wanting urges) and physiological changes in relevant systems including the HPA and SAM axes, which are especially important in behavior like addiction relapse (Ungless, Argilli, & Bonci, 2010).

It is important to understand the neural structures that mediate cognitive control of addiction-based affect (sometimes referred to as cognitive down-regulation) in addition to understanding the networks involved in addiction-based affect and behavior itself (that involves neural structures including those in the SEEKING system as well as the VMPFC/OFC). Recent evidence suggests that the network primarily mediating cognitive down-regulation of addiction involves the DLPFC, VLPFC, ACC, VMPFC/OFC, and the NAcc and OT (Hayashi, Ko, Strafella, & Dagher, 2013; Kober et al., 2010). In ACT-R/Φ, activity in different buffers and modules is associated with these neural structures/networks:

Goal buffer/module - associated with activity in the DLPFC and ACC

Retrieval buffer - associated with activity in the VLPFC

Affective-associations module – associated with activity in the VMPFC/OFC

SEEKING module – associated with activity in the NAcc and OT.

A process model of psychopathological behavior related to addiction (e.g., a model that uses cognitive strategies to reduce ones addictive urge) would use these modules/buffers to simulate the process of addiction regulation. An ACT-R/Φ-based model of this sort seems fairly intuitive given that the Goal buffer and module maintain the current goal state of the architecture and evidence that DLPFC activity increases with the perceived availability of the substance of addiction (indicating more cognitive control may be needed in this situation) and transmagnetic stimulation (TMS) deactivation of the DLPFC attenuates ACC activity and affects subjective craving related to proximity/availability (perhaps a form of *goal-proximity*) of the particular substance (Hayashi et al., 2013). A process-model of addiction psychopathology that gives representations on the physiological, affective, and cognitive levels would be useful for providing predictions of efficacy of treatments, both cognitive/behavioral treatments and drug-related treatments.

Developing realistic and tractable agents for military simulations

At what point does performance and decision-making begin to degrade while under a variety of stressors, and what are the mechanisms that cause these degradations? Given that the ACT-R/Φ architecture can represent physiology over longer periods of time, it seems a natural fit for developing more realistic behavioral models and agents to operate within military simulations. Now agents and models can dynamically change their behavior according to a realistic internal physiology and get stressed tired, hungry, or thirsty over the course of a mission. These changes in physiological and affective states can have severe effects on cognition. Understanding behavior on multiple levels in a military context is especially useful given the often-occurring interaction between psychostimulants like caffeine and especially stressful situations.

Process models of human cognitive behavior have been developed to run within military-like simulations in the past (e.g., Best & Lebiere, 2006; Evertsz, Ritter, Russell, & Shepherdson, 2007; Ritter et al., 2012). Having an architecture in which to develop process models that predict behavior in different contexts is important for understanding both friendly and adversarial behavior. It can be highly beneficial to understand the points at which adversarial decision-making behavior may change due to physiological or affective changes (e.g., being tired, stressed, and/or hungry). Understanding and identifying what internal and external environments may cause behavioral biases, and identifying what those biases are, can lead to saving lives.

Stress and post-traumatic stress disorder (PTSD) are also important to understand, combat, and recognize in any context, but especially in a military context, where live ammunition and deadly force comes into play. The hybrid architecture can be used to develop models that get stressed, experience fear, and even experience PTSD-like effects on physiological, affective, and cognitive processes. One can then simulate and observe how these processes can affect overall leaning and decision-making. Process models can be developed with the aim of understanding how, for example, monitoring and changing the internal and external environment of military personnel can result in more optimal, potentially life-changing, decisions.

Limitations

Given the depth and breadth of portions of the dissertation work, it follows that some limitations and caveats will apply to the work. There were limitations in the architectural development, model development, and the experiment. Some limitations were expected and some emerged as the work was completed.

Architectural limitations

The ACT-R/Φ architecture combines three models that describe human behavior on different levels: HumMod ─ physiological, Primary-Process Affect ─ affective, ACT-R ─ cognitive. In combining these models into a computational architecture, it follows that equations and data are needed to define interacting mechanisms beyond a purely descriptive model. Two of these models (HumMod and ACT-R) were already in computational form and provided quantifiable (but separate) predictions, and the primary-process affect theory did not directly provide any quantifiable predictions, but other related research did give some computational accounts of mechanisms (e.g., incentive salience theory).

Development of the ACT-R/Φ architecture revealed a gap in the existing data and quantitative models that made it much more difficult to implement specific mechanisms into the architecture. Thus, while I was able to combine some existing results from incentive salience theory, theories and results from studies of stress, and studies of learning and memory, the task proved more difficult and less straightforward than I would have anticipated. I used theories and models that were grounded in, or previously related to, physiological structures (e.g., neurological structures) to narrow down interacting mechanisms. Nonetheless, once the interaction of mechanisms and processes had to be quantified, relatively few were explicitly in a form that could be readily combined into a computational model.

It follows that some of the equations in the architecture can be improved upon, or combined. Some of the equations had to be developed with a general idea of how the processes may operate given related (but not necessarily direct) experimental data. One of the advantages of the architectural approach is this limitation becomes much more explicit and the path to understanding how these processes interact in a complex computational system becomes somewhat more clear. ACT-R/Φ, and more generally computational behavior architectures, can help us understand and specify what we don't understand about human behavior.

On a more practical side, there also exists a limitation in how fast the architecture software may take to accomplish a task. ACT-R/Φ is limited by hardware (e.g., I/O speed and CPU speed) and the speed of the ACT-R and HumMod systems individually. This speed limitation is difficult to rectify, because though the physiological model and equations underlying the HumMod simulation system are open-source, the *model solver* that processes all the equations is not. Getting the code to such a solver may be possible, and this would likely result in a big increase in architecture speed, but would also mean sacrificing having an open-source system.

Model limitations

The computational process model developed to run within the ACT-R/Φ architecture and complete a computerized version of IGT is one of the few that uses the majority of the learning (and perceptual) mechanisms in the ACT-R architecture. Nonetheless, the model-experimental data revealed that the model could still be improved in several areas.

The differences between model-predicted and experimental data indicate that the model would likely benefit from having a more intricate exploration of how parameter-sets and memory/strategy-sets result in different behavior in the task. The work by Kase (2008) indicates the usefulness of using high performance computing to explore parameter sets. Using a high-performance computing cluster would be especially useful given ACT-R/Φ typically runs slower than ACT-R due to the synchronous communication between ACT-R and HumMod in ACT-R/Φ.

The model also uses a specific equation to transform the reward and loss values to values that are reinforced using the reinforcement learning mechanism in ACT-R. There is a limitation in that only one equation was explored (e.g., see Napoli & Fum, 2010 for another possible transformation equation) and only the default ACT-R reinforcement learning mechanism was used by this model (e.g., see Fu & Anderson, 2006; Veksler, Gray, & Schoelles, 2013 for alternative models that could be used within the architecture).

Experiment limitations

Data from both men and women were collected and participant sex was not balanced, there were more men than women in the task. Given that recent studies have shown a sex-related difference in decision-making in the task (Van den Bos et al., 2013) and the data indicated that sex and group factors may have interacted and affected performance, it would have been useful to have more participants, so that groups were balanced for sex so that sex-related differences could be more reliably explored.

There is also existing evidence that there may be differences in participant performance when completing the computerized version of the task as opposed to the physical card version (Overman & Pierce, 2013). Thus, this task is limited in that the experiment only explores participant performance in the computerized version of the task.

As previously mentioned, though results from the Iowa Gambling Task were originally used as evidence for the somatic marker hypothesis (e.g., Bechara & Damasio, 2005), these interpretations of the data have been questioned (e.g., Colombetti, 2008; Maia & McClelland, 2004). The dissertation work is limited in that it only explores decision-making in the IGT. The work may benefit from exploring the use of the same cognitive, affective, and physiological mechanisms in other decision-making tasks. It has previously been argued that as complexity of a choice increases, structure that affects implicit choices becomes more important (e.g., Thaler & Sunstein, 2008), therefore, it would be useful to present a task similar to the IGT that increases complexity whether through increases deck selections or making the deck contingencies more subtle.

Lastly, though sex:group:block interactions were found in these data, no group:block interactions were found. The aim of the experiment was to see if different sets of affective subliminal stimuli would change decision-making behavior, and the lack of pervasive changes in behavior may be closer tied to the stimuli themselves. It would be useful to run a similar study, but have the stimuli more explicitly conditioned with an explicit more aversive stimulus (e.g., an electric shock). It would be interesting to see if different effects are found as a result of prior conditioning.

Summary

In this dissertation, I presented a novel hybrid computational architecture that provides representations of human behavior on the physiological, affective, and cognitive levels and begins to describe and quantify some connections between these levels. I also presented three computational process models that operate within and are constrained by this new architecture: a *stressed* model that has its performance modulated by changes in physiology due to a psychological stressor; a *thirsty* model that has its choice behavior modulated by its osmolarity (thirst); and an IGT model that has its decision-making affected by subliminal affective stimuli presented during a computerized modified version of the Iowa Gambling Task. Lastly, I presented the results from a study that explored how subliminal visual affective stimuli affect decision-making behavior during the Iowa Gambling Task. The data indicate that this subliminal treatment affected physiological and decision behavior of participants and these effects interacted with sex and personality factors.

As the understanding of the processes that mediate human behavior expand, it can become more difficult to comprehend what the interaction of these processes can mean for the entire behavioral system, whether on the physiological, affective, or cognitive level. This is especially true in decision-making behavior as it involves the interaction of several processes over-time. Having a unified theory that describes the processes that mediate decision-making is important as it provides a better understanding of the constraints and changes that can come into play given different environmental contexts, whether external or internal (e.g., physiological change).

Having an underlying theoretical framework can also make decision-making during a specific task more generalizable as the theory can live in the underlying processes that implement the behavior. ACT-R/Φ is a novel hybrid architecture that can be used to better understand the processes that underlie decision-making, and how systematic changes in physiological, affective, and cognitive processes can alter decision and choice behavior over time. The model and data presented in this dissertation, as well as the new architecture, highlight how differences in physiology, memory, and learning over time may effect decisions. The data also show that stimuli that are not consciously perceived can have noticeable effects of physiological and choice behavior and that sex differences factor into the way that these non-conscious stimuli affect the decision process. This improved understanding of the architecture that constrains our behavior gives us a better opportunity to comprehend why we make the decisions we do and how we can use this knowledge to make better decisions in our professional and personal lives.

Appendices

Appendix A: The IGT model details

Parameter values used by the model:

```
(sgp :esc t :lf .05 :show-focus t :trace-detail low :bll 0.25 :ans 0.1 :rt -3.5 :time-master-start-
  increment 1.0 :egs 1 :ncnar nil :reward-hook igt-r-hook :epl t :alpha 0.1)
(sgp :declarative-num-finsts 8 :bll 0.5 )
(sgp :phys-delay 1 :phys-enabled t)
(sgp :SEEK-enabled t :SEEK-max-rew 1 :SEEK-util-offset t :SEEK-hom-noise 0.099 :SEEK-
  delay 1)
(sgp :FEAR-enabled t)
(sgp :ul t :ult nil :ppm 30)
(sgp :md -1)
(sgp :AA-fear-vis-affect-hook igt-vision-affect :AA-seeking-vis-affect-hook igt-vision-affect
  :AA-dm-noise-switch nil :AA-util-noise-switch nil)
```

Chunk-types and declarative memory present in the model at the beginning of each task run

```
(set-hand-location right 19 5)
(chunk-type integer-value num)
(chunk-type (picture (:include visual-object)))
(chunk-type (reward (:include visual-object)))
(chunk-type (deck (:include visual-object)))
(chunk-type card color)
(chunk-type     (deck-key-pair (:include peck)) value)
(chunk-type (val (:include visual-object)) win loss)
(chunk-type IGT-decks currD aWin aLoss bWin bLoss cWin cLoss dWin dLoss)
(chunk-type (deck-value-pair (:include val)) value)
(chunk-type (card-value-pair (:include val)) value)
(chunk-type IGT-game
          state
```

```
        best-deck
        aPos  ;false or true
        bPos
        cPos
        dPos)

(add-dm
  (start isa chunk) (attend isa chunk)
  (pick-deck isa chunk) (get-vals isa chunk)
  (decide-deck isa chunk) (start-trial isa chunk)
  (start-decision isa chunk)
  (good isa chunk) (bad isa chunk)
  (true isa chunk) (false isa chunk)
  (start-game isa chunk) (find-deck isa chunk)
  (wait-for-reward isa chunk) (wait-for-next-trial isa chunk)

  (greenCard isa card color green)
  (redCard isa card color red)

  (zero isa integer-value num 0)
  (fifty isa integer-value num 50)
  (one-hundred isa integer-value num 100)
  (negative-twenty-five isa integer-value num -25)
  (negative-fifty isa integer-value num -50)
  (negative-seventy-five isa integer-value num -75)
  (negative-one-hundred-fifty isa integer-value num -150)
  (negative-two-hundred isa integer-value num -200)
  (negative-two-hundred-fifty isa integer-value num -250)
  (negative-three-hundred isa integer-value num -300)
  (negative-three-hundred-fifty isa integer-value num -350)
  (negative-one-thousand-two-hundred-fifty isa integer-value num -1250)
  (zero-vals isa val      win 0 loss 0)
  (redCard-zero isa card-value-pair color red win 0 loss 0)
  (greenCard-zero isa card-value-pair color green win 0 loss 0)
  (goal isa IGT-game state start-decision))
  (set-similarities (true false -0.99))
(add-dm-fct
  (list
  (list 'deckA-pos 'isa 'visual-location 'screen-x *deckAX* 'screen-y *deckAY* 'Color 'blue)
  (list 'deckB-pos 'isa 'visual-location 'screen-x *deckBX* 'screen-y *deckBY* 'Color 'blue)
  (list 'deckC-pos 'isa 'visual-location 'screen-x *deckCX* 'screen-y *deckCY* 'Color 'blue)
  (list 'deckD-pos 'isa 'visual-location 'screen-x *deckDX* 'screen-y *deckDY* 'Color 'blue)
```

```
(list 'deckA 'isa 'deck 'value "A" 'screen-pos 'deckA-pos);'screen-x *deckAX* 'screen-y
*deckAY*);
(list 'deckB 'isa 'deck 'value "B" 'screen-pos 'deckB-pos);'screen-x *deckBX* 'screen-y
*deckBY*);
(list 'deckC 'isa 'deck 'value "C" 'screen-pos 'deckC-pos);'screen-x *deckCX* 'screen-y
*deckCY*);
(list 'deckD 'isa 'deck 'value "D" 'screen-pos 'deckD-pos);'screen-x *deckDX* 'screen-y
*deckDY*);
(list 'deckA-key1 'isa 'deck-key-pair 'value "A" 'hand 'right 'finger 'index 'r 0 'theta 0)
(list 'deckB-key2 'isa 'deck-key-pair 'value "B" 'hand 'right 'finger 'middle 'r 0 'theta 0)
(list 'deckC-key3 'isa 'deck-key-pair 'value "C" 'hand 'right 'finger 'ring 'r 0 'theta 0)
(list 'deckD-key4 'isa 'deck-key-pair 'value "D" 'hand 'right 'finger 'index 'r 1 'theta (* 1.5 pi))
(list 'isa 'deck-value-pair 'value "A" 'screen-pos 'deckA-pos 'win 0 'loss 0)
(list 'isa 'deck-value-pair 'value "B" 'screen-pos 'deckB-pos 'win 0 'loss 0)
(list 'isa 'deck-value-pair 'value "C" 'screen-pos 'deckC-pos 'win 0 'loss 0)
(list 'isa 'deck-value-pair 'value "D" 'screen-pos 'deckD-pos 'win 0 'loss 0)))
```

Appendix B: Experimental Materials

Below, I list versions of the questionnaires used for the experiment and parameters used in the EDA data analysis.

PANAS Questionnaire

This scale consists of a number of words that describe different feelings and emotions. Read each word then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way right now, that is, at the present moment

1	2	3	4	5
very slightly or not at all	a little	moderately	quite a bit	extremely

______ interested	______ irritable
______ distressed	______ alert
______ excited	______ ashamed
______ upset	______ inspired
______ strong	______ nervous
______ guilty	______ determined
______ scared	______ attentive
______ hostile	______ jittery
______ enthusiastic	______ active
______ proud	______ afraid

ANPS Questionnaire

Affective Neuroscience Personality Scale 2.4 Name: ____________

Age: [ ] [ ] **Sex:** ○ ○ Please mark bubbles like this ●

	Str Disagree	Disagree	Agree	Str Agree
1. Almost any little problem or puzzle stimulates my interest.	○	○	○	○
2. People who know me well would say I am an anxious person.	○	○	○	○
3. I often feel a strong need to take care of others.	○	○	○	○
4. When I am frustrated, I usually get angry.	○	○	○	○
5. I am a person who is easily amused and laughs a lot.	○	○	○	○
6. I often feel sad.	○	○	○	○
7. Feeling a oneness with all of creation helps give more meaning to my life.	○	○	○	○
8. I like to be the one in a group making the decisions.	○	○	○	○
9. I do not get much pleasure out of looking forward to special events.	○	○	○	○
10. I am not frequently jittery and nervous.	○	○	○	○
11. I think it is ridiculous the way some people carry on around baby animals.	○	○	○	○
12. I never stay irritated at anyone for very long.	○	○	○	○
13. My friends would probably describe me as being too serious.	○	○	○	○
14. I seem to be affected very little by personal rejection.	○	○	○	○
15. Feeling like a part of creation is not an important source of meaning for my life.	○	○	○	○
16. I will gossip a little at times.	○	○	○	○
17. I really enjoy looking forward to new experiences.	○	○	○	○
18. I often think of what I should have done after the opportunity has passed.	○	○	○	○
19. I like taking care of children.	○	○	○	○
20. My friends would probably describe me as hotheaded.	○	○	○	○
21. I am known as one who keeps work fun.	○	○	○	○
22. I often have the feeling that I am going to cry.	○	○	○	○
23. I am often spiritually touched by the beauty of creation.	○	○	○	○
24. I usually avoid activities in which I would be the center of attention.	○	○	○	○
25. I am usually not highly curious.	○	○	○	○
26. I would not describe myself as a worrier.	○	○	○	○
27. Caring for a sick person would be a burden for me.	○	○	○	○
28. I cannot remember a time when I became so angry that I wanted to break something.	○	○	○	○
29. I generally do not like vigorous games which require physical contact.	○	○	○	○
30. I rarely become sad.	○	○	○	○
31. I rarely rely on spiritual inspiration to help me meet important challenges.	○	○	○	○
32. I always tell the truth.	○	○	○	○
33. Seeking an answer is as enjoyable as finding the solution.	○	○	○	○
34. I often cannot fall right to sleep because something is troubling me.	○	○	○	○
35. I love being around baby animals.	○	○	○	○
36. When I get angry, I often feel like swearing.	○	○	○	○
37. I like to joke around with other people.	○	○	○	○
38. I often feel lonely.	○	○	○	○
39. For me, experiencing a connection to all of life is an important source of inspiration.	○	○	○	○
40. When I play games, it is important for me to win.	○	○	○	○
41. I usually feel little eagerness or anticipation.	○	○	○	○
42. I have very few fears in my life.	○	○	○	○
43. I do not especially like being around children.	○	○	○	○
44. When I am frustrated, I rarely become angry.	○	○	○	○
45. I dislike humor that gets really silly.	○	○	○	○
46. I never become homesick.	○	○	○	○
47. For me, spirituality is not a primary source of inner peace and harmony.	○	○	○	○
48. Sometimes I feel like swearing.	○	○	○	○
49. I enjoy anticipating and working towards a goal almost as much as achieving it.	○	○	○	○
50. I sometimes cannot stop worrying about my problems.	○	○	○	○
51. I feel softhearted towards stray animals.	○	○	○	○
52. When someone makes me angry, I tend to remain fired up for a long time.	○	○	○	○
53. People who know me would say I am a very fun-loving person.	○	○	○	○
54. I often think about people I have loved who are no longer with me.	○	○	○	○
55. Contemplating spiritual issues often fills me with a sense of intense awe and possibility.	○	○	○	○
56. If my peers have outperformed me, I would still be happy, if I have nearly met my goals.	○	○	○	○

47321

mdanps24

Please mark bubbles like this ● and not like this ☒ or ◉

Str Disagree | Disagree | Agree | Str Agree

57. I am usually not interested in solving problems and puzzles just for the sake of solving them.
58. My friends would say that it takes a lot to frighten me.
59. I would generally consider pets in my home to be more trouble than they are worth.
60. People who know me well would say I almost never become angry.
61. I do not particularly enjoy kidding around and exchanging "wisecracks."
62. It does not particularly sadden me when friends or family members are disapproving of me.
63. My sense of significance and purpose in life does not come from my spiritual beliefs.
64. I have never "played sick" to get out of something.
65. My curiosity often drives me to do things.
66. I often worry about the future.
67. I feel sorry for the homeless.
68. I tend to get irritated if someone tries to stop me from doing what I want to do.
69. I am very playful.
70. I tend to think about losing loved ones often.
71. Feeling a connection with the rest of humanity motivates me to make more ethical choices.
72. When I play games, I do not mind losing.
73. I rarely feel the need just to get out and explore things.
74. There are very few things that make me anxious.
75. I do not like to feel "needed" by other people.
76. I rarely get angry enough to want to hit someone.
77. I do not tend to see the humor in things many people consider funny.
78. I rarely have the feeling that I am close to tears.
79. The goals I set for myself are not influenced by my spirituality.
80. There have been times in my life when I was afraid of the dark.
81. Whenever I am in a new place, I like to explore the area and get a better feel for my surroundings.
82. I often worry about whether I am making the correct decision.
83. I am the kind of person that likes to touch and hug people.
84. When things do not work out the way I want, I sometimes feel like kicking or hitting something.
85. I like all kinds of games including those with physical contact.
86. I frequently feel downhearted when I cannot be with my friends or loved ones.
87. Spiritual inspiration helps me transcend my limitations.
88. I am not satisfied unless I can stay ahead of my peers.
89. I am not the kind of person that likes probing and investigating problems.
90. I rarely worry about my future.
91. I do not especially want people to be emotionally close to me.
92. I hardly ever become so angry at someone that I feel like yelling at them.
93. I do not frequently ask other people to join me for fun activities.
94. I rarely think about people or relationships I have lost.
95. My choices are not guided by a sense of connectedness with all of life.
96. I have never intentionally told a lie.
97. I often feel like I could accomplish almost anything.
98. I often feel nervous and have difficulty relaxing.
99. I am a person who strongly feels the pain of other people.
100. Sometimes little quirky things people do really annoy me.
101. I see life as being full of opportunities to have fun.
102. I am a person who strongly feels the pain from my personal losses.
103. When working on a project, I like having authority over others.
104. Being embarrassed or looking stupid are among my worst fears.
105. I am not an extremely inquisitive person.
106. I almost never lose sleep worrying about things.
107. I am not particularly affectionate.
108. When people irritate me, I rarely feel the urge to say nasty things to them.
109. Playing games with other people is not especially enjoyable for me.
110. It would not bother me to spend the holidays away from family and friends.
111. Striving to be better than my peers is not important for me.
112. Fear of embarrassment often causes me to avoid doing things or speaking to others.

47321

mdanps24

EDA Batch parameters

EDA data was analyzed using the Ledalab matlab SC analysis script. This allows one to process EDA data individually or using an included batch mode. I used this batch mode to process EDA data from all participants that used the Q-Sensor. Equation B-1 gives the specific command used to run the Ledalab batch script:

$dataDir$ The directory location of the data

$'text2'$ The format of the EDA file

$'downsample', 2$ How much to downsample the data (this makes the data half as large)

$'analyze', 'CDA'$ Type of analysis used for EDA (can be CDA or DDA)

$'export\_era', [0\ 4\ 0.05\ 1]$ Export event-related data into individual Matlab files with 1 response window of 0 – 4 seconds, and a minimum amplitude threshold criterion of $0.05\mu S$

$$Ledalab(dataDir, 'open', 'text2', 'downsample', 2, 'analyze', 'CDA', 'export\_era', [0\ 4\ 0.05\ 1], 1);$$

Equation B-1. The script used to run the Ledalab Matlab EDA analysis script

References

Aarøe, L., & Petersen, M. B. (2013). Hunger games: Fluctuations in blood glucose levels influence support for social welfare. *Psychological Science, 24*(12), 2550-2556.

Adams, R. B., Franklin, R. G., Kveraga, K., Ambady, N., Kleck, R. E., Whalen, P. J., . . . Nelson, A. J. (In press). Amygdala responses to averted vs direct gaze fear vary as a function of presentation speed. *Social Cognitive and Affective Neuroscience*.

Ahn, W., Busemeyer, J. R., Wagenmakers, E., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science, 32*(8), 1376-1402.

Aïte, A., Borst, G., Moutier, S., Varescon, I., Brown, I., Houdé, O., & Cassotti, M. (2013). Impact of emotional context congruency on decision making under ambiguity. *Emotion, 13*(2), 177-182.

Alcaro, A., & Panksepp, J. (2011). The SEEKING mind: Primal neuro-affective substrates for appetitive incentive states and their pathological dynamics in addictions and depression. *Neuroscience & Biobehavioral Reviews, 35*(9), 1805-1820.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: OUP.

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing problem solving in real time: fMRI analysis of the subject-paced Tower of Hanoi. *Journal of Cognitive Neuroscience, 17*(8), 1261-1274.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

Anderson, J. R., Byrne, D., Fincham, J. M., & Gunn, P. (2008). Role of prefrontal and parietal cortices in associative learning. *Cerebral cortex, 18*(4), 904-914.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*(2), 167-207.

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences, 12*(4), 136-143.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128*(2), 186-197.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*(1), 403-450.

Bach, J. (2009). *Principles of synthetic intelligence: PSI: an architecture of motivated cognition*. New York, NY: OUP.

Ballard, K., & Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *NeuroImage, 45*(1), 143-150.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (Vol. 2, pp. 142-162). New York, NY: Cambridge University Press.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: a neurocognitive perspective. [10.1038/nn1584]. *Nat Neurosci, 8*(11), 1458-1463.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior, 52*(2), 336-372.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral cortex, 10*(3), 295.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*(13), 5473.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.

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.

Belavkin, R. V. (2001). *The role of emotion in problem solving.* In Proceedings of the the AISB '01 Symposium on emotion, cognition, and affective computing, pp. 49-57. University of York, United Kingdom.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80-91.

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: 'liking', 'wanting', and learning. *Current Opinion in Pharmacology, 9*(1), 65-73.

Berthoud, H., & Münzberg, H. (2011). The lateral hypothalamus as integrator of metabolic and environmental needs: From electrical self-stimulation to opto-genetics. *Physiology & Behavior, 104*(1), 29-39.

Best, B. J., & Lebiere, C. (2006). Cognitive agents interacting in real and virtual worlds *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (pp. 186-218).

Bourque, C. W. (2008). Central mechanisms of osmosensation and systemic osmoregulation. *Nature Reviews Neuroscience, 9*(7), 519-531.

Bradley, M. M., & Lang, P. J. (2007a). *The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual*. Technical Report B-3. University of Florida, Gainesville, FL.

Bradley, M. M., & Lang, P. J. (2007b). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 29-46): OUP.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: Common and differential networks. *NeuroImage, 31*(2), 906-919.

Büchel, C., Morris, J., Dolan, R. J., & Friston, K. J. (1998). Brain Systems Mediating Aversive Conditioning: an Event-Related fMRI Study. *Neuron, 20*(5), 947-957.

Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies, 55*(1), 41-84.

Cahill, L. (2006). Why sex matters for neuroscience. [10.1038/nrn1909]. *Nat Rev Neurosci, 7*(6), 477-484.

Cahill, L., & Alkire, M. T. (2003). Epinephrine enhancement of human memory consolidation: Interaction with arousal at encoding. *Neurobiology of Learning and Memory, 79*(2), 194-198.

Charmandari, E., Tsigos, C., & Chrousos, G. (2005). Endocrinology of the Stress Response. *Annual Review of Physiology, 67*(1), 259-284.

Cohen, M. A., Ritter, F. E., & Haynes, S. R. (2010). Applying software engineering to agent development. *AI Magazine, 31,* 25-44.

Colombetti, G. (2008). The Somatic Marker Hypotheses, and What the Iowa Gambling Task Does and Does not Show. *The British Journal for the Philosophy of Science, 59*(1), 51-71.
Damasio, A. R. (1994). *Descartes error: Emotion, rationality and the human brain*. New York, NY: Putnam.
Damasio, A. R., & Carvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. [10.1038/nrn3403]. *Nat Rev Neurosci, 14*(2), 143-152.
Dancy, C. L. (2013). ACT-RΦ: A cognitive architecture with physiology and affect. *Biologically Inspired Cognitive Architectures, 6*.
Dancy, C. L., & Kaulakis, R. (2013). *Towards Adding Bottom-Up Homeostatic Affect to ACT-R.* In Proceedings of the 12th International Conference on Cognitive Modeling, Ottawa, Canada.
Dancy, C. L., Ritter, F. E., & Berry, K. (2012). *Towards adding a physiological substrate to ACT-R.* In Proceedings of the 21st Conference on Behavior Representation in Modeling and Simulation, pp. 78-85. Amelia Island, FL.
Dancy, C. L., Ritter, F. E., Berry, K., & Klein, L. C. (in press). Using a cognitive architecture with a physiological substrate to represent effects of psychological stress on cognition. *Computational and Mathematical Organization Theory*.
Davies, J. L., & Turnbull, O. H. (2011). Affective bias in complex decision making: Modulating sensitivity to aversive feedback. *Motivation and Emotion, 35*(2), 235-248.
Davis, K. L., & Panksepp, J. (2011). The brain's emotional foundations of human personality and the Affective Neuroscience Personality Scales. *Neuroscience & Biobehavioral Reviews, 35*(9), 1946-1958.
De Quervain, D. J. F., Aerni, A., Schelling, G., & Roozendaal, B. (2009). Glucocorticoids and the regulation of memory in health and disease. *Frontiers in Neuroendocrinology, 30*(3), 358-370.
Delgado, M. R., Jou, R. L., & Phelps, E. A. (2011). Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. [Original Research]. *Frontiers in Neuroscience, 5*.
Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience & Biobehavioral Reviews, 30*(2), 239-271.
Eichenbaum, H. (2010). Memory systems. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 478-490.
Evertsz, R., Ritter, F. E., Busetta, P., Pedrotti, M., & Bittner, J. L. (2008). *CoJACK - Achieving principled behaviour variation in a moderated cognitive architecture.* In Proceedings of the 17th Conference on Behavior Representation in Modeling and Simulation, pp. 80-89. Orlando, FL.
Evertsz, R., Ritter, F. E., Russell, S., & Shepherdson, D. (2007). *Modeling Rules of Engagement in Computer Generated Forces.* In Proceedings of the 16th Conference on Behavior Representation in Modeling and Simulation, pp. 123-134. Orlando, FL.
Fanselow, M. S. (2010). From contextual fear to a dynamic view of memory systems. *Trends in Cognitive Sciences, 14*(1), 7-15.
Feldman Barrett, L. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science, 1*(1), 28-58.
Feldman Barrett, L. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*(1), 20-46.
Fleetwood, M. D., & Byrne, M. D. (2002). Modeling icon search in ACT-R/PM. *Cognitive Systems Research, 3*(1), 25-33.

Friedman, M. I., & Stricker, E. M. (1976). The physiological psychology of hunger: A physiological perspective. *Psychological Review, 83*(6), 409-431.

Fu, W., & Anderson, J. R. (2006). From recurrent choice to skill learning: A reinforcement-learning model. *Journal of Experimental Psychology: General, 135*(2), 184-206.

Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., & Hanakawa, T. (2005). Functional activity related to risk anticipation during performance of the Iowa gambling task. *NeuroImage, 24*(1), 253-259.

Gillath, O., & Canterberry, M. (2012). Neural correlates of exposure to subliminal and supraliminal sexual cues. *Social Cognitive and Affective Neuroscience, 7*(8), 924-936.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences, 108*(Supplement 3), 15647-15654.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523-551.

Groeneweg, F. L., Karst, H., de Kloet, E. R., & Joëls, M. (2011). Rapid non-genomic effects of corticosteroids and their role in the central stress response. *Journal of Endocrinology, 209*(2), 153-167.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science, 33*(5), 880-910.

Gunzelmann, G., Moore Jr, R. L., Salvucci, D. D., & Gluck, K. A. (2011). Sleep loss and driver performance: Quantitative predictions with zero free parameters. *Cognitive Systems Research, 12*(2), 154-163.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367-388.

Haber, S. N., & Knutson, B. (2009). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4-26.

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology, 14*(2), 233-238.

Hamann, S., Herman, R. A., Nolan, C. L., & Wallen, K. (2004). Men and women differ in amygdala response to visual sexual stimuli. [10.1038/nn1208]. *Nat Neurosci, 7*(4), 411-416.

Harris, R. B. S. (2000). Leptin - Much More Than a Satiety Signal. *Annual Review of Nutrition, 20*(1), 45-75.

Hayashi, T., Ko, J. H., Strafella, A. P., & Dagher, A. (2013). Dorsolateral prefrontal and orbitofrontal cortex interactions during self-control of cigarette craving. *Proceedings of the National Academy of Sciences, 110*(11), 4422-4427.

Hester, R. L., Brown, A. J., Husband, L., Iliescu, R., Pruett, D., Summers, R., & Coleman, T. G. (2011). HumMod: A modeling environment for the simulation of integrative human physiology. *Frontiers in Physiology, 2*(12).

Hester, R. L., Iliescu, R., Summers, R., & Coleman, T. G. (2011). Systems biology and integrative physiological modelling. *Journal of Physiology, 589*(5), 1053-1060.

Hope, R. M., Schoelles, M. J., & Gray, W. D. (2013). Simplifying the interaction between cognitive models and task environments with the JSON Network Interface. *Behavior research methods*, 1-6.

James, W. (1890). *The principles of psychology* (Vol. 1): Dover Publications.

Jennings, J. H., Rizzi, G., Stamatakis, A. M., Ung, R. L., & Stuber, G. D. (2013). The Inhibitory Circuit Architecture of the Lateral Hypothalamus Orchestrates Feeding. *Science, 341*(6153), 1517-1521.

Joëls, M., & Baram, T. Z. (2009). The neuro-symphony of stress. [10.1038/nrn2632]. *Nature Review in Neuroscience, 10*(6), 459-466.

Joëls, M., Fernandez, G., & Roozendaal, B. (2011). Stress and emotional memory: a matter of timing. *Trends in Cognitive Sciences, 15*(6), 280-288.

Johansen, J. P., Tarpley, J. W., LeDoux, J. E., & Blair, H. T. (2010). Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. [10.1038/nn.2594]. *Nat Neurosci, 13*(8), 979-986.

Johnson, A. K. (2007). The sensory psychobiology of thirst and salt appetite. *Medicine & Science in Sports & Exercise, 39*(8), 1388-1400 1310.1249/mss.1380b1013e3180686de3180688.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *The Journal of Economic Perspectives, 5*(1), 193-206.

Kahneman, D., & Thaler, R. H. (2006). Anomalies: Utility maximization and experienced utility. *The Journal of Economic Perspectives, 20*(1), 221-234.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-291.

Kase, S. E. (2008). *Parallel genetic algorithm optimization of a cognitive model: Investigating group and individual performance on a math stressor task.* (Ph.D.), Penn State, University Park, PA.

Keightley, M. L., Winocur, G., Graham, S. J., Mayberg, H. S., Hevenor, S. J., & Grady, C. L. (2003). An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli. *Neuropsychologia, 41*(5), 585-596.

Kennedy, W. G. (2011). *Learning from Others: Soar's memory extensions*. PowerPoint Presentation. George Mason University. ACT-R Post-Graduate Summer School. Retrieved from http://act-r.psy.cmu.edu/workshops/workshop-2011/Kennedy.pdf

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science, 303*(5660), 1023-1026.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The "Trier Social Stress Test": A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81.

Kleiner, M. (2010). Visual stimulus timing precision in psychtoolbox-3: tests, pitfalls and solutions. *Perception, 39*, 189.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of Increasing Monetary Reward Selectively Recruits Nucleus Accumbens. *The Journal of Neuroscience, 21*(16), RC159.

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: Neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1511), 3771-3786.

Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. F. (2007). Neural predictors of purchases. *Neuron, 53*(1), 147-156.

Kobayakawa, M., Koyama, S., Mimura, M., & Kawamura, M. (2008). Decision making in Parkinson's disease: Analysis of behavioral and physiological patterns in the Iowa gambling task. *Movement Disorders, 23*(4), 547-552.

Kober, H., Mende-Siedlecki, P., Kross, E. F., Weber, J., Mischel, W., Hart, C. L., & Ochsner, K. N. (2010). Prefrontal–striatal pathway underlies cognitive regulation of craving. *Proceedings of the National Academy of Sciences, 107*(33), 14811-14816.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. [10.1038/nrn1825]. *Nature Reviews Neuroscience, 7*(1), 54-64.

Laird, J. E. (2008). *Extending the Soar Cognitive Architecture.* In Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference, pp. 224-235.

Laird, J. E. (2012). *The Soar cognitive architecture*. Cambridge, MA: MIT Press.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50*(5), 372-385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. The Center for Research in Psychophysiology, University of Florida. Gainesville, FL.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer Publishing Company.

LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York, NY: Simon & Schuster.

LeDoux, J. E. (2012). Rethinking the Emotional Brain. *Neuron, 73*(4), 653-676.

LeDoux, J. E., & Phelps, E. A. (2008). Emotional Networks in the Brain. In M. S. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 159-179). New York, NY: Guilford Press.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology, 22*(6), 1027-1038.

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. [10.1038/nn.2904]. *Nat Neurosci, 14*(10), 1250-1252.

Li, X., Lu, Z.-L., D'Argembeau, A., Ng, M., & Bechara, A. (2010). The Iowa Gambling Task in fMRI images. *Human brain mapping, 31*(3), 410-423.

Liddell, B. J., Brown, K. J., Kemp, A. H., Barton, M. J., Das, P., Peduto, A., . . . Williams, L. M. (2005). A direct brainstem–amygdala–cortical ‘alarm’ system for subliminal signals of fear. *NeuroImage, 24*(1), 235-243.

Liotti, M., & Panksepp, J. (2004). Imaging human emotions and affective feelings: Implications for biological psychiatry. In M. Liotti & J. Panksepp (Eds.), *Textbook of Biological Psychiatry* (pp. 33-75). Hoboken, NJ: Wiley.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological bulletin, 127*(2), 267.

Lovallo, W. R., Al'absi, M., Blick, K., Whitsett, T. L., & Wilson, M. F. (1996). Stress-like adrenocorticotropin responses to caffeine in young healthy men. *Pharmacology Biochemistry and Behavior, 55*(3), 365-369.

Lovallo, W. R., Whitsett, T. L., al Absi, M., Sung, B. H., Vincent, A. S., & Wilson, M. F. (2005). Caffeine stimulation of cortisol secretion across the waking hours in relation to caffeine intake levels. *Psychosomatic medicine, 67*(5), 734.

Maia, T. V., & McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America, 101*(45), 16075-16080.

Maren, S., & Quirk, G. J. (2004). Neuronal signalling of fear memory. [10.1038/nrn1535]. *Nat Rev Neurosci, 5*(11), 844-852.

Marewski, J. N., & Link, D. (2014). Strategy selection: An introduction to the modeling challenge. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(1), 39-59.

Marewski, J. N., & Mehlhorn, K. (2011). Using the ACT-R architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making, 6*(6), 439–519.

Marinier III, R. P., Laird, J. E., & Lewis, R. L. (2009). A computational unification of cognitive behavior and emotion. *Cognitive Systems Research, 10*(1), 48-69.

Marsella, S., & Gratch, J. (2009). EMA: A process model of appraisal dynamics. *Cognitive Systems Research, 10*(1), 70-90.

McClure, S. M., Daw, N. D., & Read Montague, P. (2003). A computational substrate for incentive salience. *Trends in Neurosciences, 26*(8), 423-428.

McClure, S. M., Gilzenrat, M. S., & Cohen, J. D. (2006). *An exploration-exploitation model based on norepinepherine and dopamine activity.* In Proceedings of the Advances in neural information processing systems, pp. 867-874. Cambridge, MA.

McMorris, T. (2009). Exercise, dehydration and, cognitive function *Exercise and Cognitive Function* (pp. 115-134): John Wiley & Sons, Ltd.

Mikels, J., Fredrickson, B., Larkin, G., Lindberg, C., Maglio, S., & Reuter-Lorenz, P. (2005). Emotional category data on images from the international affective picture system. *Behavior research methods, 37*(4), 626-630.

Miyashita, T., & Williams, C. L. (2006). Epinephrine administration increases neural impulses propagated along the vagus nerve: Role of peripheral β-adrenergic receptors. *Neurobiology of Learning and Memory, 85*(2), 116-124.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., . . . Frith, C. D. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science, 317*(5841), 1079-1083.

Mogg, K., Bradley, B. P., Hyare, H., & Lee, S. (1998). Selective attention to food-related stimuli in hunger: Are attentional biases specific to emotional and psychopathological states, or are they also found in normal drive states? *Behaviour Research and Therapy, 36*(2), 227-237.

Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393*(6684), 467-470.

Moscarello, J. M., & LeDoux, J. E. (2013). The Contribution of the Amygdala to Aversive and Appetitive Pavlovian Processes. *Emotion Review, 5*(3), 248-253.

Napoli, A., & Fum, D. (2010). *Rewards and punishments in iterated decision making: An explanation for the frequency of the contingent event effect.* In Proceedings of the International Conference on Cognitive Modeling, pp. 175-180. Philadelphia, PA.

Newell, A. (1973). You Can't Play 20 Questions with Nature and Win: Projective Comments on the Papers of this Symposium. *Visual information processing: proceedings*, 283-309.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, Massachusetts: Harvard University Press.

Newell, A. (1991). Desires and diversions. Carnegie Mellon University.

Öhman, A. (2008). Fear and anxiety: Overlaps and dissociations. In M. S. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 709-729). New York, NY: Guilford Press.

Öhman, A., Carlsson, K., Lundqvist, D., & Ingvar, M. (2007). On the unconscious subcortical origin of human fear. *Physiology & Behavior, 92*(1-2), 180-185.

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483-522.

Overman, W. H., & Pierce, A. (2013). Iowa Gambling Task with non-clinical participants: Effects of using real + virtual cards and additional trials. [Review]. *Frontiers in Psychology, 4*.

Packard, M. G., & McGaugh, J. L. (1996). Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. *Neurobiology of Learning and Memory, 65*(1), 65-72.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: OUP.

Panksepp, J. (2011). Empathy and the laws of affect. *Science, 334*(6061), 1358-1359.

Panksepp, J., & Biven, L. (2012). *The Archeology of Mind: Neuroevoloutionary Origins of Human Emotions*. New York, NY: W.W. Norton & Company.

Panksepp, J., Fuchs, T., & Iacobucci, P. (2011). The basic neuroscience of emotional experiences in mammals: The case of subcortical FEAR circuitry and implications for clinical anxiety. *Applied Animal Behaviour Science, 129*(1), 1-17.

Pavlik Jr, P. I., & Anderson, J. R. (2005). Practice and Forgetting Effects on Vocabulary Memory: An Activation-Based Model of the Spacing Effect. *Cognitive Science, 29*(4), 559-586.

Pew, R. W., & Mavor, A. S. (Eds.). (1998). *Modeling human and organizational behavior: Application to military simulations*. Washington, DC: National Academies Press.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI. *NeuroImage, 16*(2), 331-348.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology, 57*, 27-53.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2013). *R Development Core Team. 2013. nlme: linear and nonlinear mixed effects models. R package version 3.1-113*.

Raio, C. M., Carmel, D., Carrasco, M., & Phelps, E. A. (2012). Nonconscious fear is quickly acquired but swiftly forgotten. *Current biology : CB, 22*(12), R477-R479.

Reyes, B. A. S., Carvalho, A. F., Vakharia, K., & Van Bockstaele, E. J. (2011). Amygdalar peptidergic circuits regulating noradrenergic locus coeruleus neurons: Linking limbic and arousal centers. *Experimental Neurology, 230*(1), 96-105.

Richard, J. M., Castro, D. C., DiFeliceantonio, A. G., Robinson, M. J. F., & Berridge, K. C. (in press). Mapping brain circuits of reward and motivation: In the footsteps of Ann Kelley. *Neuroscience & Biobehavioral Reviews*.

Rick, S., & Loewenstein, G. F. (2008). The role of emotion in economic behavior. In M. S. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 138-156). New York, NY: Guilford Press.

Ritter, F. E., Bittner, J. L., Kase, S. E., Evertsz, R., Pedrotti, M., & Busetta, P. (2012). CoJACK: A high-level cognitive architecture with demonstrations of moderators, variability, and implications for situation awareness. *Biologically Inspired Cognitive Architectures, 1*, 2-13.

Ritter, F. E., Kase, S. E., Klein, L. C., Bennett, J., & Schoelles, M. (2009). *Fitting a model to behavior tells us what changes cognitively when under stress and with caffeine.* In Proceedings of the the Biologically Inspired Cognitive Architectures Symposium at the AAAI Fall Symposium. Keynote presentation, pp. 109-115. Washington, DC.

Ritter, F. E., Reifers, A. L., Klein, L. C., & Schoelles, M. J. (2007). Lessons from defining theories of stress. In W. D. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 254-262). New York, NY: OUP.

Ritter, F. E., Shadbolt, N. R., Elliman, D., Young, R. M., Gobet, F., & Baxter, G. D. (2003). *Techniques for modeling human performance in synthetic environments: A supplementary review*. DTIC Document. Wright-Patterson Air Force Base, OH: Human Systems Information Analysis Center.

Ritter, F. E., Van Rooy, D., St. Amant, R., & Simpson, K. (2006). Providing user models direct access to interfaces: An exploratory study of a simple interface with implications for HRI and HCI. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 36*(3), 592-601.

Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1507), 3137-3146.

Roesch, M. R., Calu, D. J., Esber, G. R., & Schoenbaum, G. (2010). Neural Correlates of Variations in Event Processing during Learning in Basolateral Amygdala. *The Journal of Neuroscience, 30*(7), 2464-2471.

Rolls, E. T. (2013). What are Emotional States, and Why Do We Have Them? *Emotion Review, 5*(3), 241-247.

Roozendaal, B., Hahn, E. L., Nathan, S. V., de Quervain, D. J.-F., & McGaugh, J. L. (2004). Glucocorticoid Effects on Memory Retrieval Require Concurrent Noradrenergic Activity in the Hippocampus and Basolateral Amygdala. *The Journal of Neuroscience, 24*(37), 8161-8169.

Roozendaal, B., Okuda, S., Van der Zee, E. A., & McGaugh, J. L. (2006). Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. *Proceedings of the National Academy of Sciences, 103*(17), 6741-6746.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. [10.1038/nrn2573]. *Nature Reviews Neuroscience, 10*(3), 211-223.

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: The locus coeruleus mediates cognition through arousal. *Neuron, 76*(1), 130-141.

Scherer, K. R., Schorr, A., & Johnstone, T. (2001). *Appraisal processes in emotion: Theory, methods, research*. New York, NY: OUP.

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: An update and integration. *Neuroscience & Biobehavioral Reviews, 36*(7), 1740-1749.

Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: from 'thinking' to 'doing'. *Trends in Cognitive Sciences, 17*(2), 60-68.

Sen, A. (1993). Internal consistency of choice. *Econometrica, 61*(3), 495-521.

Shurman, B., Horan, W. P., & Nuechterlein, K. H. (2005). Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the Iowa Gambling Task. *Schizophrenia Research, 72*(2–3), 215-224.

Silverman, B. G., Pietrocola, D., Nye, B., Weyer, N., Osin, O., Johnson, D., & Weaver, R. (2012). Rich socio-cognitive agents for immersive training environments: case of NonKin Village. *Autonomous Agents and Multi-Agent Systems, 24*(2), 312-343.

Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and Organization*. Amsterdam: North-Holland.

Simon, H. A. (1996). *The science of the artificial* (3rd ed.). Cambridge, MA: MIT press.

Smith, C. A., & Kirby, L. D. (2009). Putting appraisal in context: Toward a relational model of appraisal and emotion. *Cognition & Emotion, 23*(7), 1352-1372.

Smith, C. A., & Lazarus, R. S. (1990). Emotion and adaptation. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 609-637). New York, NY: Guilford.

Smith, K. S., Berridge, K. C., & Aldridge, J. W. (2011). Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proceedings of the National Academy of Sciences, 108*(27), E255–E264.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*(3), 171-177.

Stanovich, K. E. (2009). *Decision making and rationality in the modern world*. New York, NY: OUP.

Stocco, A., Fum, D., & Napoli, A. (2009). Dissociable processes underlying decisions in the Iowa Gambling Task: a new integrative framework. *Behavioral and Brain Functions, 5*(1), 1.

Suhr, J. A., & Tsanadis, J. (2007). Affect and personality correlates of the Iowa Gambling Task. *Personality and Individual Differences, 43*(1), 27-36.

Sun, R. (2006). The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In R. Sun (Ed.), *Cognition and MultiAgent Interaction* (pp. 79-99). New York, NY: Cambridge University Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge, MA: MIT Press.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*: Yale University Press.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*(4), 273-286.

Tuk, M. A., Trampe, D., & Warlop, L. (2011). Inhibitory Spillover. *Psychological Science, 22*(5), 627-633.

Turnbull, O. H., Evans, C. E. Y., Bunce, A., Carzolio, B., & O'Connor, J. (2005). Emotion-based learning and central executive resources: An investigation of intuition and the Iowa Gambling Task. *Brain and Cognition, 57*(3), 244-247.

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

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. [10.1038/nrn2647]. *Nature Reviews Neuroscience, 10*(6), 397-409.

Ungless, M. A., Argilli, E., & Bonci, A. (2010). Effects of stress and aversion on dopamine neurons: Implications for addiction. *Neuroscience & Biobehavioral Reviews, 35*(2), 151-156.

Van Bockstaele, E. J., Reyes, B. A. S., & Valentino, R. J. (2010). The locus coeruleus: A key nucleus where stress and opioids intersect to mediate vulnerability to opiate abuse. *Brain Research, 1314*(0), 162-174.

Van den Bos, R., Harteveld, M., & Stoop, H. (2009). Stress and decision-making in humans: Performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology, 34*(10), 1449-1458.

Van den Bos, R., Homberg, J., & De Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the Iowa Gambling Task. *Behavioural Brain Research, 238*(0), 95-108.

Varma, S. (2011). Criteria for the design and evaluation of cognitive architectures. *Cognitive Science, 35*(7), 1329-1351.

Veksler, V. D., Gray, W. D., & Schoelles, M. J. (2013). Goal-Proximity Decision-Making. *Cognitive Science, 37*(4), 757-774.

Verdejo-García, A., & Bechara, A. (2009). A somatic marker theory of addiction. *Neuropharmacology, 56, Supplement 1*(0), 48-62.

Verdejo-Garcia, A., Benbrook, A., Funderburk, F., David, P., Cadet, J., & Bolla, K. I. (2007). The differential relationship between cocaine use and marijuana use on decision-making performance over repeat testing with the Iowa Gambling Task. *Drug and Alcohol Dependence, 90*(1), 2-11.

Walsh, B. J., Buonocore, M. H., Carter, C. S., & Mangun, G. R. (2010). Integrating conflict detection and attentional control mechanisms. *Journal of Cognitive Neuroscience, 23*(9), 2211-2221.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious Affective Reactions to Masked Happy Versus Angry Faces Influence Consumption Behavior and Judgments of Value. *Personality and Social Psychology Bulletin, 31*(1), 121-135.

Wolf, O. T. (2009). Stress and memory in humans: Twelve years of progress? *Brain Research, 1293*(0), 142-154.

Wright, J., & Panksepp, J. (2012). An Evolutionary Framework to Understand Foraging, Wanting, and Desire: The Neuropsychology of the SEEKING System. *Neuropsychoanalysis, 14*(1), 5-39.

Wright, N. D., Hodgson, K., Fleming, S. M., Symmonds, M., Guitart-Masip, M., & Dolan, R. J. (2012). Human responses to unfairness with primary rewards and their biological limits. [10.1038/srep00593]. *Scientific Reports, 2*.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459-482.

Zhang, J., Berridge, K. C., Tindell, A. J., Smith, K. S., & Aldridge, J. W. (2009). A Neural Computational Model of Incentive Salience. *PLoS Comput Biol, 5*(7), e1000437.

Zhao, C., Kaulakis, R., Morgan, J. H., Hiam, J. W., Ritter, F. E., Sanford, J. P., & Morgan, G. P. (Accepted). Building social networks out of cognitive blocks: Factors of interest in agent-based socio-cognitive simulations. *Computational and Mathematical Organization Theory*.

Zhao, C., Kaulakis, R., Morgan, J. H., Hiam, J. W., Sanford, J. P., Ritter, F. E., & Morgan, G. P. (2012). *Socio-cognitive networks: Modeling the effects of space and memory on generative social structure.* In Proceedings of the 21st Conference on Behavior Representation in Modeling and Simulation, pp. 24-31. Amelia Island, FL.

Christopher L. Dancy II

Education

Ph.D. in Information Sciences and Technology, AI & Cognitive Science focus
The Pennsylvania State University – University Park, PA
May, 2014

B.S. in Computer Science
The Pennsylvania State University – University Park, PA
May, 2010

Awards and Honors

Awards

Best Paper, *Behavior Representation in Modeling and Simulation (BRIMS) Conference*, 2012
Best Student Paper, *Behavior Representation in Modeling and Simulation (BRIMS) Conference*, 2012

Selected Fellowships/Scholarships

The Penn State Alumni Association Dissertation Award (Nominated), 2014
IST Graduate Travel Award, 2013
The Sloan Scholarship, 2013
The Ford Foundation Dissertation Fellowship, Honorable Mention, 2013
The Pennsylvania Space Grant Consortium Graduate Fellowship, 2011
CNA Corporate Scholarship, 2011
Bunton-Waller Graduate Fellowship, 2010
The Science, Mathematics And Research for Transformation (S.M.A.R.T.) Scholarship, 2008

Selected Journal Articles and Conference Publications

*Denotes an award winning paper

- Dancy, C. L., Ritter, F. E., & Berry, K. (in press). Using a cognitive architecture with a physiological substrate to represent effects of psychological stress on cognition. *Computational & Mathematical Organization Theory*.
- Dancy, C. L. (2013). ACT-RΦ: A cognitive architecture with physiology and affect. *Biologically Inspired Cognitive Architectures*, 6.
- Dancy, C. L., & Kaulakis, R. (2013). Towards adding bottom-up homeostatic affect to ACT-R. *Proceedings of the Twelfth International Conference on Cognitive Modeling*: Ottawa, Canada.
- *Dancy, C. L., Ritter, F. E., & Berry, K. (2012). Towards adding a physiological substrate to ACT-R. In *Proceedings of the 21st Conference on Behavior Representation in Modeling and Simulation*, 78-85. The BRiMS Society:Amelia Island, FL.
- Tyworth, M., Giacobe, N. A., Mancuso, V., Dancy, C.L. (2012) The Distributed Nature of Cyber Situation Awareness. *IEEE 2nd International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA).* New Orleans, LA.
- Medaglia, J.D., Dancy, C.L., Bochnakova, T., DePinto, N., & Hillary, F.G (2013). A Hybrid Data Reduction Approach for Examining Node Structures in fMRI Data. (Abstract) In *Proceedings of the 41st Annual Meeting of the International Neuropsychological Society*.