NORTHWESTERN UNIVERSITY

The Dynamics of Affect: Using Newtonian Mechanics, Reinforcement Sensitivity Theory, and the Cues-Tendencies-Actions Model to Simulate Individual Differences in Emotional Experience

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Ashley D. Brown

EVANSTON, ILLINOIS

September 2017

ABSTRACT

The Dynamics of Affect: Using Newtonian Mechanics,
Reinforcement Sensitivity Theory, and the Cues-Tendencies-Actions Model
to Simulate Individual Differences in Emotional Experience

Ashley D. Brown

Reinforcement sensitivity theory (RST) has enormous potential to become a paradigmatic model of individual differences. However, while its foundations in experimental genetic and neurophysiological research on nonhuman animals are among the strongest in personality psychology, it has perhaps not gained the foothold within the field that it deserves. It is the purpose of this research to introduce a computational model of RST system functioning capable of generating hypotheses and resolving conceptual disagreements among the theory's proponents.

Here, I first outline RST's development, focusing on the revised version of the theory. Notably, the revised RST seems to suffer from a dearth of good self-report measures, perhaps because of confusion regarding certain of the theory's central constructs. I then propose a means of bridging the gap between RST's 'bottom-up' neuroscience roots and contemporary practitioners' efforts to build 'top-down' self-report questionnaires that capture its central constructs; namely, a dynamic computational model that generates the observable between-persons differences in behavior and affect on which self-report measures are based from the unobservable within-persons differences on which RST is premised. Next, I describe the process by which I created the 'CTA-RST' model, which unites the Cues-Tendencies-Actions model (CTA, Revelle & Condon, 2015) and a simulation of RST developed by Alan Pickering (2008) . Finally, I demonstrate the model's efficacy by simulating three real studies; no previous simulations of RST or related theories have modeled real data in any detail.

Specifically, in Chapter 4 I recreate two experiments by Smillie, Cooper, Wilt, and Revelle (2012), showing that extraversion is related to the experience of energetic affect under reward-pursuit conditions, but not of pleasant positive affect under reward-consumption conditions. In Chapter 5, I model research by Wilt, Bleidorn, and Revelle (2016), replicating the finding that approach goal 'velocity' mediates the relationships between state personality and affect. In Chapter 6, I simulate a study by Wilt, Funkhouser, and Revelle (2011) and show, as they did, that individuals differ in the tendency to experience activated and tense affect simultaneously as a function of the interaction between affective traits. The median correlation between real and simulated studies' tabulated effect sizes was $r = 0.73$.

Contents

List of Figures

List of Tables

Chapter 1

Reinforcement Sensitivity and its Discontents

> The weather ... is a matter about which a great deal is said and very little done.
>
> Attributed to Charles Dudley Warner, 1889 (The Book Buyer)

As centuries have passed scholars have, at the very least, been able to agree that their purpose in studying personality is to understand how and why people differ in the spatiotemporal patterning of their thoughts, feelings, behaviors, and desires (Allport, 1937; Ortony, Norman, & Revelle, 2005); at most, they know much more now about personality than they did in the 5th century BCE (Aristotle, 2014; Confucius, 2005; Hippocrates & Galen, 1846; Theophrastus, 1927). Nevertheless, personality, like the weather, is one of those subjects about which people talk much and do little. At least one model of personality seems, however, to contain the potential to allow us to move further and do more.

Gray and colleagues' (2000) Reinforcement Sensitivity Theory's (RST's) foundations in

experimental research are among the strongest in personality psychology. However, RST remains a relatively marginal theory, compared to flourishing five-factor trait paradigms like the Big Five and the Five Factor Model (B5, FFM; Costa & McCrae, 1992; Goldberg, 1990, respectively). Specifically, RST is plagued by general conceptual confusion, which has led to mixed empirical support and poor self-report measures. I claim that computational modeling is a powerful tool that can be used to solve many, if not all, of RST's theoretical problems; in what follows, I will discuss a new neural network model of RST and present three simulations of real research studies that provide support for my claim.

1.1 A Problem and a Potential Solution

The study of individual differences has long been fraught with intellectual tribalism, and the field remains divided despite efforts to integrate and systematize divergent approaches (e.g., Corr, DeYoung, & McNaughton, 2013; Elliot & Thrash, 2002; Revelle, 1995). In order to understand why RST hasn't gained the foothold among models of human personality that it seems to deserve, it will be beneficial to explore two orthogonal conceptual distinctions; namely, the 'top-down/bottom-up' and 'between-/within-persons' dichotomies (Poropat & Corr, 2015). I will begin here by elaborating upon the top-down/bottom-up divide and move on in later sections to explain how the missing link between these divergent approaches is related to the as-yet untrodden middle ground between the study of interindividual and intraindividual variability; i.e. between structure and process models of personality (Collins, Jackson, Walker, O'connor, & Gardiner, 2017).

Theorists who tackle personality from a top-down perspective focus first on describing and cataloging extant individual differences, usually operationalized as sets of items in

self-report questionnaires. Next, factor analysis is used to ascertain patterns of covariance among items in order to identify item clusters or 'traits' and to deduce the mechanisms (social, cognitive, neurological, or genetic) that underlie the latter. Such theories tend to, but need not (H. J. Eysenck & Eysenck, 1967), be heavily reliant on Galton's (1884) lexical hypothesis, which posits that the adjectives people use to describe their own and others' personalities reflect real variations therein.

Factor-analytic, top-down models range from that of the dubious (Revelle & Wilt, 2013) 'General Factor' (Rushton, Bons, & Hur, 2008) to the more credible 'Giant Two' (DeYoung, Quilty, & Peterson, 2007; Digman, 1997) to Ashton and Lee's 'Big Six' (2007). Such theories are primarily descriptive, but they do not fail to address basic questions regarding traits' causes. For instance, Eysenck's (1981) three-factor PEN model proposed well-articulated physiological mechanisms to account for the emergence of the factor-analytically derived traits Psychoticism, Extraversion, and Neuroticism. Similarly, Ashton, Lee, and colleagues (2001; 2007) have done their best to ground the Big Six in (ever-tenuous) evolutionary argumentation.

Other theorists, for instance Jeffrey Gray (1982) and his intellectual descendants (Corr, 2004; Corr, DeYoung, & McNaughton, 2013; Smillie, 2008), have approached the problem of personality from a bottom-up perspective. This approach first posits neurophysiological systems that differ subtly among individuals; the traits that are likely to emerge may then be deduced from the nature of the systems involved and the manner in which they differ. Such systems exhibit varying degrees of biological specificity, ranging from single, discrete brain structures (e.g. the nucleus accumbens) to widely-distributed networks encompassing or thought to encompass many such structures (e.g. RST's BAS, on which I will elaborate later). A thorough understanding of causality remains elusive

within psychology (Pearl, 2009); nevertheless, RST has a firm basis in experimental animal research (Gray & McNaughton, 2000) and may therefore make stronger causal claims than top-down theories like the B5.

Bottom-up theories like Gray's RST (2000) represent a valuable complement to top-down theories like the B5 (1990) or FFM (1992). In the next section, I will review some basic information about RST's utility and history. Keep in mind, however, that RST has its drawbacks; these will be discussed at greater length in Chapter 2.

1.2 Reinforcement Sensitivity Theory: Utility and History

1.2.1 Theoretical Utility

A primary advantage of bottom-up personality theories like RST (Gray, 1982; Gray & McNaughton, 2000) is that they are derived from bodies of research that include experiments. In general, the more experimental research there is informing a theory, the more credible its causal claims. In particular, experimental research that involves brain lesioning or selective breeding allows strong causal claims to be made about fundamental neurological or genetically-determined systems that might differ among individuals and give rise to personality. Naturally, such research can, ethically speaking, only be performed on nonhuman animals. Although it is true that this makes research findings more difficult to generalize to humans (Matthews, 2008), there is also good reason to suppose that humans have much in common with other animals (McNaughton & Corr, 2008a; Revelle, 2008). Therefore, it seems safe to say that the strongest causal claims of all are made by bottom-up theories of personality that are founded in animal research. Gray's RST epitomizes this neurophysiological experimental approach.

1.2.2 RST's Origin Story

Gray built his theory around his experimental work on rats, which was directed in large measure toward understanding the neurobiological underpinnings of anxiety. Using a wide variety of behavioral measures, Gray and his team established convergent and divergent validity for an operational definition of (rat) anxiety. A complete list of these behaviors, which often involve resolving conflicts between goals (e.g. passive avoidance, two-way avoidance) may be found in Gray (1982), Table 6.5 (p.178); similarly, one may refer to Gray and McNaughton (2000), Table 4.2 (p.74) and online Appendix 8, which is currently only available upon request from Neil McNaughton (2017, personal communication). Crucially, Gray noted that when rats' septohippocampal systems were lesioned, they behaved less anxiously; similar diminutions in anxious behavior were observed when rats with unlesioned brains were given anxiolytic drugs.

For example, in several studies Gray used the open field test, developed by Hall (1934), to operationalize anxiety. The open field is a small (1.2 m in diameter, 0.45 m high), circular, walled enclosure that's brightly lit and open at the top, allowing researchers to observe the behavior of rats within (one is rather reminded of the Roman Colosseum). While they are in the open field, more anxious (or, in Hall's words, 'emotional') rats spend less time at the center of the enclosure (in 'center field'), eat less, and defecate more (Hall, 1934; Prut & Belzung, 2003). In another series of experiments, Broadhurst (1975) selectively bred the rats that defecated the most in 'center field' (but not more elsewhere); these ultra-anxious rats were dubbed 'the Maudsley reactive strains,' and they were used to further investigate the neural basis of anxiety and the palliative effects of psychotropic medication (Smillie, Loxton, & Avery, 2011).

Gray's foray into human personality (and, eventually, top-down theory-building) was

largely a reaction to his encounter with Hans Eysenck's theory (Corr, 2004). Gray argued that the fundamental dimensions of personality are not arousal-based, as Eysenck proposed (1967), but are instead predicated on brain systems that respond to cues (conditioned or unconditioned) for reward and punishment. Buttressing his claims with evidence from his own lab (Gray, 1982; Gray & McNaughton, 2000), Gray speculated that Eysenck's factor-analytically derived traits, neuroticism (N) and extraversion (E), ought to be rotated roughly 45 (later, 30) degrees in order to reflect an alternative (hypothetically more basic) pair, anxiety and impulsivity. In Gray's model, the high end of the anxiety dimension (reflecting sensitivity to cues for punishment) is located between Eysenck's high N and low E, whereas high impulsivity (or sensitivity to cues for reward) is located between Eysenck's high N and high E (Corr, 2004; H. J. Eysenck, 1999; Smillie, Loxton, & Avery, 2011).

RST is a reasonable and, in many ways, an elegant theory. Gray's interest in the behaviorist learning literature informed his research and his theories (Corr, 2004); for instance, RST's exclusive emphasis on rewards and punishments reflects its behaviorist foundations. However, the idea that reward and punishment are the ultimate determinants of behavior substantially predates behaviorism; conceptual antecedents include Freud's 'pleasure principle' (1924) and the 'hedonic principle' of ancient Greek philosophy (Smillie, Loxton, & Avery, 2011). These ideas' enduring intuitive appeal has doubtless a great deal to do with the fact that any understanding of the differences we observe between or within species must necessarily be grounded in an understanding of what are the fundamental similarities among all living organisms (Revelle, 2008). For this and many other reasons, RST deserves a closer look; let's give it one.

1.2.3 The Original RST

There are two incarnations of Gray's RST, an 'original' version (1982) and a 'revised' version (2000); I will refer to these as RST1 and RST2, respectively. RST1 postulated three systems responsible for determining organisms' behavior. The Behavioral Approach System (BAS) governs responsivity to conditioned cues for reward (or the termination of punishment). The Fight-Flight System (FFS) governs responsivity to unconditioned cues for punishment (i.e., innately painful stimuli). The Behavioral Inhibition System (BIS) governs responsivity to conditioned cues for punishment (or the termination of reward) and extreme novelty, as well as to high intensity stimuli and innate fear stimuli. Specifically, the BIS performs a 'comparator function;' in 'just checking' mode, the comparator has not detected any threatening cues and thus remains 'offline;' in 'control mode,' it is 'online,' having detected a threat.

Behaviorally, the FFS activates behaviors designed to defend the organism against imminent threats, creating fight or flight. The RST1 BAS activates behaviors intended to secure rewards; it is responsible not only for approach behavior and learning, but also for active avoidance, directed escape, and 'offensive fighting' or predatory aggression. The RST1 BIS inhibits behaviors that might incur punishment, increases arousal, and heightens attention (e.g., creating 'scanning' or 'orienting' behavior in rats). In addition, the RST1 BIS is responsible for passive avoidance behaviors and extinction of conditioned appetitive responses (Torrubia, Ávila, Moltó, & Caseras, 2001). Although the RST1 BIS and BAS are theoretically independent, they are also capable of interacting, primarily by way of mutual inhibition (Gray & Smith, 1969). Gray suggested that such interactions were most likely to occur in instances of approach-avoidance conflict and discriminative learning.

With regard to RST1 as applied to human personality, Gray's modification of Eysenck's theory is discussed above; furthermore, Gray proposed that the BAS governs 'impulsivity,' whereas the BIS governs 'anxiety.' Eysenck's third trait, Psychoticism (P), was thought to correlate negatively with BIS sensitivity and positively with BAS and FFS sensitivity. Affectively, the RST1 BAS is responsible for positive affect, while the the RST1 BIS and FFS are responsible for negative affect. Allegedly, the RST1 BIS was associated with both anxiety and fear; the FFS was linked with more intense affects like panic and rage (Corr, 2004).

1.2.4 The Revised RST

RST2 is quite similar to RST1. There are still three systems; namely, the BAS, BIS, and FFFS (all acronyms except for the FFFS, discussed below, stand for exactly the same things as they did in RST1). RST2 BAS still functions as the primary 'reward sensitivity' system, but now it handles unconditioned as well as conditioned cues for reward. It is also no longer the system responsible for 'active avoidance' (the FFFS has subsumed that role). The BAS still creates positive affect and governs approach behavior, including offensive or predatory aggression. The FFFS and RST2 BIS make larger departures from their original incarnations, as they incorporate and account for results of recent research that emphasize the neurophysiological distinctness of fear and anxiety (Corr, 2004; Gray & McNaughton, 2000).

The revised version of the FFS, the Fight-Flight-Freeze System (FFFS), controls all active defensive behaviors; in RST1, it was responsible only for undirected, automatic fleeing or defensive fighting behavior. In RST2, the FFFS mediates responses to *all* aversive, punishing, or threatening stimuli, whether such cues are conditioned or

unconditioned (Corr, 2004; Gray & McNaughton, 2000); the FFS only mediated responses to unconditioned cues. Neurologically, the FFFS is complex and widely-distributed throughout the evolutionarily older and newer parts of the brain; in ascending hierarchical order, structures associated with the FFFS include the periaqueductal grey, medial hypothalamus, amygdala, and anterior cingulate cortex. More 'primitive' parts of the brain are responsible for handling more immediate, more intense threats.

Affectively, however, the FFFS still does not produce all negative emotions, only fear and panic (Corr, 2004; Corr, DeYoung, & McNaughton, 2013; Smillie, 2008). Anxiety is as much a BIS domain in RST2 as it was in RST1, and the reason for introducing the distinction has to do primarily with the fact that fear and anxiety may be directionally differentiated. To wit, fear occurs when one is actively avoiding or escaping from a threat, whereas anxiety occurs when one must approach or anticipate a threat or, more broadly, when one's goals are in conflict (Corr, 2004; Corr, DeYoung, & McNaughton, 2013; Smillie, 2008; Smillie, Loxton, & Avery, 2011). Moreover, BIS anxiety and FFFS panic respond differently to different anxiolytic drugs; those anxiolytics that also affect panic are said to have 'panicolytic' properties (Corr, DeYoung, & McNaughton, 2013; Gray & McNaughton, 2000; McNaughton & Corr, 2008b).

The RST2 BIS is associated with anxious affect (albeit no longer with fear), but it does not itself mediate responses to cues for punishment. Instead, it serves as a conflict detector and resolver, taking signals from both the BAS and the FFFS. When the latter enjoy a certain measure of co-activation, the BIS itself is activated. Specifically, the RST2 BIS resolves conflict among systems by recursively amplifying the weights of the conflicting goals' negative associations until one goal is the clear winner (Gray & McNaughton, 2000). The BIS has the same behavioral effects as it did in the original

RST, inhibiting ongoing behavior, increasing arousal and attention, and creating other behaviors (sometimes termed 'passive avoidance' or 'defensive approach') designed to prepare the organism to respond to threat (Corr, DeYoung, & McNaughton, 2013; Gray & McNaughton, 2000). Neurologically, the BIS is confined to limbic structures, including the septo-hippocampal system and the amygdala (Smillie, 2008).

The RST2 BIS and FFFS are typically conceived as linked yet separable systems. They are linked in the sense that both are thought to mediate avoidance processes, both are thought to contribute to 'punishment sensitivity' in human personality, and both are thought to relate to trait neuroticism or negative emotionality (Smillie, 2008). The two systems utilize the same pair of neurotransmitters, serotonin and noradrenalin, which are released when threatening stimuli are encountered (Smillie, 2008). Furthermore, both theory and evidence suggest that the BIS interacts with the FFFS, amplifying signals sent to the latter during goal conflict resolution and thus biasing affect, cognition, and behavior in favor of avoidance (Smillie, Loxton, & Avery, 2011).

1.3 Important (Albeit Neglected) Concepts in RST

Several ideas that Gray and McNaughton (2000) proposed are essential to a thorough understanding of Gray's theory, as they link all three RST2 systems. I discuss them here not only because they are so important, but also because they seem occasionally to be forgotten or underplayed in the literature.

1.3.1 Goals

Gray and McNaughton define a *goal* as a stimulus-response pairing; for example, 'the goal of the rat running down the runway has both a stimulus (in this case, place) component and a response (the animal's tendency to run toward it) or motivational component' (Gray & McNaughton, 2000, p.23). The authors further claim that 'the strength of (a goal's) activation is determined by integration of the prior affectively positive and affectively negative associations of that goal' (2000, p.28, my parentheses). Although Gray and McNaughton (2000) leave these associations' nature and the details of how 'integration' leads to goal activation largely unspecified, there seem to be a few safe, yet illuminating, conclusions that may be drawn from the little that *is* known with relative certainty.

First, a goal's 'approach' or 'avoidance' nature is determined by the direction in which it causes an animal to move in space and time relative to the goal-stimulus (a concrete place, an abstract end state, etc.). Approach goals bring the animal closer to the goal-stimulus (perhaps increasing goal activation), whereas avoidance (defensive) goals tend to increase the distance between the animal and the goal-stimulus (decreasing goal activation, and assuming all activations are treated as positive). Note that the idea of goal activation is an old one, and its relationship to distance makes it quite similar if not identical to Miller's (1959) 'goal gradient' construct.

The distance between an animal and a goal object has also been studied in some detail (Hull, 1932; Miller, 1959). Most recently, the Blanchards (1990b) have built a career around the idea of 'defensive distance,' a subjective, monotonically increasing (but not necessarily linear) function of proximity to an unpleasant stimulus. A complementary concept might be dubbed 'appetitive distance,' a subjective measure of reward proximity.

I will refer to appetitive and defensive distances collectively as 'goal distances.'

Goal distance influences goal activation, and the latter is probably inversely proportional to the former. However, consider the observation that defensive distance determines the specific behavioral response (fight, flight, freezing) that the FFFS will produce in a threatening situation (Blanchard & Blanchard, 1990a). Specifically, an animal flees if it can escape (the threat is relatively distant), freezes if it can't and the threat is distant, and fights if it can't escape (the threat is nearby). This example illustrates not only the importance of spatiotemporal distance in determining FFFS action selection (and, therefore, goal activation), but also of other factors, such as availability of escape (in this case, lower availability increases FFFS activation). Note that Gray and McNaughton (2000) refer to this factor as whether a threat is avoidable or unavoidable (e.g., see Fig. 11.2 on p. 295); if an actual threat is avoidable, an animal may flee or fight; if unavoidable, it may freeze or fight. BAS activation follows analogous principles, and RST system (goal) activation writ large depends on at least three factors: spatiotemporal distance, environmental affordances for approach/avoidance (which influence reward/punishment expectancies), and cue potency. All these factors themselves interact with reinforcement sensitivities to determine goal activation.

1.3.2 Conflicts and Their Resolution

Gray and McNaughton (2000) identify two distinct species of conflicts: goal conflicts and motor conflicts. *Goal conflicts* are conflicts between goals with incompatible response tendencies (e.g. between the tendencies to approach or avoid the same goal-stimulus, to approach one stimulus at the cost of not approaching another one, etc.). *Motor conflicts* are conflicts between different means of enacting the same response tendency (e.g. a

conflict between impulsive vs. controlled approach to the same goal-stimulus).

The BIS resolves goal conflicts, but not motor conflicts; the latter are resolved within the system, BAS or FFFS, that corresponds to the approach or avoidance nature of the goal. BAS and FFFS goals are more activated at lower appetitive and defensive distances, respectively. The BIS, which compares goals' activations in order to determine whether a goal conflict that it needs to resolve is taking place, is more active at higher goal activations *while goals are in conflict*. Past its activation threshold, the BIS is more active the greater the conflict is; that is, the more equally activated and the more highly activated the conflicting goals.

Any kind of goal conflict, if sufficiently intense, may activate the BIS; these include conflicts between the FFFS or BAS and itself as well as those between the FFFS and BAS (Corr, DeYoung, & McNaughton, 2013; Smillie, 2008). In order to resolve goal conflicts, Gray and McNaughton (2000, p.28) tell us that the BIS recursively adds weight to or excites additional negative associations linked to the conflicting goals, and that 'In many cases this will allow resolution of the conflict in favour of that goal which has the lesser negative associations.' Still, if only sufficiently intense goal conflicts activate the BIS, some goal conflicts must be resolved in other ways. What, then, is the BIS really good for?

1.3.3 Positivity Offset, Negativity Bias, and a Cross-Species Speculation

Two RST-related concepts; namely, positivity offset (Cacioppo, Gardner, & Berntson, 1997) and negativity bias (Miller, 1959), are helpful in better understanding the BIS. 'Positivity offset' is used to denote the principle that animals' behavioral or motivational approach systems have higher baseline activations than do their behavioral or

motivational avoidance systems. On the other hand, 'negativity bias' is the name given the principle that avoidance systems are more sensitive to input than are approach systems. Together, these principles predict that, in mixed-incentive conditions, behavior will be biased toward approach when goal stimuli are far away and toward avoidance when the latter are closer at hand.

Furthermore, the purpose of the BIS becomes clearer if one considers the ways in which RST systems likely differ across species. Two primary constituents of the BIS are the septo-hippocampal system and the amygdala (Gray & McNaughton, 2000, the latter is used by the BAS and FFFS, as well), which could roughly be considered cognitive and affective components of the BIS, respectively. Because the hippocampus is a structure found only in vertebrates, only vertebrates possess any neurophysiological system *akin to* Gray's BIS. Because the amygdala is a structure found only in complex vertebrates (e.g., mammals), only complex vertebrates possess a neurophysiological system *equivalent to* Gray's BIS. Nevertheless, animals that may lack a BIS probably still have a BAS and FFFS. Although studies of individual differences among lower vertebrates (e.g. fish) and invertebrates are relatively scarce, there are enough to demonstrate that these simple animals display interindividual differences in approach and avoidance tendencies (Gosling & John, 1999; Shearer & Pruitt, 2014).

Even extremely simple organisms (e.g. corals), which are more likely to be immobile, approach things and defend themselves after a fashion. Indeed, a body of any kind serves as a defense, even if it's nothing more than a layer of lipids. Consider, too, that the first self-reproducing molecules had no such defenses; they simply collected the chemical constituents necessary to build themselves anew (a basic form of approach). Defenses had to evolve and, from this perspective, approach may be considered more basic than

defense. The 'positivity offset' principle (Cacioppo et al., 1997), then, seems to be a quite understandable and general one. Given the indisputable selective success of simple 'BIS-less' species, it is clear that a BIS is not necessary for survival and reproduction.

A BIS is not even necessary for the resolution of goal conflicts. Among BIS-less species, goal and motor conflicts alike are probably resolved via nothing more than direct competition between (or, for motor conflicts, within) BAS and FFFS systems or, what's more neurally likely, between different activated goals (S-R pairings) that could be categorized as either Approach or Avoidance. Here, both positivity offset and negativity bias (Cacioppo et al., 1997) play essential roles; the latter ensures that defensive behaviors are more likely to win when threatening cues are present, enabling life to take precedence, so to speak, over lunch. But if BIS-less species are doing so well, why do so-called 'higher' species tend to have those systems that we label 'BIS'? Why is the BIS needed to detect and aid in the resolution of goal conflicts when other species perform the task so much more simply? If we can answer this question, then we should be well on our way to understanding the BIS and, by extension, RST as a whole.

Here's how we might approach the answer: BIS-less species, though capable of many complex behaviors, employ a reproductive strategy of 'quantity over quality' (E. O. Wilson, 1975), which renders behavioral mistakes less threatening to selective success, obviating the need for sophisticated learning capabilities that enable individuals to acquire complex behaviors within a single lifetime, rather than evolutionarily, across many lifetimes. Among species with a BIS, the septo-hippocampal system permits a special kind of learning: it enables organisms to alter their motor behavior in response to errors. Essentially, an error is both an expectation violation (one's action didn't bring about the outcome one expected) and a source of goal conflict (between the response that

initially caused the error and any new response that one enacts in hopes of correcting it). The BIS has often been characterized as an error corrector, a goal conflict detector, and a detector of expectation violations (Fua, Revelle, & Ortony, 2010; Gray & McNaughton, 2000), but the logical links among these characterizations are sometimes ignored.

Any cognitive apparatus capable (as, indeed, the septo-hippocampal system is) of serving as an error detector-and-corrector must in turn incorporate or interact with memory and expectation functions for stimulus/skeletomotor response pairs (see Morsella, 2005, for an explanation of why I refer only to 'skeletomotor' responses). These functions have the potential to make cognition vastly more sophisticated, as they undergird networks of S-R associations that make it possible for cues to signal much more spatially and temporally distant rewards and punishments (Ortony, Clore, & Collins, 1988). This signalling-at-a-distance introduces an element of uncertainty into goal pursuit, which in turn creates a need for additional cognitive complexity.

Both greater cognitive complexity and greater uncertainty increase the number of goals that an organism could pursue at a given time; in addition, greater cognitive complexity increases the organism's ability to deal intelligently with threats. Thus, BAS-related positivity offset is effectively increased and FFFS-related negativity bias is effectively decreased among such species. These factors in turn increase the probability of goal conflict and produce the need for an additional 'negativity bias system' that will settle conflicts between goals that simple inhibitory opponent processes would take too long to resolve (opponent processes are, however, probably still used to resolve motor and less evenly-matched goal conflicts even among creatures with a BIS).

Chapter 2

Literature Review

> In our daily lives we are confronted at every turn with systems whose internal mechanisms are not fully open to inspection, and which must be treated by the methods appropriate to the Black Box.
>
> W. Ross Ashby, 1956 (An Introduction to Cybernetics)

2.1 Empirical Support for RST

Despite its firm foundation in experimental research, multiple literature reviews have noted that empirical support for RST (revised or not) in humans is mixed at best (Corr, 2004; Matthews & Gilliland, 1999; Pickering et al., 1997). Theories that receive mixed support are unlikely to be embraced en masse; indeed, the reasons for RST's equivocal standing are probably related to several other factors that preclude its universal adoption. Pickering (1997) states that the results of RST research fall into one of four categories; namely, appropriate trait effects, complementary trait effects, inverse trait effects, and null effects. Appropriate trait effects are those that are unambiguously predicted by RST;

e.g., individuals high in reward sensitivity learn best when they are rewarded. The same outcome as is predicted for the 'appropriate trait,' if caused by a low level of the opposing trait, is a complementary trait effect; e.g., individuals low in punishment sensitivity learn best when they are rewarded. Inverse trait effects occur when the 'appropriate trait' results in an outcome opposite to that which RST predicts; e.g., individuals high in reward sensitivity learn best when they are punished. Null effects are null results.

The reasons for RST's unimpressive empirical performance almost certainly boil down to the fact that, as Pickering (2008) remarks, 'RST has always been a complex theory, although it masquerades as a simple one' (p. 455). From the very beginning, RST has been a more dynamic, mechanistic theory than top-down competitors like the Big Five. Before either RST1 (1982) or RST2 (2000) were fully-fledged, Gray and Smith (1969) were depicting the theory in circuit diagrams described by multiple differential equations. Such models present conceptual challenges for even the most nimble of unaided human brains, and the fact that people tend to favor energy-efficient strategies of all kinds over those that are more effortful and accurate only increases the probability of conceptual carelessness. In particular, the omission of even a few seemingly insignificant details accidentally or for brevity's sake can produce far-reaching consequences.

For instance, Corr (2004) reminds readers that even the earliest incarnations of Gray's RST1 allowed the punishment system (BIS) and reward system (BAS) to mutually inhibit one another (Gray & Smith, 1969), a feature that he refers to as the Joint Subsystems Hypothesis (JSH). He further notes that, at least since the publication of Gray and McNaughton's revision (2000), some RST scholars seem to be laboring under an alternative assumption, which he calls the Separable Subsystems Hypothesis (SSH); i.e. that the reward and punishment systems function entirely independently. Corr (2004)

believes that researchers' failure to consider the consequences of the JSH accounts for a number of apparently discrepant effects in the RST literature. Unfortunately, the evidence in favor of the JSH is as mixed as is the evidence in favor of RST itself, and effects consistent with the SSH are insufficient to rule out the JSH, since the latter makes allowance for systemic decoupling under certain conditions (albeit not those that typify human-subjects laboratory environments). Corr (2004) himself admits that demonstrating the validity of the JSH is likely to be a delicate task, and, as will be discussed in greater detail below, he may not have defined the problem entirely correctly (Pickering, 2008).

The difficulties (of which the JSH issue is only one example) surrounding RST's key concepts and operational definitions, especially as they apply to humans, are likely the source of many if not all of its other difficulties; e.g., disagreements regarding its applicability to humans (Matthews, 2008), partially successful efforts to integrate it with five-factor and other top-down models (Mitchell et al., 2007; Poropat & Corr, 2015; Van Egeren, 2009), and, perhaps most notably, failed (Torrubia et al., 2001; G. D. Wilson, Gray, & Barrett, 1990) or misguided attempts to create a 'gold-standard' self-report measure of reinforcement-sensitivity-related personality traits (Smillie, 2008).

2.2 Measuring RST

RST's self-report measures are among its most substantial shortcomings. Although the concerns regarding RST's inadequate coverage of human personality, which reflect the bandwidth-fidelity dilemma (Cronbach & Gleser, 1965), might apply even if the theory were to produce a successful self-report measure, such an achievement would likely put RST on better footing as a competitor with five-factor models, especially given that the

latter appear neither to be as real nor as uniquely useful as many suppose (Condon, 2014; Yarkoni, 2013). A few 'high fidelity,' bottom-up traits like those that RST describes (i.e. sensitivity to cues for reward, punishment, and conflict) might be preferable in self-report assessment to a larger number of potentially less valid, top-down traits. Moreover, high-fidelity RST traits could be used to set the axes of rotation for high-bandwidth factor analyses. In fact, this application seems to have been the impetus behind Gray's (1970) suggestion that Eysenck rotate his trait axes by 30-45 degrees.

Given that most measurement and structural equation models are based on self-report measures, 'constructs' and 'questionnaires' have become nearly synonymous in contemporary human personality research, which is necessarily limited to observational methods. Because self-report measures are cheap and easy to administer, the latter have become a preferred means of data collection as well as of communicating the substance of many if not all personality theories, top-down or otherwise. Any personality theory unaccompanied by a statistically sound questionnaire is empirically disadvantaged and, as others have noted, RST's self-report measures tend to be both inconsistent (Torrubia et al., 2001) and psychometrically suspect (Smillie, 2008).

2.3 Dynamic Processes in RST

Even if collaboration and construction were as fruitful and rigorous as they could possibly be, one idea would remain that every RST2 scale-development team seems to have ignored. Specifically, RST2 scale developers stand to profit by acknowledging that there is a difference between a bottom-up *reinforcement sensitivity* and a top-down *reinforcement-sensitivity-related personality trait* (Pickering, 2008). Reinforcement sensitivities are neurophysiological parameters that interact dynamically within persons

over time so as to give rise to relatively stable, between-persons, RST-related traits. Reinforcement sensitivity parameters are independent, yet interacting, internal factors (whose degrees of stability may indeed be traitlike) that are not available to introspection. On the other hand, the between-persons traits that such interactions produce are easily self-reported and often non-independent.

It is for this reason, among others, that Smillie (2008) advocates a move away from questionnaires and toward more direct measures of reinforcement sensitivity, such as those that employ psychogenomic, psychopharmacologic, neuroimaging, and behavioral paradigms. However, when Smillie (2008) claims that it may be impossible to self-report reinforcement sensitivities, he does not necessarily mean that it is impossible to create what we might call 'a good self-report measure of RST.' Instead, he is merely agreeing with Pickering (2008) that a *reinforcement sensitivity* is likely some set of neurophysiological parameters which interact with other parameters and on which it is unreasonable to expect individuals to be capable of introspecting. On the other hand, a *reinforcement sensitivity trait* assessed by a self-report inventory will tend only to *indirectly* reflect these internal parameters; instead, such a measure taps the 'functional outcomes of system activation' (Pickering, 2008, p. 459) rather than the systems' sensitivities themselves. Thus, traits like Neuroticism or Extraversion cannot be *identified with* reinforcement sensitivities operationalized as genotypes or brain structures; at most, the latter 'may manifest as or partly underlie' the former (Smillie, 2008, p. 367).

Although it is theoretically possible to create a self-report measure that could uncover individual differences in reinforcement sensitivity parameters, we are a long way from having such a measure. What would be more useful still would be to determine the precise manner in which RST parameters give rise to self-reportable, RST-related traits.

To this end, it is illuminating to survey research that attempts to link RST to top-down theories, like the Big Five, that are associated with popular high-bandwidth, top-down trait measures. Here again, though, we run into the aforementioned mysteriously missing link between top-down and bottom-up theories – a link that Poropat and Corr (2015) argue won't be found unless personality researchers model *all* sources of method variance in their research. These sources have to do not only with the individuals assessed (targets), but also with those doing the assessing (judges).

Each target and judge are affected by three factors: process, situation (which, for the judge, includes the target's observed behavior), and outcome. 'Process' refers to endophenotypic personality processes like those that arise from reinforcement sensitivities (Poropat & Corr, 2015, p. 65, Equation (4)); therefore, in order to develop a model that explains how bottom-up factors like reinforcement sensitivities give rise to top-down phenomena like Big Five or RST-related self-report traits, scholars need to focus on the *time-dependent processes* that give rise to targets' behaviors and judges' assessments. Such a redirection of attention is sorely needed, as intraindividual variability and personality dynamics are often neglected in both the top-down and bottom-up literatures.

Understanding the manner in which dynamic intraindividual differences give rise to static interindividual differences may lessen the probability that researchers will make several common conceptual errors. These are analogous to the geneticist's OGOSH (one-gene-one-system-hypothesis) fallacy, which neglects the fact that even small physiological systems are typically controlled by many genes. RST scale developers (and others stymied in their efforts to integrate bottom-up and top-down models) often commit what might be called the OBOSH (one-behavior-one-system) or OBOTH (one-behavior-one-trait) fallacy, when it is likely that it takes multiple reinforcement

sensitivities (or traits) to create a single behavior. Indeed, this is the point that Corr (2004) makes when he argues in favor of the Joint Subsystems Hypothesis, although, as Pickering (2008) points out, Corr (probably incorrectly) assumes that self-reported RST traits directly, rather than indirectly, reflect reinforcement sensitivity parameters.

If different combinations of RST parameters can give rise to the same behavior, then they could conceivably also give rise to the same level of self-reported RST trait. This possibility recalls Gray's (1970) suggestion that, given the rotation of RST traits relative to Eysenckian traits, Extraversion scores could be calculated by subtracting punishment sensitivity (Anxiety) scores from reward sensitivity (Impulsivity) scores. For instance, an individual who scored a standardized "1" on Extraversion could do so either by having Anxiety = 1 and Impulsivity = 2 or by having Anxiety = 0 and Impulsivity = 1. Even more interesting, however, than the possibility that the same behavior or trait may be created differently in different people, is the equally strong possibility that different behaviors or traits could, given different learning histories or different stages within a single history, result from identical sets of reinforcement sensitivities. Here, dynamics are directly responsible for the disparities; e.g., a certain individual may initially engage in impulsive approach, but later learn to favor more controlled, cautious forms of approach.

That multiple parameter combinations may result in the same behavior (or trait) or that the same combination may result in multiple behaviors (or traits) are not ideas with which psychologists are unfamiliar; they are, in fact, the old developmental concepts 'equifinality' and 'multipotentiality,' respectively. However, while such concepts may be comprehensible enough in the abstract, they are often devastatingly difficult to apply, especially when a theory has as many dynamically-interacting features (i.e. 'moving parts') as RST does. What RST plainly needs is a research program that, concordant

with Poropat and Corr's suggestion (2015), focuses not only on bottom-up structures or parameters, but also on the dynamic, mechanistic processes in which the latter are involved. In other words, if we want a grand unified theory of personality (or, indeed, all of psychology), then we need a precise model of the intraindividual processes that *cause* individual differences of all kinds.

2.4 Personality Dynamics and Computational Modeling

Of course, experimental research is the touchstone for causal knowledge in psychology, but because the in vivo study of human personality is restricted to using correlational rather than experimental methods, researchers have devised a number of 'shortcut' techniques for inferring causal effects where they cannot be conclusively proven. These methods fall into three broad categories, analogical, empirical, and computational, which I will review briefly in ascending order of relevance.

2.4.1 Analogical Methods

The simplest way to lend one's correlational data theoretical heft is to tell a good story about them. I refer to this kind of inference as 'analogical' because one's data must 'look like' they fit the story one tells, just as the pairs of relationships in an analogy share certain features and therefore 'look like' one another. However, just as analogical pairs look alike but need not be (and in practice never are) exactly the same, the stories one tells to explain one's data may fit or look like the data but need not be true. There are at least three kinds of storytelling that personality psychologists employ or study: (1) inference based on personal narratives (e.g., McAdams, 1995; McAdams & Pals, 2006), (2) inference based on evolutionary psychology (e.g., Nettle, 2006), and (3) inference

based on comparative psychology (e.g., Gosling & John, 1999).

Personal narratives seem to describe the causes of events in narrators' lives, but they are rarely studied with the intent to deduce the strength of their causal claims. Evolutionary psychology reminds us that the selection pressures that shaped the human brain are essential to a full causal explanation of its structure and function; however, this insight provides no information about process mechanisms and is, moreover, a weak means of inferring causality because knowledge of human evolution is itself constantly evolving and, worse, is not subject to rigorous falsification. Though in some ways similar to evolutionary psychology, comparative psychology takes a cross-sectional, cross-species, rather than a longitudinal, within-species approach and is couched in data that are easier to falsify than those which inform evolutionary theorizing. Comparative theorizing has been used in concert with Eysenck's model (1981; 1997) and the Big Five (Nettle & Penke, 2010) as well as in arguments justifying RST's generalization to human personality (McNaughton & Corr, 2008a).

As compelling as the stories that analogical inference creates may be, however, they are still *only stories*. As such, any story one may choose to tell (a) might be incorrect; (b) is likely to be influenced by 'common sense' or folk wisdom, and therefore is at best uninformative and at worst a means of fostering theoretical complacency or reinforcing stereotypes; and (c) still doesn't say much about the precise mechanisms of change, even if it is true. The third problem is likely the most serious; knowledge of mechanisms is exactly what is needed if one hopes to understand the manner in which top-down and bottom-up theories interlock. For these reasons, analogical methods of inferring causality are best used to supplement empirical or computational techniques.

2.4.2 Empirical Methods

Many personality psychologists use statistical techniques and specially-designed research studies to infer personality causality. These methods seem to fall into seven broad categories: (1) simple correlation-based causal inference, (2) factor analysis, (3) studies of genetics and heritability (Johnson, 2007), (4) studies of the effects of psychopharmacologic agents on behavior, (5) neuroimaging studies, (6) longitudinal techniques like growth-mixture modeling (Leszko, Elleman, Bastarache, Graham, & Mroczek, 2016), and (7) studies of within-person variability (Wilt et al., 2011). The first five of these don't really provide much stronger evidence for causality than the analogical methods, as all such studies provide correlations extracted from data taken during a few circumscribed sessions over a relatively short period of time; correlations encourage storytelling. Of these five, psychopharmacologic studies probably provide the strongest means of inferring causality, as drugs may be experimentally administered. Factor analysis, the weakest of the five methods, in fact provides no grounds for inferring causality (Revelle, 1983), but it is included here because it is often treated as if it did.

Longitudinal research, which many acknowledge as a gold standard for non-experimental investigations of psychological causality, has enabled scholars to answer a number of important questions about personality, including 'does it change?' (Mroczek & Spiro, 2003), 'how does it change (on average)?' (Roberts, Walton, & Viechtbauer, 2006), and 'does that change follow different groupings of developmental trajectories?' (Leszko et al., 2016). Nevertheless, longitudinal data often fail to supply precise, mechanistic explanations for the changes they document, no doubt because they are insufficiently granular; a great deal more information is required in order to move from describing between-persons patterns to within-persons processes. Indeed, those who study

intraindividual variability (i.e., the manner in which an individual's personality varies relative to itself) do not conduct long-term longitudinal studies precisely because within-persons personality research generates an enormous amount of data even when the studies are brief. Such studies typically employ a simple, inexpensive technique called 'experience sampling' which prompts participants to contribute data at regular intervals throughout the day, over a period of several days (Fleeson, 2007b).

Intraindividual differences research is potentially the most useful of the seven empirical methods for investigating personality processes and reconciling top-down and bottom-up theories of personality. However, most such research to date has emphasized a different reconciliation; namely, that of between-persons personality traits and within-persons personality states, both of which are top-down constructs. At least two theories purport to explain traits' relation to states: Fleeson's (2004) 'state-trait isomorphism' postulates that traits are the time-averages of states, whereas Revelle and Condon's CTA theory (2015) holds that traits are the relatively stable rates at which tendencies to behave in certain ways change. Note here that longitudinal studies and CTA both consider time a variable; state-trait isomorphism does not.

Although within-person variability data may lead to the development of a 'state-trait calculus' that could articulate the manner in which states lead either to mean trait levels or to trait change, three problems remain for those who would use it to reconcile bottom-up and top-down models. First, such studies would have to be prohibitively lengthy and data-rich in order to document important processes like trait development and change. Second, understanding the bottom-up mechanisms that support even brief processes is difficult given the current state of experience-sampling research, which tends to collect self-report data and not much else. Third, there's still the nagging question of

causality: even if the study is brief, and even if we were able to collect data on processes about which it is effectively impossible to introspect, we still could not make a strong causal claim unless we were able to randomly assign different personality parameters to different individuals. With human beings, this simply cannot be done.

2.4.3 Computational Methods

Directly manipulating computer-simulated personality traits is the closest researchers are likely to get anytime soon to directly manipulating human personality traits. Although by no means a widely-used method, computer simulations are slowly gaining acceptance as a research tool in the social sciences (e.g. Read et al., 2010). Broadly speaking, of course, the same criticism applies to computer simulations as to factor-analytically derived measures of human behavior; i.e. that 'garbage in equals garbage out.' Nevertheless, even imperfect simulations improve upon questionnaires as aids to theory-building because the former allow us more easily to see personality as a dynamic process. Below, I will summarize several simulation efforts that seem particularly pertinent to informing future computational models of RST.

2.5 Indirectly-Related Computational Models

A number of theories regarding intraindividual personality dynamics have been proposed over the years. Several of these, while promising, have yet to be developed into explicit computational models. For instance, Kuhl's (2000) Personality Systems Interaction theory posits that personality and motivated behavior arise from energy flow among four basic systems: object recognition, intuitive behavior control, feeling, and thinking. Alternatively, Cervone's Knowledge and Appraisal Personality Architecture model (2004)

suggests that individual differences in behavioral consistency across situations arise from differences in knowledge structures and situational appraisals. Both models favor social-cognitive explanations of personality and focus little on the possible ways in which intraindividual variation might give rise to interindividual variation.

In fact, even those models of personality dynamics that have been translated into computational simulations often do not explore one or the other of the between-persons/within-persons poles. Unsurprisingly, most such models have been developed outside the RST tradition; nevertheless, I will briefly consider three that I believe are of interest. Specifically, Shoda and Mischel's CAPS simulation (1998), Poznanski and Thagard's SPOT model (2005), and Quek and Moskowitz's neural network simulation of workplace behavior (2007) have some bearing on problems related to those I will be tackling with my own simulation.

2.5.1 Shoda and Mischel's Models

The Cognitive-Affective Process System Model (CAPS, Mischel & Shoda, 1995; Shoda & Mischel, 1998) was one of the first neural network models used to understand personality dynamics. To stand in for personality, the model uses a network of cognitive-affective units (goals, plans, expectancies) which receive situational input and produce behavioral output. These goal-based units are quite loosely organized, inasmuch as they are in no way parsed or bundled into any structure that might represent a superordinate (neural, motivational, etc.) system. Moreover, as might be expected from a team whose PI is notoriously critical of the trait construct, Shoda and Mischel (1998) do not claim that their goal-based units comprise personality *traits*. Thus, CAPS does not capture interindividual personality structure and cannot be used to infer the manner in which

between-persons traits arise from individuals' ever-changing goals, states, etc.

2.5.2 Poznanski and Thagard's Model

In addition to providing students of personality with a new tool, the Simulating Personality Over Time model (SPOT, Poznanski & Thagard, 2005) serves as a platform for creating personality in video game characters. These agents' personalities are simulated via a set of seven baseline activations for personality nodes, three of which correspond to a Big Five trait and four of which correspond to trait poles (Introversion, Extraversion, Agreeableness, Disagreeablenss). The model includes six kinds of nodes in all; namely, situational, relationship, mood, emotion, personality, and behavior nodes. Behavior nodes, which represent the model's outputs, are affected by situation, relationship, emotion, and personality nodes. Emotion nodes, the only nodes to both influence and be influenced by other nodes, receive inputs from situation, relationship, mood, and personality nodes. Although the authors developed 'rules' governing personality change within the model (essentially, hard-coded connection weights in the neural net), the latter are not particularly realistic. Additionally, most non-personality nodes are characterized rather arbitrarily; for example, Mood can be *good*, *bad*, or *neutral*; Relationship can be *like* or *dislike*; Behaviors include *persist*, *assault*, *cry*, *help*, etc. Although sample simulations show that the model can replicate some familiar results (e.g., different people do different things in the same situation or behave in ways that are consistent with their traits) and mimic personality change in accordance with its circumscribed set of rules, it still does not provide meaningful insight into personality structure or dynamics.

2.5.3 Quek and Moskowitz's Model

Quek and Moskowitz's Workplace Roles model (2007) is a three-layer (input, hidden, output) feed-forward neural network that the authors used to simulate two archival datasets pertaining to professional roles and behaviors. In the first study, the network learned the relationship between work role (i.e., the network's input nodes were set to *boss*, *coworker*, and *supervisee*) and either *dominant* or *submissive* behavior (output node settings). In the second study, the network replicated results in the literature regarding the relationship between gender and communal behavior in different types of relationships. Although the simulations certainly showed what they were intended to show, the Quek and Moskowitz (2007) network modeled only a limited number of situations and goals. Nevertheless, this is the only computational model of individual differences and behavior I was able to find that modeled data from real studies (Moskowitz, Suh, & Desaulniers, 1994; Suh, Moskowitz, Fournier, & Zuroff, 2004).

2.6 Computational Models of RST

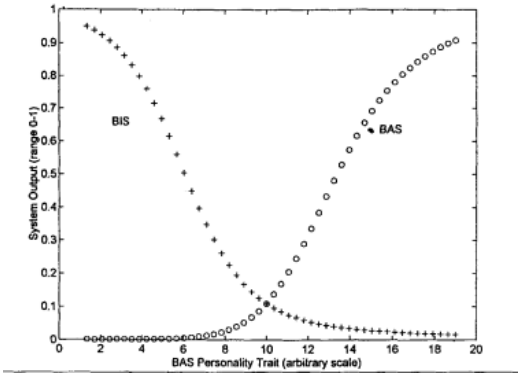
2.6.1 Pickering's Models

The foremost producer of purpose-built computational models of RST has been Alan Pickering, a student of Jeffrey Gray's (Pickering, 1997, 2008; Smillie, Pickering, & Jackson, 2006). Even before Corr (2004) began championing the Joint Subsystems Hypothesis, Pickering (1997) had made the claim that apparently discrepant RST research findings could, in fact, still arise even if the theory as specified by Gray and Smith (1969) was correct in all its particulars. Pickering argued further that computational models could aid in prediction-making, given that RST is a deceptively

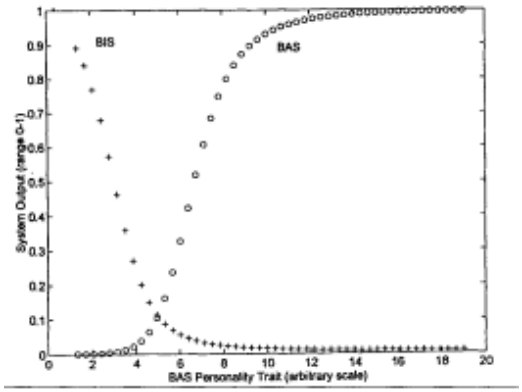
simple theory that often tricks researchers into thinking that they've devised hypotheses in line with its tenets when, in fact, they haven't.

Pickering (1997) made a compelling case for his claim when he used two slightly different computational models (see Figures 2.3a and 2.3b) to replicate various 'unexpected' experimental outcomes, including a null effect (no correlation between reward sensitivity and change in BAS output as a result of reward) and an inverse trait effect (a negative correlation between reward sensitivity and change in BAS output as a result of reward). In order to obtain the null effect, Pickering randomly selected the BAS and BIS sensitivities for each subject from independent normal distributions (M = 10, SD = 3), then randomly assigned subjects to either a 'control condition' (where BAS cue = FFFS cue = 1.0; see Figure 2.1a) or a 'reward condition' (where BAS cue = 2.0 and FFFS cue = 1.0; see Figure 2.1b). Participants' BAS output in the control condition was subtracted from their BAS output in the reward condition to obtain a dependent measure of 'change in learning'. Plotting this change as a function of BAS sensitivity resulted in a curvilinear effect in which participants with medium BAS sensitivities exhibited the greatest change in learning in response to reward (Figure 2.1c); this inverted-U-shaped curve translated into a nonsignificant linear correlation or, when simulated subjects were subjected to a median split in BAS sensitivity, a nonexistent interaction between sensitivity and condition (Figure 2.1d).

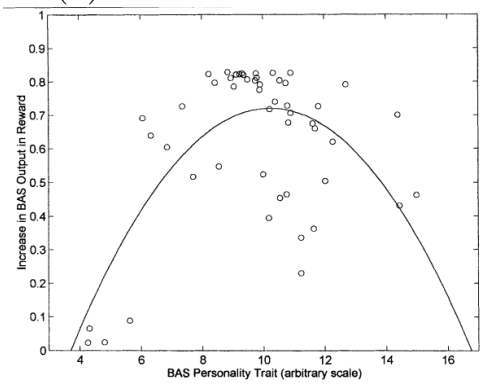
Essentially, this curvilinear relationship between BAS sensitivity and change in BAS output is a by-product of the sigmoid functions that govern this model's BAS and BIS system outputs, a result that echoes Humphreys and Revelle's (1984) observation regarding the derivation of curvilinear effects from opposing monotonic processes. For instance, when BIS sensitivity is set at a constant, medium value; cues for reward and



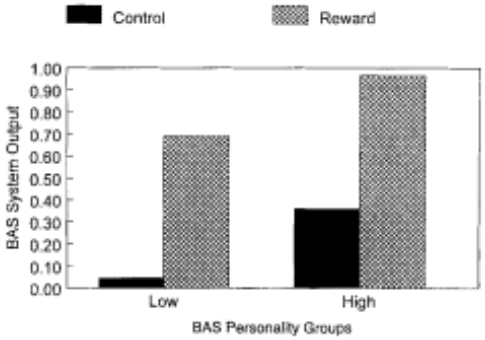
(a) Null Control: Rew = Pun = 1



(b) Null Reward: Rew = 2, Pun = 1



(c) Null Effect: BAS change v. BAS weight

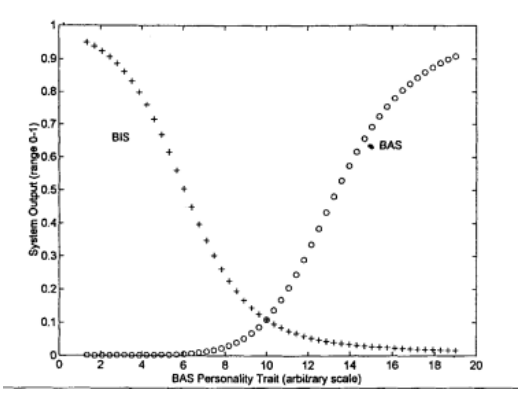


(d) Function in (c) as a median split

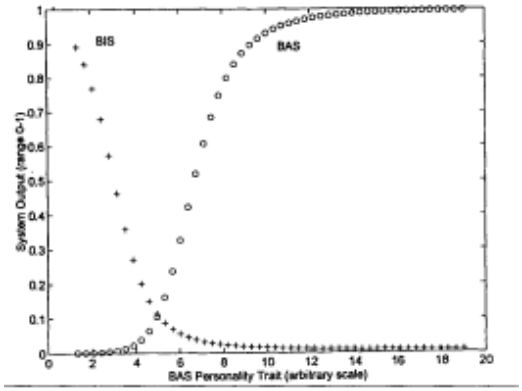
Figure 2.1: Adapted from Pickering, 1997: Simulating a Null Effect

punishment are equal; and BAS sensitivity is varied from low to high (as in Figure 2.1a), BIS output is a decreasing and BAS output an increasing sigmoid function of that sensitivity. The two functions intersect at a medium BAS sensitivity when cues for reward and punishment are equal, but when cues for reward are increased, the functions become steeper and their point of intersection shifts to the left (as the graph in Figure 2.1b does relative to the graph in Figure 2.1a). As a result, change in BAS output is greatest for those with moderate levels of reward sensitivity. Pickering's (1997) inverse

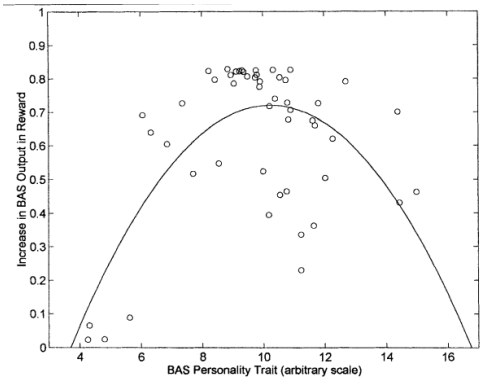
trait effect was due to the same principle. In this case, however, a negative correlation between BAS sensitivity and change in learning in response to reward was obtained when the intensities of all rewarding and punishing stimuli in the Control and Reward conditions were doubled, pushing the inverted-U-shaped curve further to the left (see Figures 2.2c and 2.2d).



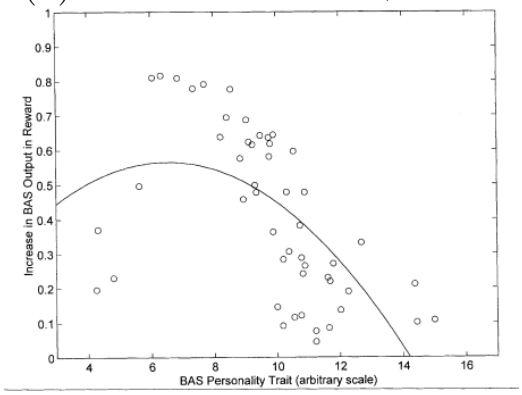
(a) Inverse Control: Rew = Pun = 2



(b) Inverse Reward: Rew = 4, Pun = 2



(c) Null Effect: BAS change v. BAS weight



(d) Inverse Effect: BAS change v. BAS weight

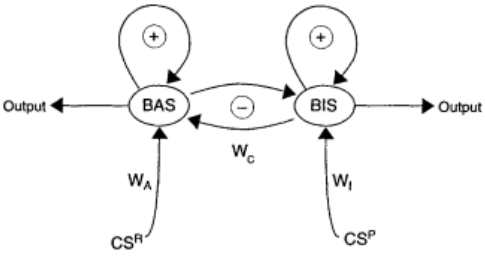
Figure 2.2: Adapted from Pickering, 1997: Simulating an Inverse Effect

Using an 'arousal-only' model of RST1 (i.e. BAS and BIS do not directly activate goals,

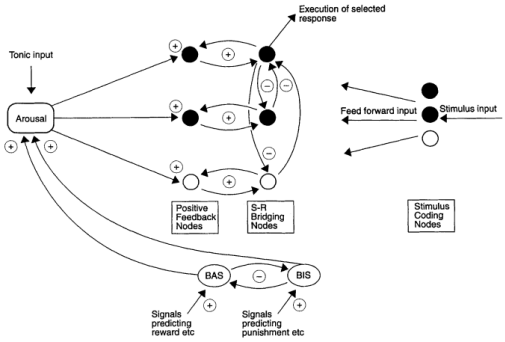
but rather act on an arousal system, which activates goals; see Figure 2.3b), Pickering (1997) was also able to produce a complementary trait effect (those low in BIS sensitivity exhibited greater gains in learning in response to reward than those high in BIS sensitivity; see Figure 2.3c), due to the fact that both BIS and BAS activation contribute in the arousal-only model to goal activation. Unfortunately, however, Pickering's complementary trait effects were always accompanied by appropriate trait effects which, as Corr (2004) noted, rarely happens in real life. Again, an example of an appropriate trait effect would be higher reactivity to reward in participants high in BAS sensitivity, as shown in Figure 2.3d.

Later, Pickering (2004) retooled the original (non-arousal-only) simulation, bringing it into agreement with RST2. The reciprocal inhibitory connections between reward- and punishment-sensitivity systems proposed by Gray and Smith (1969) were retained, but the FFFS became the punishment-sensitivity system and the BIS was rewired to receive excitatory inputs from BAS and FFFS and to inhibit the BAS while exciting the FFFS. Smillie et al. (2006) used the revised model to support Corr's (2004) claims regarding joint effects of RST systems under mixed-incentive laboratory conditions, while making it clear that the simulation's results reflected the 'functional outcomes' of the system in question (in this case, the BAS), rather than the system's sensitivity per se.

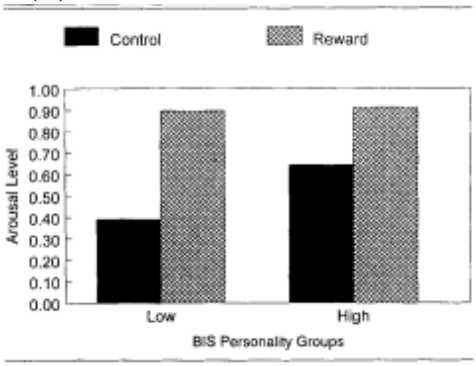
Specifically, simulated participants with randomly-drawn BAS, BIS, and FFFS sensitivities were each subjected to 200 simulated 'events' (pairs of randomly-drawn reward and punishment stimuli). Opponent processes between the BAS and FFFS were modeled using a Grossberg (1976) shunting inhibition equation. A full description of the model's details may be found in the article's appendix (Smillie et al., 2006). The final calculated outcome of each situation was influenced jointly by BAS, BIS, and FFFS,



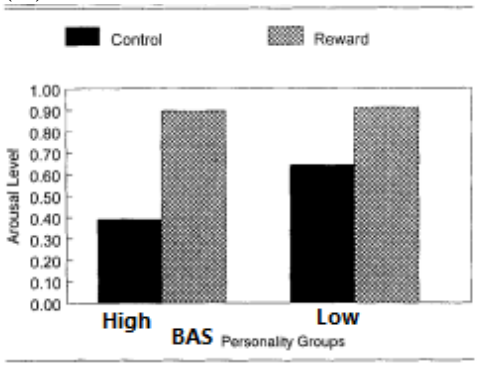
(a) Pickering, 1997: Standard Model



(b) Pickering, 1997: Arousal-Only Model



(c) Complementary Trait Effect



(d) Appropriate Trait Effect

Figure 2.3: Adapted from Pickering, 1997: Complementary vs. Appropriate Trait Effects

resulting in outputs for each that could be construed as emotion, self-reported trait scores, etc. A strong negative correlation (-.53) between reward-trait (BAS) output and punishment-trait (FFFS) output was observed; correlations between the BIS and BAS, and between the FFFS and BIS, were significant and positive, albeit somewhat weaker.

Pickering's RST2 simulation reappeared (in a slightly altered technically but essentially equivalent guise) in a (2008) book chapter. Here, Pickering added a statistical translation of his simulation, dubbing the former a 'decision model' and using it to demonstrate the differences between two-system and three-system models of RST (adding a BIS decreases

the likelihood that the BAS will gain control of behavior, with the effect being greatest for those with moderate reward sensitivities). Pickering also showed that, given environments containing equal reward and punishment cues, maximal input to the BIS (or conflict resolution system) occurs in individuals with moderate reward sensitivities. As the intensity of reward and punishment cues increases, so does the input to the BIS. Finally, BIS activation is a nonlinear function of the probability of the BAS gaining control of behavior; activation increases when the probability of approach behavior is low, then decreases past a probability of 25% or so.

Importantly, Pickering (2008) is quite clear that he no longer believes, as he stated previously (1997), that the sensitivity weights on the model's systems can be construed as the systems' 'functional outcomes' (e.g., responses on self-report measures). It is here that Pickering differentiates most clearly between reinforcement sensitivity parameters and the traits (or other outcomes) to which they, in concert with other factors, give rise. The chapter's simulations often refer to system output as 'probability of reward or punishment system control of behavior,' but this is not the only possible functional outcome that Pickering discusses. Others include reward/punishment-related arousal system output, reward/punishment-related behavioral intensity, and reward/punishment system activation. Once again, Pickering (2008) shows that reward system and punishment system output should be negatively correlated.

In addition to the aforementioned models, Pickering's team has generated other simulations which are less directly related, but still relevant, to modeling RST as a whole. For instance, Pickering and Pesola (2014) simulated feedback related negativity (FRN), a dopaminergic EEG signature that acts as an index of the degree to which reward magnitude conforms to expectations. FRN is most negative in response to unpredicted

nonreward; it is least negative in response to unpredicted reward. In general, extraverts or individuals with high self-report BAS scores react more strongly (negatively) to unpredicted nonreward and less strongly negatively (or more strongly positively) to unpredicted reward than introverts or individuals with low self-report BAS scores (Cooper, Duke, Pickering, & Smillie, 2015; Smillie, Cooper, & Pickering, 2011). Pickering is currently working on a model of RST2 that is rather different from the one that featured in his earlier papers; among other differences, this model places mutually inhibitory connections between potential FFFS and BAS outputs, rather than between the systems themselves. It is intended to demonstrate that the BIS is the system responsible for risk aversion in prospect theory as described by Kahneman and Tversky (1979) and demonstrated computationally by Tom, Fox, Trepel, and Poldrack (2007).

2.6.2 Corr and McNaughton's Model

Not to be outdone, Corr and McNaughton (2008) presented their own computational model of RST2 in the same book as Pickering (2008) did. It is also based on Gray and Smith's (1969) original blueprint and is in many ways very similar to Pickering's RST2 models. Sensitivity weights similar to Pickering's translate information about external rewarding and punishing stimuli into internal representations thereof and, in a departure from the Gray and Smith template, the mutual inhibition between reward and punishment systems is achieved by establishing feed-forward inhibitory links from rewarding or punishing external stimuli to punishing or rewarding internal representations of stimuli, respectively.

The Corr and McNaughton (2008) model's Arousal (sum of internal reward and punishment activation) and Decision Mechanism (difference between internal reward and

punishment activation) functions make it more similar to the Gray and Smith (1969) model than Pickering's (2008) is, but, unlike Gray and Smith (1969), Corr and McNaughton (2008) do not allow internal representation activations to take negative values. BIS activation is lower when the two internal representations' activations differ more and greater the greater the conflicting activations' sum; BIS activation increases arousal and inhibits the decision mechanism. The arousal and decision systems' activations are multiplied in order to determine observed behavioral output. Corr and McNaughton (2008) use a series of simulations in which reward and punishment input and reward and punishment sensitivity are varied to show that the effect on behavior of losing the BIS is to increase the amount of reward-related behavioral output at intermediate to high levels of punishment input.

2.6.3 Interim Summary

Ultimately, I selected Pickering's model rather than Corr and McNaughton's to provide my own simulation with a firm grounding in RST's central tenets. My reason for choosing as I did was that computational models of RST are scarce, and Pickering has been developing them longer than Corr and McNaughton have; ultimately, I elected to follow in the more experienced research team's footsteps. Note here, too, that no previous computational models of RST have simulated data from specific studies; the most they have been used to do is generate data that are consistent with those collected in previous 'real' research. I, however, intend to use my model to simulate data from three real studies.

2.7 Read's Computational Model of Personality Dynamics

Read et al.'s (2010) ambitious neural network model is one of the most recent in a series to emerge from the Read and Miller labs (Read & Miller, 2002; Read et al., 2006). Although not designed to agree with the theory in all particulars, the model bears a close resemblance to RST2. In fact, Read et al. (2010) used several lines of research to inform the network's construction. Along with Gray and McNaughton's (2000) RST2, Read et al. (2010) drew upon research in the lexical, factor-analytic tradition (especially on that which supports five-factor trait models); upon personality theories emphasizing temperament and biological underpinnings of behavior (e.g., those of Clark & Watson, 1999; Cloninger, 1987; Zuckerman, 2002); upon goal-based models developed by Read and Miller (1989) and Shoda and Mischel (1998); and upon works in evolutionary psychology and affective neuroscience regarding human motivation (Fiske, 1992; Panksepp, 1998).

Like RST2, the Read et al. (2010) model includes three major systems: one that governs approach motivation, one that governs avoidance motivation, and one that governs general inhibition and accounts for individual differences in constraint (because it is not a variable-calculator or 'layer,' the authors refer to the latter as a 'process' rather than a 'system'). On the face of it, these systems sound like the RST2 BAS, FFFS, and BIS, respectively; however, Read et al. (2010) take a position that is perhaps closer to that of RST1 and liken their avoidance system to BIS, not FFFS. Their approach system is likened to BAS.

In addition to the three primary systems, the Read et al. model includes specific goals or motives that the approach and avoidance motivation systems influence. Goals include finding mates, establishing and maintaining good hierarchical and communal social

relationships, and avoiding physical harm; all motives fall into 'communal' or 'agentic' categories. The network's parameters (i.e. strengths and sensitivities of the three major systems as well as specific goals' baseline activation levels) can be varied to simulate human personality traits. Situational features (where the simulated person is, who it's with, what it's doing) and individual resources (e.g. money, adaptive personal attributes, etc.; note that time and money are considered 'part of the person,' whereas other extrapersonal physical resources, like alcohol, are considered parts of the situation) may also be varied (Read et al., 2010, p. 65).

In Read et al.'s model, a motive's degree of activation is determined by situation, experience (settings or weights established during training), the motive's baseline activation, the strength of the inhibition system, and the sensitivity of the relevant motivational system. In turn, motive activation produces behavioral output (drawn from a list of 43 possible behaviors, which included drinking alcohol, working extra-hard, and telling jokes). Two 'hidden layers,' one of which was placed between the situational and goal layers, and one of which was placed between all other layers and the behavioral layer, allowed the model to learn to predict and enact appropriate behaviors on the basis of feedback regarding combinations of situations and traits. Once the model was built, it could be trained and tested. 'Training events' involved presenting the network with pairings of situations and goals, situations and behaviors, etc. so that it could learn which situations were likely to allow for the fulfillment of various goals and which behaviors lead to said fulfillment. The first few tests administered to the model showed simply that it learned the appropriate relationships between situations and behaviors and that its behavioral output did not vary wildly when small changes were made to the situations on which it had been trained.

Read et al. (2010) ran the network through eight simulations designed to test how well it reproduced familiar research findings. The first three simulations explored the impact of individual differences in reinforcement sensitivities (or, perhaps, of broad traits like Extraversion and Neuroticism), with the following results. (1) The Approach and Avoidance systems' sensitivity settings strongly influenced the systems' activations and the frequency with which they generated behaviors. (2) The model's built-in positivity offset and negativity bias generated more approach than avoidance behaviors when goals were far away, but also produced more avoidance behaviors as goals drew nearer. (3) When avoidance goals were more highly activated than approach goals during training (i.e. when punishment sensitivity was greater than reward sensitivity), the network produced more avoidance behaviors when tested on the same situations, reflecting the tendency of those high in avoidance motivation to learn more quickly about punishment cues (as demonstrated, for instance, by Zinbarg & Mohlman, 1998).

The other five simulations focused on interactions between the approach and avoidance systems and on how the latter's interactions with other network elements might mimic narrower traits. In simulation (4), Read et al. (2010) showed, again supporting Miller (1959), that approach and avoidance systems compete for control of behavior, resulting in nonlinear relationships between system activation and behavior. Simulation (5) revealed that higher levels of inhibition led to lower levels of activation in approach and avoidance systems, as well as to less behavior-switching. The sixth (6) simulation demonstrated that behavioral output for facets of Extraversion (sociability), Conscientiousness (industriousness), and Neuroticism (confidence) could be mimicked simply by varying the biases of specific goal and resource settings, rather than varying the approach, avoidance, and inhibition settings. Simulation (7) manipulated goal/resource and motivational system settings (Read et al., 2010, p. 78) to create a facsimilie of 'the communal

component of Extraversion' (akin to DeYoung's (2007) Enthusiasm aspect of Extraversion) and replicate the situationally influenced within-persons variability discussed by Fleeson (2007a). Finally, simulation (8) modeled an even more specific trait, rejection sensitivity; by varying the strengths of only two avoidance goals and making Avoidance slightly more sensitive than Approach, the model behaved in a more socially withdrawn manner.

2.8 The Cues-Tendencies-Actions (CTA) Model

2.8.1 Theory: Cues, Tendencies, Actions

The Cues-Tendencies-Actions model (Revelle & Condon, 2015) is simpler than Read et al.'s (2010) neural network model of personality (CTA is deterministic whereas the Read et al. model is probabilistic); however, its simplicity may prove to be an asset to future attempts to model RST2, inasmuch as things that are easier to understand are easier to manipulate along theoretical lines. CTA is instantiated in the open-source programming language R (R Core Team, 2015), and the 'cta function' is included in R's *psych* package (Revelle, 2016). R and *psych* may be downloaded free of charge from the Comprehensive R Archive Network (CRAN, http://cran.r-project.org).

CTA is, in fact, an updated version of a program that Revelle (1986) described, which was itself inspired by Atkinson's Dynamics of Action model (DOA, Atkinson & Birch, 1970). The DOA was prescient in its focus on the temporal dynamics of psychological processes, and it essentially ascribed the properties of physical inertia as stated in Newton's First Law of Motion to constructs called *action tendencies* (i.e. wishes or desires). Just as a body in motion tends to remain in motion, a desire persists until it is fulfilled; just as a

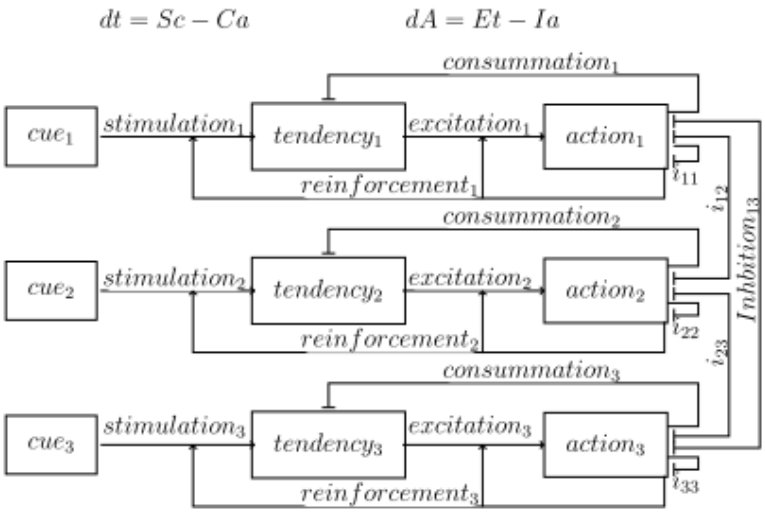


Figure 2.4: Adapted from Revelle and Condon, 2015: CTA Flow-chart. Note that t = tendency, a = action, c = cue, S = stimulation, C = consummation, E = excitation, I = inhibition.

body at rest tends to remain at rest, a desire neither persists nor grows until it is instigated. The DOA modeled action tendencies as differential equations; however, its fusion of desire and action created the need for additional extraneous mechanisms, some of which were introduced to solve the problems caused by others of their kind. The CTA model improves upon the DOA by constructing separate differential equations for tendencies (desires, wishes) and actions; i.e. tendencies (which mimic the function of Read et al.'s hidden layers) and actions each have their own independent inertial parameters. The equations for dynamically changing tendencies (i.e. a vector of tendency strengths, **t**) and actions (a vector of action strengths, **a**), respectively, are as follows:

$$d\mathbf{t}=\mathbf{Sc}-\mathbf{Ca} \tag{2.1}$$

and

$$d\mathbf{a}=\mathbf{Et}-\mathbf{Ia} \tag{2.2}$$

These equations are depicted as a flow-chart in Figure 2.4. Specifically, CTA tendency strengths change over time not only as relevant situational cues change (such that tendency strength increases with cue strength **c** or intensity, in a manner that is more specifically determined by innate or learned association strengths or connection weights between cue-tendency pairs), but also as the tendency's associated action strength increases (such that tendency strengths decrease with increasing action strengths, in a manner specifically determined by that action's 'consummatory strength,' or the efficiency with which it satisfies the associated desire or tendency). Action strengths increase as tendency strengths do, according to tendency-action weights, and decrease as they themselves increase, according to the degree to which the action inhibits itself (and other actions inhibit it).

Actions are initiated when their strengths are large enough to overcome the inhibiting influence of other actions, and, in theory, performing an action not only reduces the tendency to act but also strengthens the cue-tendency and tendency-action associations, which in turn slightly increases the absolute values of both the tendency's and the action's strengths (note well that the current version of the CTA model does not learn in this way, although it does simulate actions' self-inhibiting effects). The differential equations linking tendencies to cues and actions, and actions to tendencies and actions, are most generally expressed as vector, rather than scalar, formulas; that is, multiple cues and actions influence multiple tendencies simultaneously, as multiple tendencies and actions influence multiple actions simultaneously. Unlike other proposed cybernetic models of personality (DeYoung, 2015), the CTA model lacks a 'set point' (like the

homeostatic temperature on a thermostat); instead, equilibrium is the result when both differential equations are equal to zero, representing a situation in which cues' excitatory effects are matched by actions' inhibitory effects.

Within the CTA model, it is the cue-tendency and tendency-action association strength matrices, as well as the consummatory strength and self-inhibition matrices (labeled **S**, **E**, **C**, and **I**, respectively), that are treated as stable and are proxies for personality traits (this is not to say, of course, that association strengths and other connection weights do not change over time; rather, the cta function in *psych* as yet lacks the ability to learn and cannot be used to model trait development). The changing cue, tendency, and action strengths (**c**, **t**, and **a**) are state, rather than trait, variables; trait matrices are thus derivatives of states.

Revelle and Condon (2015) used CTA to model the effects of setting **c**, **S**, and **I** to different values to create three mutually compatible actions (all tendencies and actions equilibrate, reaching their asymptotic values) and compare those to the effects of setting **c** to different initial values while leaving the **S** and **I** matrices equal by default to create three mutually incompatible actions (action tendencies rise and fall in parallel, while actions are carried out in series due to inhibitory influences). Interestingly, the latter may be used to simulate the temporal patterning *either* of a group of many mutually exclusive behaviors competing for expression within a single individual *or* a single behavior as expressed within a group of individuals for whom one person's performance of the action in question acts as an inhibitor of same among all the others (e.g., a conversation in which only one person speaks at a time). Furthermore, varying the input combinations of cue strengths (**c**) was used to simulate conversations within groups containing different proportions of introverted and extraverted individuals.

2.8.2 Practice: The AAC Triad Model

Like Read et al.'s neural network model, CTA dovetails neatly with RST. Indeed, as Revelle and Condon (2015) point out, sensitivities to cues for reward and punishment are reflected in the growth rates of tendency strengths. Furthermore, the decay rates of tendency strengths may differ with individual differences in systems governing learning (e.g., habituation) and satiety. Revelle (2012) has suggested that CTA be used in the modeling of emotion dynamics, and Revelle and Condon (2015) describe CTA's power as a tool for exploring individual differences at three levels of analysis: between-persons, within-persons, and between-groups.

CTA has, in its various guises, gathered promising theoretical and empirical support over the last three decades. Early versions of CTA (and, earlier, DOA) were used to model trial-to-trial behavior in response to success and failure feedback in achievement settings; simulations showed that effort decreases following success feedback and increases following failure feedback within a given set of trials, even though the correlation between success and performance over the long term (i.e. between multiple sets of trials) is positive (Revelle, 1986; Revelle & Michaels, 1976).

As yet, the best concrete example of CTA's efficacy as a platform for modeling RST is the Fua et al. (2010) Approach-Avoidance-Conflict (AAC) Triad simulation, an individual differences framework that combines the insights of Gray and McNaughton's RST2 (2000) and Ortony, Norman and Revelle's (2005) three-level 'ONR' theory, which claims that the key components of personality (affect, behavior, cognition, and desire, personality's 'ABCD's') must each be considered at three levels of information processing, also known as the 'Three Rs' of the Conceptual Nervous System which, in ascending order of complexity, refer to the Reactive, Routine, and Reflective levels of processing (roughly

corresponding to the cerebellum and brain stem, limbic system, and cortex, respectively). A few words of explanation regarding both the Three R's and the ABCD's may be useful here.

Of the three ONR levels of processing, the reactive level receives perceptual input and produces hard-wired behavioral outputs; forms of learning thought to take place on this level include habituation and some classical conditioning. The routine level also receives perceptual input, but its output consists of more flexible, if well-learned, behaviors. Unlike the reactive level, which only processes information about the present and (primitive) past, the routine level is able (primitively) to predict the future. The routine level encompasses learning algorithms that govern operant conditioning, case-based reasoning, and more complex forms of classical conditioning. The reflective level neither receives perceptual input nor is solely responsible for producing behavioral output; it may only receive information from, and pass it to, the routine level. However, the reflective level is the most complex of the three (mentalization is not limited to the past, present, and primitive future, but elaborates upon the latter and extends it to hypothetical scenarios) and is probably the level responsible for producing consciousness. All higher-order learning and cognition take place there, enabling this level to act as a check on the routine level.

The ABCD's of personality (Allport, 1937; Hilgard, 1980; Revelle, 2008; Wilt, 2014) will be defined here as they are in the ONR paper (Ortony et al., 2005). Affect encompasses all valenced feeling states, including 'emotions, moods, feelings, and preferences' (Ortony et al., 2005, p.174); it becomes more akin to recognizably human emotion at higher ONR levels and as organisms' cognitive complexity increases (reactive-level affect is termed 'proto-affect;' routine-level affects comprise 'primitive emotions;' and reflective-level

affects are 'true emotions'). Whereas affect 'has to do with value' (what's good), cognition 'is essentially concerned with meaning' (what's true; Ortony et al., 2005, p.174). Cognition and affect may interact, and neither provides a perfect guide to adaptive functioning. Desire (motivation) 'concerns tendencies to behave in certain ways' (Ortony et al., 2005, p.174); as does the CTA model, ONR distinguishes between the *tendency* to behave in a certain way and the behavior itself (tendencies are stimulus-response pairings, whereas behaviors are a class of responses or actions). Finally, Ortony et al. (2005) define behavior as any internal or external physical action that an organism's body may perform (excluding physical phenomena occurring in non-motor areas of the brain).

The ideas of ONR, ABCD, and RST intersect in interesting ways. Ortony et al.'s (2005) notion of motivations as stimulus-response (S-R) pairs corresponds directly to Gray and McNaughton's (2000) *goal* construct (described in detail elsewhere). Moreover, thinking of the ABCD's in terms of their status as either stimulus-side (input) or response-side (output) elements in S-R pairs provides insight into their similarities and differences. For instance, affect and cognition (in the sense of what an organism 'knows, thinks, and believes,' Ortony et al., p.174) may be seen as either inputs or outputs, depending on level of processing. Nevertheless, because the most primitive S-R pairings consist of sensations (S) and behaviors (R), cognition (an elaboration on non-cognitive, reactive-level sensation) is more like a stimulus-side process, whereas affect (which, due to its directional, imperative quality, serves as an efficient guide to behavior) is more like a response-side process.

When Fua and his colleagues modeled RST at the ONR reactive level, the CTA cues consisted of other environmental objects' names, actions, and distances from the subject, which were then translated into information about the reward (R+), non-reward (R-),

punishment (P+), and non-punishment (P-) values of both the cues themselves and of their predicted actions. These valuations in turn activated the AAC model's approach/BAS (*play*, *friend*, *safety*) and avoidance/FFFS (*offensive aggress*, *avoid bad*) action tendencies (with individual differences in BAS and FFFS sensitivity modeled by the sensitivity with which the tendencies react to the cues; i.e. by values in the **S** matrix). An *info-gather* tendency was activated by a BIS module that was itself activated, and which inhibited conflicting tendencies, whenever a pair of BAS and FFFS tendencies (say R+ and P+) were active. The sensitivity of the BIS to differently coactivated BAS and FFFS tendencies also varied, and served as a model of BIS sensitivity. In addition, the BIS module allowed for expectation and action modification when expectations were violated.

2.9 Summary and Statement of Purpose

The foregoing discussion has demonstrated that RST is an empirically sound theory of individual differences with enormous potential to lend much-needed neurophysiological precision to personality psychology. Currently, however, it appears to be experiencing 'growing pains.' Nearly two decades have passed since its last revision (Gray & McNaughton, 2000), and the intervening years have seen little more in the way of development than a straggling parade of semi-valid RST self-report inventories. I propose that what needs to be done in order to spur progress is to create a precise computational model of RST that will allow researchers to study the time-dependent relationships between RST's many dynamically-interacting parameters and processes. Such a model would be capable not only of helping scale-developers build better self-report measures, but also, and more significantly, of enabling workers in the field to refine the theory itself.

After reviewing the relevant computational modeling literature, I combined the best aspects of Pickering's models of RST2 (Pickering, 2008; Smillie, Pickering, & Jackson, 2006) and Revelle and colleagues' CTA model (Fua, Revelle, & Ortony, 2010; Revelle, 1986; Revelle & Condon, 2015) to create a new, CTA-based model of RST (CTA-RST) capable of generating realistic affective and behavioral data. These two models were chosen for their relative simplicity, compatibility, and accessibility. To reiterate, Pickering's RST simulations were preferable to those of Corr and McNaughton because the former belong to a better-established research tradition.

I crossed Revelle's CTA algorithm with Pickering's simulated RST for two reasons. First, CTA has its roots in a distinguished line of inquiry outside, yet with similar goals to, those of RST; therefore, CTA was a good candidate to infuse Pickering's models with 'fresh blood,' as it were. Second, CTA and Pickering's computational formulation of RST have much in common, making it relatively simple to combine the two programs. The enhanced capabilities that the CTA code offers include the ability to model multiple actions or goals simultaneously, alternative graphing tools, and a more general conceptual framework in which individual difference variables other than the three RST systems' sensitivities may be specified.

Relatively little has been retained in CTA-RST of the Read et al. (2010) neural network, which is more complex and arguably more biologically plausible than either Pickering's RST or Revelle's CTA. Indeed, the 'positivity offset' principle is the only idea from Read et al.'s (2010) model that found its way into CTA-RST. Including positivity offset in CTA-RST will make the simulation more realistic in a functionally meaningful way, whereas other details of the Read et al. model seem less likely to make an important difference in the simulation's output. Ultimately, I used simple, easily-manipulable

programs to build CTA-RST, the first in an intended series of models (ideally of increasing sophistication). Biological verisimilitude, although admirable and worth pursuing in future projects, is not essential to a good computational model of RST. After all, even if a simulation does not function in precisely the same way as the brain does, the two are considered cybernetically equivalent or 'isomorphic' as long as the former's outputs cannot be distinguished from the latter's (Ashby et al., 1956). Naturally, it may be impossible to establish isomorphism quickly or at all; however, that will not prevent me from trying.

In the remainder of this paper, I outline the procedure I used to develop CTA-RST and describe three studies that I simulated in order to match simulated data to real data using similar operational definition parameters across studies. I focus on studies that collected affect data because, whereas previous work on the CTA model has demonstrated that it is capable of simulating choice data (Revelle & Condon, 2015), there is as yet no evidence that directly attests to CTA's facility at mimicking affective data. The best-known and -validated of the top-down personality traits (i.e. Extraversion, Neuroticism) have strong affective components, and if a model can generate convincing affect and choice data, then it can reasonably be said to have begun faithfully to reflect the dynamic intraindividual processes that give rise to many if not most interindividual differences in personality traits. Importantly, any model that satisfies these criteria in turn becomes a powerful tool for ameliorating the conceptual 'discontents' that plague RST.

Chapter 3

A CTA Model of RST

> If it looks like a duck, and quacks like a duck, we have at least to consider the possibility that we have a small aquatic bird of the family Anatidae on our hands.
>
> Douglas Adams, 1987 (Dirk Gently's Holistic Detective Agency)

In this chapter, I describe the construction of the CTA-RST program and provide a detailed rationale for selecting the three studies I have modeled. Afterward, I discuss the nomenclature used in subsequent chapters to refer to the operational definitions of affect and personality employed in each simulated study.

3.1 Developing CTA-RST

The open-source code for Pickering's published RST2 simulations is available online at *http://homepages.gold.ac.uk/aphome/newrst4.m* and is reproduced, in abridged form (i.e. without some comments and plotting functions, omitted from CTA-RST, that are included in the online version), in Appendix A. As Pickering's models are written in

MatLab, I first translated his code into R (R Core Team, 2015) so that it would be in the same language as CTA (the original code for which may be found in Appendix B). I ran a series of replication simulations in order to debug the translated Pickering code and ensure that it produced results commensurate with those previously published. Specifically, Pickering and I agreed (personal communication) on the values of the correlations among BAS, BIS, and FFFS functional outputs, as well as on the regression weights relating functional outputs to sensitivity parameters for three sets of inputs: those of Smillie et al. (2006), those of Pickering (2008), and a novel third set whose outputs are as yet unpublished. Table 3.1 compares the regression weights of the Pickering model I ultimately chose to imitate to those produced by its R translation.

In line with many of Pickering's other creations, my CTA-RST simulations' parameters include subjects, events (each of which includes reward and punishment cue strengths), reinforcement sensitivity weights, and time (i.e., number of iterations). The BAS, FFFS, and BIS 'weight' parameters are particularly crucial, as they are what represent individual differences in the systems' sensitivities. A detailed discussion of all such settings may be found in this chapter's third section.

Of interest here, as well, are the versions of Pickering's RST-system equations that I ultimately modified to create CTA-RST. Pickering's original program contains four alternative equations for the BAS and FFFS (another was added in the course of our correspondence) and two alternative equations for the BIS. I altered the following set of BAS, FFFS, and BIS equations (shown here in Pickering's original notation and depicted graphically in Figure 3.1) for use in my own model:

$$dx_{BAS} = (k_{1.BAS} * s_{BAS} * w_{BAS}) * (Max_E - x_{BAS.OUT}) - [(k_{4.BAS} * x_{FFFS.OUT}) + (k_{2.BAS} * x_{BIS.OUT})] * (Max_I + x_{BAS}) - (k_{3.BAS} * x_{BAS}) \quad (3.1)$$

$$dx_{FFFS} = (k_{1.FFFS} * s_{BFFFS} * w_{FFFS}) * (Max_E - x_{FFFS.OUT}) - (k_{4.FFFS} * x_{BAS.OUT}) * (Max_I + x_{FFFS}) + (k_{2.FFFS} * x_{BIS.OUT}) * (Max_E - x_{FFFS.OUT}) - (k_{3.FFFS} * x_{FFFS}) \quad (3.2)$$

$$dx_{BIS} = (k_{1.BIS} * x_{BAS.OUT} * x_{FFFS.OUT} * w_{BIS}) * (Max_{E.BIS} - x_{BIS}) - (k_{3.BIS} * x_{BIS}) \quad (3.3)$$

In particular, this version of the BIS was used by both Smillie et al. (2006) and Pickering (2008), whereas this version of the BAS and FFFS were used only by Pickering (2008). In CTA parlance, the BAS, BIS, and FFFS were modeled as separate cue/tendency/action triplets, and a complete set of BAS, BIS, and FFFS triplets constituted what I decided to call one *goal*. Multiple goals were modeled with multiple BAS, BIS, and FFFS triplets. A full copy of the CTA-RST code may be found in Appendices C (which contains the model's core processing code) and D (which contains the functions called by the core code).

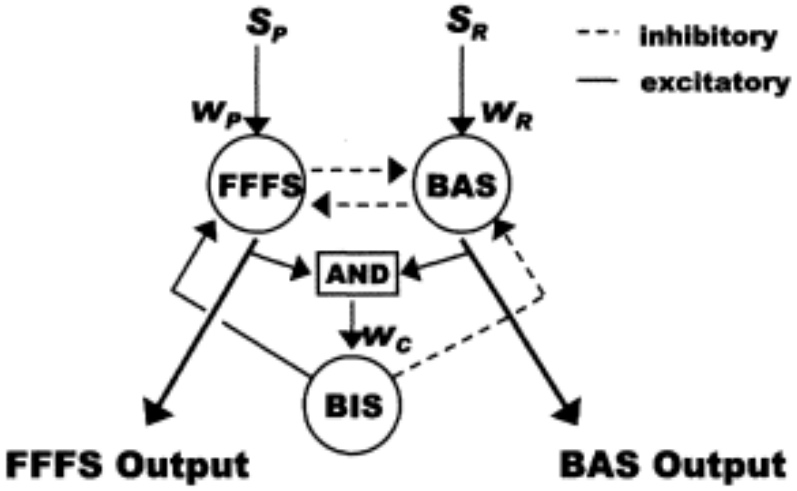


Figure 3.1: Adapted from Pickering, 2008: Model of Revised RST

Table 3.1: Beta Weights for Regressions of System Output on Instigating Cue Sensitivities in Pickering's RST Simulation vs. Brown's R Translation

P = Pickering; B = Brown

	BAS_w		$FFFS_w$		BIS_w	
Output	P	B	P	B	P	B
BAS	.45	.47	-.27	-.29	-.15	-.10
FFFS	.45	.47	-.27	-.29	-.15	-.10
BIS	-.16	-.19	.51	.54	.15	.10

3.2 Rationale for Three Simulations of Real Studies

In selecting the three studies I would use to demonstrate CTA-RST's utility, I looked specifically for research that examined personality and affect as dynamic processes, rather than as static structures. In addition, I gave priority to studies conducted by my own lab and its collaborators, so that data (and answers to my occasional questions) would be easily accessible. Ultimately, I selected one experimental study and two observational studies; the former used mood manipulations to explore the interactions between affective

states and personality traits, whereas the latter used experience sampling, or daily diary data on personality and affective states, to investigate questions directly related to personality dynamics. These three studies were, to my knowledge, the ones that most directly addressed the issue of the dynamics of state-trait (i.e. structure-process) links within personality psychology. Happily, all three were conducted by members of or collaborators with Northwestern University's Personality, Motivation, and Cognition (PMC) lab.

3.2.1 Study 1: Smillie et al., 2012: Do Extraverts Get More Bang for the Buck?

The first study that I elected to model used experimental manipulations to examine how the affective impact (activated versus pleasant positive affect) of two ostensibly different kinds of behavior (pursuing versus enjoying or consuming rewards) differed as a function of a top-down trait, extraversion, which is thought to reflect individual differences in reward sensitivity. In an effort to explain why extraverts are, in general, happier than other people, Smillie et al. (2012) conducted five experiments; in each, participants experiencing mood inductions (positive, negative, pleasant, appetitive, or neutral) self-reported personality and affect data. Pleasant stimuli improved extraverts' and introverts' moods similarly, demonstrating that extraverts aren't happier because they react more strongly to pleasant cues (e.g. visualizing relaxing on a beach, watching a 'feel-good' clip from the movie *Parenthood*, etc.).

However, when participants were asked to visualize, watch, or participate in appetitive situations that involved active reward pursuit (e.g. to watch clips from a James Bond film or imagine themselves winning the lottery), extraverts' emotional reactions were stronger than introverts'. Specifically, extraverts experienced increased activated affect,

i.e. feelings of interest and enthusiasm, *not* increased pleasant affect (contentment), as a result of appetitive mood inductions.

Study 1 examines a single reinforcement-sensitivity-related trait, extraversion. Its results indicate that if reward sensitivity is the primary determinant of extraversion, then reward sensitivity is unlikely to be the only individual difference involved in producing positive affect. Instead, *energetic* positive affect is related to extraversion, but *pleasant* positive affect is not. The CTA-RST model, with its emphasis on both the sensitivity to rewards and the consummatory value of approach behaviors, seems uniquely suited to disentangling these disparate aspects of positive emotion.

3.2.2 Study 2: Wilt et al., 2016: Velocity Links Personality States and Affect

The prototype for Study 2 explored affective dynamics from an observational perspective, showing that affect is related to one's rate of progress toward goals (i.e. velocity; Wilt et al., 2016). In particular, the study confirmed that state positive affect and state extraversion are positively related, whereas state negative affect is positively related to state neuroticism. Crucially, approach-goal velocity mediated both relationships; approaching goals more quickly partially accounted for the relationship between extraversion and positive affect, whereas approaching goals more slowly partially accounted for the neuroticism-negative affect link. Note here that slowed *approach behavior* was associated with negative affect; thus, depending on the dynamics of approach, either positive or negative affect may accompany approach behavior.

Not only may negative affect sometimes accompany approach behavior, positive affect may sometimes accompany avoidance behavior. Carver (2006) and his colleagues (2009) have long claimed that not all negative affects belong to punishment sensitivity and vice

versa, an assertion supported by Wilt et al. (2011), in regards to a study that will be simulated in Chapter 6. Although this principle may be counterintuitive, it is more easily understood when one considers that a goal-directed action's nature (approach or avoidance) is independent of its consequences (good or bad). It is both the actions and their consequences that generate the affect. That said, 'bad' affects may be FFFS/threat-related (i.e. fear-producing), BIS/goal-conflict-related (anxiety-producing), or BAS/motor-conflict-related (typically sadness- or anger-producing).

Study 2 examines the mechanistic determinants of affect and strongly supports the position that approach-goal velocity is one such determinant. CTA-RST is capable of generating position, velocity, and acceleration data for tendencies and actions associated with each of the three RST systems for any number of goals. Therefore, CTA-RST should be capable of reproducing the Wilt et al. (2016) data, as well as of generating novel predictions regarding other spatiotemporal determinants of affect.

3.2.3 Study 3: Wilt et al., 2011: Affective Synchrony and Personality

With the third and final simulation, I modeled variability in the within-persons experience of affect. In the original observational study, Wilt et al. (2011) demonstrated that individual differences in affective synchrony can be predicted by interactions between traits thought to reflect reinforcement sensitivities.

Affective synchrony is the positive correlation between one's experience of positive affects (e.g. energetic arousal) and negative affects (e.g. tense arousal). If you tend to feel positive and negative affect at the same time, then the latter are positively correlated and you are affectively *synchronous*. If you tend to feel either positive or negative but not both, then your affective experiences are negatively correlated and you are affectively

desynchronous. If your emotions are sometimes mixed, sometimes not, then your positive and negative affective experiences are uncorrelated and you are affectively *asynchronous*. Most people are affectively asynchronous, but individual differences in affective synchrony are nontrivial.

Wilt et al. (2011) found that when scores on affective traits like extraversion and neuroticism were more similar, participants were more likely to experience synchronous emotions. I therefore chose to use CTA-RST to model this study's data because the latter are consistent with what Pickering (2008) and Corr (2004) say about the joint effects of reinforcement sensitivities. Note here that, although it is tempting to suppose that individuals with more evenly-matched reward-as-BAS and punishment-as-FFFS sensitivities will be more likely to experience mixed emotions, a point of caution is in order. Evenly-matched BAS and FFFS sensitivities *are* more likely to give rise to goal conflict and general behavioral inhibition. However, such behavior (cessation of action coupled with goal conflict) corresponds to a *negative* affect, anxiety, not *mixed* negative and positive affects. On the other hand, if self-reported reward and punishment sensitivities are not one-to-one analogs of BAS and FFFS sensitivities, but rather are composites of BAS and BIS, or FFFS and BIS parameters, then it is more likely that matched self-reported sensitivities will correspond to mixed affect rather than negative affect.

3.2.4 Summary

As stated, the three studies I have chosen to simulate are those that I believe best interrogate the processes underlying the relationships between state and trait individual differences. Both experimental and observational (experience-sampling) approaches are

represented, and the original data were conveniently available from members of the PMC lab or its collaborators. Additionally, these studies may be construed as building upon each other to create a relatively complete picture of the affective consequences of reinforcement sensitivities. Study 1 focuses on positive (BAS-related) affect, demonstrating that changes in the latter are independent of changes in negative affect and that different kinds of positive affect are differently dependent on the reward-sensitivity-related trait, extraversion. Study 2 considers negative (FFFS-related) as well as positive affect, linking these to state neuroticism and extraversion, respectively. In addition, Study 2 delves into action's relationship to affect, suggesting that the direction of and changes in goal-approach velocity mediate the links between affect and state personality. Finally, Study 3 investigates the manner in which positive and negative affect co-occur within persons, a species of interaction which may depend on the BIS.

3.3 Rationale for Parameter Inputs

In constructing CTA-RST, all input parameters were carefully considered, as the benefits of computational models are balanced by and ultimately products of the high level of control afforded the modeler in constraining simulated 'reality.' Without narrowing the range of possible variations to the program's basic design in a manner informed by both logic and empirical research, a simulator runs the risk of invalidating her model via confirmation bias or excessive researcher degrees of freedom (Simmons, Nelson, & Simonsohn, 2011). The possible sources of variation in the CTA-RST model include, at the very least, its essential input functions; i.e. sensitivities, cues, tendencies, actions, consummations, excitations, and inhibitions. Further, if the model is to learn, as would be desirable albeit unnecessary (previous simulations without the ability to learn have

produced solid data), then we must add expectation and reinforcement functions to the list.

Recall that cue (and, if the model learns, reinforcement) functions are meant to represent sensory input from the environment, tendencies and actions are output (analogous to the subject's desires or affects and behaviors, respectively), and sensitivities, consummations, excitations, and inhibitions (plus expectations in a learning model) are person variables; i.e. possible sources of individual differences. In CTA-RST, only the sensitivity and consummation functions differ between simulated persons; these reflect the subjects' sensitivity to instigating cues in the environment and to the satiating effects of actions, respectively. In what follows, I will describe the input functions employed in the model and explain why each was chosen. Figure 3.2 and Table 3.2 summarize the role of each of CTA-RST's input functions.

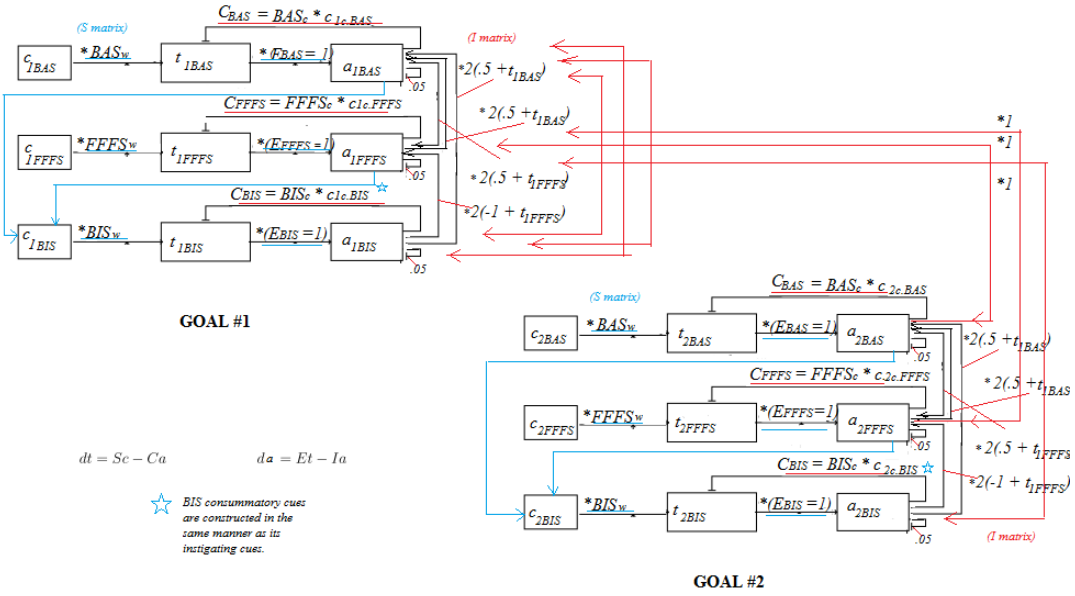


Figure 3.2: Adapted from Revelle and Condon, 2015: CTA-RST Flow-Chart

Table 3.2: Notation for 2-Goal CTA-RST Flow-chart

Note: Goal-number subscripts omitted in text; untabulated BIS and FFFS notation similar

Function	System, goal	Fixed/Free	CTA-RST Notation (for Goal g)
*d***t**		Input/Output	
Action (**a**)	BAS, g	Output	a_{gBAS}
Instigating cue (**c**)	BAS, g	Free, Input	c_{gBAS}
Instigating cue sensitivity (**S**)	BAS	Free, Input	BAS_w
Sc	BAS	(Intermediate Output)	c_{gBAS}*BAS_w
Consummatory cue	BAS, g	Free, Input	$c_{gc.BAS}$
Consummatory cue sensitivity	BAS	Free, Input	BAS_c
Consummation (**C**)	BAS, g	(Intermediate Output)	C_{gBAS}=$c_{gc.BAS}$*BAS_c
−**Ca**	BAS, g	(Intermediate Output)	-C_{gBAS}*a_{gBAS}
*d***a**		Input/Output	
Tendency (t)	BAS, g	Output	t_{gBAS}
Excitation	BAS	Fixed, Input	$E_{BAS} = 1$
Et	BAS, g	(Intermediate Output)	$E_{BAS} * t_{gBAS}$
Inhibition (**I**)	BAS-on-BAS, g	Fixed, Input	$I_{11} = I_{44}$ =.05
	FFFS-on-BAS, g	(Intermediate Output)	$I_{12} = I_{45}$ =2(.5+t_{gBAS})
	BIS-on-BAS, g	(Intermediate Output)	$I_{13} = I_{46}$ =2(.5+t_{gBAS})
	BAS-on-FFFS, g	(Intermediate Output)	$I_{21} = I_{54}$ =2(.5+t_{gFFFS})
	FFFS-on-FFFS, g	Fixed, Input	$I_{22} = I_{55}$ =.05
	BIS-on-FFFS, g	(Intermediate Output)	$I_{23} = I_{56}$ =2(-1+t_{gFFFS})
	BAS-on-BIS, g	Fixed, Input	$I_{31} = I_{64} = 0$
	FFFS-on-BIS, g	Fixed, Input	$I_{32} = I_{65} = 0$
	BIS-on-BIS, g	Fixed, Input	$I_{33} = I_{66} = .05$
	System-on-System, different goals	Fixed, Input	All other $I_{jk} = 1$
−**Ia**	BAS, 1	(Intermediate Output)	-$(I_{11} + I_{12} + ... + I_{16}) * a_{1BAS}$

3.3.1 Cues

Cue inputs describe the intensity of the rewarding and punishing aspects of the different goals that a simulated individual may act on. The cue input function requires specification of a distribution type, distribution parameters, and a rule for creating BIS cues from BAS and FFFS cues. The simulation's cues may be normally- or uniformly-distributed; custom cue selections are also possible. Evenly-spaced custom- and uniformly-distributed cues demand maximum and minimum settings for all systems with nonzero initial values; normal distributions require that means and standard deviations be specified. In all cases, BIS cues may be set to start at zero or at the values of the aforementioned input function. Below, I outline the reasoning regarding my choices of BAS and FFFS cue distributions, BAS and FFFS distribution parameters, and BIS cue input.

Cue Distribution:

Different cue distributions were most appropriate for different stages of simulation development and testing. During the initial phases of model construction, evenly-distributed custom cues were most useful, as these allowed for a clear understanding of how systematic changes in cue strength influenced other variables. However, in testing the simulation's ability to model real observational or experimental data, cue distributions were chosen for their verisimilitude; for instance, Study 1, in which a baseline condition's affects (presumably reflective of subjects' experiences immediately prior to participating in the study) were compared to those following a positive (or appetitive or pleasant), negative, or neutral mood induction, required one set of random cues ('baseline condition') and three sets of custom cues ('experimental conditions') per participant group. I elected to model the baseline conditions using random normal cue distributions.

Random normal distributions were chosen not only to represent baselines in Study 1, but also to represent all data in Studies 2 and 3, in the original versions of which participants texted 'daily diary entries' of affect and personality ratings to researchers. While it might be argued that a random uniform distribution of stimulus cues would serve just as well as and, indeed, be more likely to yield statistically significant effects than a normal distribution, I use the latter for two reasons. First, statistical significance is of somewhat secondary importance in modeling data via computer simulation, as all such data are clean and 'noiseless' unless the programmer decides otherwise. Second, normal distributions are probably more realistic emulators than uniform distributions of everyday experience, particularly in the lives of the college students from whom most of the psychological data that I am modeling were collected.

That normal distributions are faithful analogs of university students' typical encounters with rewarding and punishing cues can be effectively defended from several angles. First, college student samples are relatively homogeneous (Henrich, Heine, & Norenzayan, 2010); therefore, certain problems that might arise from normal distributions' use are less likely to do so under such conditions. For example, it is less likely that subjects in a smaller, more homogeneous group will be familiar with the full range of valenced experiences, from the extremely positive (e.g., getting a large pay raise) to the extremely negative (e.g., grappling with the challenges of hunger and homelessness). Uniform cue distributions will display such extremes more frequently than cue distributions for the average student population will.

Similarly, to the extent that all college students have chosen the same niche, there is no great need to use the model's cue distribution to account for 'niche-picking' (Cole, Hood, & McDermott, 1994), the process by which different individuals' sensitivities to rewards and punishments lead them to choose environments that offer a preferred distribution of reinforcing and non-reinforcing stimuli. Indeed, a common, constant uniform cue distribution approximates this state of affairs less well than does a common normal distribution. Ideally, of course, each subject would have a unique random-normal cue distribution whose mean and standard deviation were some function of his or her reward and punishment sensitivities. Again, however, since most of the participants modeled were university students who had, to some extent, already selected the same environment, I decided that it was reasonable to assume single, common normal distributions for BAS and FFFS cues.

Finally, normal distributions may easily be used to indirectly represent positivity offset (Cacioppo et al., 1997) in either the cue or sensitivity functions. However, incorporating

positivity offset into cue or sensitivity distributions implies that the mean of the BAS distribution would be higher than the mean of the FFFS distribution, making positivity offset essentially a between-subjects individual difference; that is, *on average*, people would be more sensitive to rewards than punishments, but *not every* individual would be. I believe, however, that positivity offset is a within-subjects individual difference, such that (while individual degrees of positivity offset may differ between persons) *all* individuals' baseline sensitivity to rewards must be higher than their sensitivity to punishments. It is better, then, to build positivity offset into the simulation elsewhere (see below).

Distribution Parameters: Maxima and Minima:

The minimum and maximum values of a distribution specify its range and, for these distributions in particular, quantify the extrema of rewarding and punishing stimuli. Evenly-distributed custom BAS and FFFS cue distributions' maxima and minima were set at 0.95 and 0.05, respectively. I made the assumption here that the full range of stimulus intensity varies between 0 and 1, with these extremes marking the dimension's negative- and positive-infinite poles after standardization. Given their low likelihood of occurring (and the fact that values equal to 1 or 0 sometimes threw calculations into disarray), the error created by eliminating values greater than 0.95 and less than 0.05 is likely to be within tolerance for simulation-development and -testing purposes. Furthermore, I can think of no research-supported grounds for making BAS and FFFS minima and maxima differ.

Distribution Parameters: Means and Standard Deviations:

In accordance with the reasoning behind my decision to use a single random normal distribution to generate all BAS or FFFS cue input for baseline (Study 1) and experiential (Studies 2 and 3) participant data, I will be assuming that these distributions' means and standard deviations refer only to the participant population and not to any general population from which the former might differ. Thus, the choice of the first mean is an arbitrary one, and intuition suggests that it be set at 0.5, halfway between the hypothetical extremes of 0 and 1. The choice of first standard deviation is also arbitrary, and I will set it at 0.2, so that the values 0 and 1 are roughly two standard deviations away from the distribution's mean. The choice of mean for the second distribution depends on whether or not it allows for positivity offset, regardless of whether the distribution with M = 0.5 and SD = 0.2 describes the BAS or FFFS cues. That said, I will not be using the cue distribution to instantiate positivity offset, so the means and standard distributions of the BAS and FFFS cue inputs are equal (M = 0.5, SD = 0.2).

BIS input:

The choice of function governing BIS cue input must be consistent with theory, which states that the BIS receives input from both the BAS and FFFS, and that its activation is greater the greater the conflict between different goals (e.g. a BAS goal and a FFFS goal). To reiterate, cue input to the BIS can either start at zero or at the value of the BIS input function for that experience's BAS and FFFS cues. Note that even if BIS cues start from zero (as they might if actions started at zero, as discussed below), a function still has to be chosen in order to update BIS cues on the basis of BAS and FFFS action output (a_{BAS} and a_{FFFS}, respectively) following the first cycle of iterative calculations.

In particular, I tested two such functions. The first was a simple product of BAS input and FFFS input; i.e. reward cues (c_{BAS}) times BAS sensitivity (BAS_w) and punishment cues (c_{FFFS}) times FFFS sensitivity ($FFFS_w$), respectively. That is, input to the BIS cue matrix at each iteration (dc_{BIS})) is

$$dc_{BIS} = a_{BAS} * a_{FFFS} \tag{3.4}$$

This is the function used in Pickering's simulation. Here, the cue input to the BIS is the product of the BAS and FFFS outputs, which I liken to CTA-RST BAS and FFFS actions. Thus, this function's BIS input is equal to the product of the initial outputs of the BAS and FFFS. Although the product will be larger as any one of the four variables involved increases, the range of BIS inputs will be restricted relative to that of BAS and FFFS inputs for any observed range of sensitivities and cues.

The second, slightly more complicated, function was borrowed from Corr and McNaughton's simulation (2008). Here, the change in the BIS cue matrix at each iteration is

$$dc_{BIS} = max(a_{BAS}, a_{FFFS}) - |a_{BAS} - a_{FFFS}| \tag{3.5}$$

In this case, the range of BIS inputs will be the same size as the range of BAS and FFFS inputs. However, I have chosen to use the first function because it is more faithful to the original Pickering simulation and because it is not inconsistent with CTA. Specifically, the second term in the Corr and McNaughton (2008) equation essentially duplicates a term in the CTA-RST inhibition matrix that plays a similar role. Any restriction of range

induced by use of the first equation can, if necessary, be offset by an appropriate scalar multiplication.

3.3.2 Weights

Instigating cue sensitivity weight functions assign unique sets of reinforcement sensitivities (i.e. weights) to each simulated individual; these are one of the model's 'individual differences' components. Essentially, individual differences in instigating reinforcement-cue sensitivities act like lenses, magnifying or minimizing objectively identical cues so that they *seem* to have different strengths for different individuals. These weights remain constant as the program runs, reflecting the relative stability of adult personality. As with cues, weights' distributions and attendant parameters must be decided before going forward. Also like cues, weights may be distributed normally, uniformly, or systematically with custom-built, evenly-spaced weight assignments. Uniform distributions require minima and maxima; normal distributions require means and standard deviations.

My reasons for selecting the weight functions and parameters that I have are quite similar to those I had for selecting cue functions. Evenly-spaced weight distributions were ideal for testing the simulation, while random normal distributions seemed likely to produce the most realistic participant samples in Studies 1 through 3. Personality traits are thought to be normally distributed in the population at large (Matthews, Deary, & Whiteman, 2003), and there is little cause to assume that the factors on which college students are selected (i.e., cognitive ability) are so strongly correlated with affective individual differences like reinforcement sensitivities that the resultant restriction in range would cause sensitivities to be non-normally distributed in such samples.

Parameter selection for weights also followed similar lines of reasoning to those that guided cue parameter selection. Maxima and minima for all normal and evenly-spaced custom distributions were set at .05 and .95. Because I don't use weights to instantiate positivity offset, BAS, BIS, and FFFS weight distributions all have SD = 0.2 and M = 0.5. Finally, in keeping with Pickering's model, each weight was multiplied by a system-unique growth constant (k_1) and a constant that sets a maximum value on the system's capacity for activation (Max_E). All values for k_1 and Max_E are set to 1 in Pickering's simulation and in mine.

3.3.3 Tendencies and Actions

Tendencies and actions are the CTA-RST model's dependent variables; they are roughly equivalent to variables in Pickering's simulation that describe each of the RST systems' output. For instance, Pickering's variable BAS describes output that is positive when the system is excited and negative when it is inhibited and is analogous to CTA-RST's BAS tendency, while Pickering's variable BAS_{out}, analogous to CTA-RST's BAS action, is always the maximum of 0 and BAS. Pickering's variables were, however, sometimes difficult to translate into the language of CTA, simply because his 'tendency' variables and 'action' variables sometimes appear in the same terms of his differential equations; in CTA, tendencies and actions are never multiplied together.

CTA tendencies always default to an initial value of zero, while both tendency-like and action-like variables in Pickering's program are set to start at zero. I have chosen, however, to build positivity offset into the tendencies' initial values. Again, positivity offset is the principle that BAS baseline activation is greater than FFFS or BIS baseline activation for all individuals. Thus, BIS and FFFS tendencies will always start at zero,

while BAS tendencies will always start at a constant positive value. I have chosen, arbitrarily, to start BAS tendencies at .05, a value that seemed small enough not to produce a BAS so powerful as to crush all potential FFFS and BIS output. In addition, and for the sake of realism, I decided to introduce a small amount of random-normally distributed error (M = 0 and SD = .005 in all cases) into the tendency calculations at each iteration.

The CTA model sets its default actions' initial values equivalent to the corresponding cues; however, changing the initial values to zero does not greatly alter the output, all else being equal. The same is true for the CTA-RST hybrid. Therefore, although many input patterns are possible, I have chosen for simplicity's sake to set initial actions equal to zero.

3.3.4 Consummations

Three possible consummation matrices were considered for CTA-RST; namely, one that satisfied the default conditions of the original CTA, one that satisfied the conditions of Pickering's model, and one that combined the two. The default CTA consummation matrix did not vary by simulated subject or cue condition; it was always a diagonal matrix with values of .05 on the diagonal (i.e. values describing the consummatory strengths of each system's action on itself for each goal) and all other cells' values set equal to zero. The Pickering analog to this matrix is also a diagonal one such that actions only have consummatory value for the system that generates them; however, these diagonal values vary by both the subject's system's sensitivity (Sys_w) and the strength of the cue exciting that system (c_{Sys}). In other words, the consummatory value of a system's action (C_{Sys}) on the corresponding tendency is always

$$C_{Sys} = k_{1.Sys} * Sys_w * c_{Sys} \tag{3.6}$$

where $k_{1.Sys}$ is an arbitrary constant that Pickering sets equal to 1 for all systems ($k_1 = 1$). Note, of course, that BIS cues will always be equal to the aforementioned function of BAS and FFFS actions; consequently, the consummation matrix must be constantly updated as the simulation proceeds.

Although I prefer to use consummation matrices that vary by subject and experience, I do not believe that the sensitivities and cues involved in consummation are necessarily (nor even ever) the same as the sensitivities and cues involved in system activation. Intuitively, the ease with which an individual is stimulated as a result of a reinforcing cue is not the same as the ease with which the action that results from the stimulation satisfies the tendency that underlies it. For example, the sweet taste of a soft drink is a cue that strongly incites the drink's consumption; however, the tendency that underlies consumption will not be satisfied unless the drink is sufficiently high in calories. A soft drink made with real sugar, therefore, has a higher consummatory cue-value (less action is needed to extinguish the tendency) than a soft drink made with zero-calorie synthetic sugar (more action is needed to extinguish the tendency), despite the fact that the activating cue-values (the drinks' sweetness) may be equal.

All else being equal, different people may also be differently satisfied by stimuli with equal consummatory cue-values, even if they find the stimuli equally stimulating. To recur to the soft-drink example, it is possible for two people who have the same reward sensitivities to be equally excited at the prospect of receiving a soft drink (made with real sugar), say, as a prize after successfully completing a battery of psychological tests.

However, if one has a higher reward-consummatory sensitivity, then he or she will spend less time drinking (or take longer to finish the drink) than the other, who will spend more time drinking (or take less time to finish the drink). Thus, my consummation matrices' diagonals will be comprised of the terms

$$C_{Sys} = k_{Con} * Sys_c * c_{c.Sys} \tag{3.7}$$

where Sys_c and $c_{c.Sys}$ are the consummatory sensitivity of the person (for the appropriate system) and the consummatory cue value of the stimulus (for the appropriate system), respectively. Since all constants k_1 were originally set to 1, I chose to set k_{Con} equal to 1, as well. The consummatory sensitivities' and cues' distributions will be built and parameterized in the same way as the corresponding (BAS, FFFS, or BIS) instigating cues' and instigating cue sensitivities' distributions were.

3.3.5 Excitations

The excitation function is typically a simple one, both in CTA and in Pickering's model. In CTA, the function produces a constant diagonal matrix; the most recent version sets the diagonal values equal to 1, with a step-size (proportional change-per-iteration, which is multiplied by each term in the functions governing changes in tendencies and actions) of .05. Pickering's model implicitly sets excitations equal to 1, inasmuch as his differential equations for each system are unseparated. That is, rather than passing the value for each tendency to a separate action equation to be multiplied by an excitation function and then reduced by an inhibition function, Pickering cuts out the middle man and leaves cue-sensitivity activation, consummation, excitation, and inhibition in a single function.

There is no additional constant multiplier on the cue-sensitivity and consummation terms to suggest, furthermore, that Pickering is 'exciting' his tendencies by any value other than unity. That said, my excitation function will also assign a constant diagonal excitation matrix (diagonal values = 1) to each simulated participant.

3.3.6 Inhibitions

The default inhibition matrix in the CTA program sets diagonal terms to .05 and off-diagonal terms to 1. This in essence makes all actions 'incompatible;' i.e. no two actions can take place at the same time. Each action inhibits itself weakly (.05), but inhibits all other actions totally (1). In CTA-RST, actions directed at different goals are mutually exclusive. Thus, all 'between-goals' terms in the CTA-RST inhibition matrix are set equal to 1.

The 3-by-3 on-diagonal sub-matrices of 'within-goals' terms (which represent RST-system actions directed toward the same goal) are a more complicated matter. The simplest possible model consistent with RST theory would set BAS- or FFFS-on-BIS terms equal to 0, BAS-on-FFFS terms equal to 1, the BIS-on-BAS term equal to 1, and the BIS-on-FFFS term equal to -1 (representing a negatively inhibitory or excitatory effect of BIS on FFFS). I did my best, however, to remain true to Pickering's formulation, ultimately settling on

$$I = \begin{bmatrix} k_{3A} & k_{4A} * (M_I + t_A) & k_{2A} * (M_I + t_A) \\ k_{4F} * (M_I + t_F) & k_{3F} & k_{2F} * (-M_E + t_F) \\ 0 & 0 & k_{3I} \end{bmatrix} \tag{3.8}$$

for the CTA-RST inhibition matrix's on-diagonal submatrices, where (for brevity's sake) subscripts A, F, and I denote terms (tendencies, t, and constants, k_n) referring to the BAS, FFFS, and BIS, respectively.

Here, several changes have been made to Pickering's original equations. First, Pickering's k_3 terms were multiplied by tendencies, not actions. Although I initially tried including the k_3 terms in the Excitation matrix, thus retaining their dependence on tendencies, while leaving the within-goals' submatrices' diagonal terms equal to zero, the results were much less clean when the model was framed in this manner. Second, the inhibitory effect of the BIS's action on the FFFS's action, represented here by $k_{2F} * (-M_E + t_F)$, is more faithfully translated as $k_{2F} * (-M_E + a_F)$; it is, as the simulation runs, the FFFS action, not the FFFS tendency, which affects the subsequent change in FFFS action by BIS action. Altering the BIS-on-FFFS term in the equation above from dependence on action to dependence on tendency was done for consistency's sake.

One can choose the maximum inhibitory and excitatory values for the system in question so as to make systems' actions either incompatible (one system's action entirely inhibits another's; cell value is 1 or greater) or compatible (actions may occur simultaneously; cell value is less than 1). Pickering's values for the second and fourth k constants are all equal to 0.20 and his value for M_I is 0.50; however, using these values makes BAS and FFFS actions compatible. Therefore, I have set the second and fourth k constants to 2, leaving M_I at 0.50. Otherwise, my inhibition matrix is similar to Pickering's analogous terms. Again, $M_E = Max_E = 1$, and I retained the value .05 for system-on-system inhibitions because Pickering's model also sets $k_3 = .05$ for all systems.

3.3.7 Expectancies and Reinforcements

Expectancies and reinforcements are two features of the latest version of the CTA program that allow the latter to learn, albeit in a rather crude fashion. The simulated participant begins with a number that describes the extent to which the corresponding action is expected to be reinforced; these are set by default to 1, and the resulting matrix is diagonal (i.e. only reinforcements contingent on the action in question will influence the expectancies of reward or punishment for that action). During both simulation development and testing, I used the default CTA settings for expectancies. Note that expectancies alone do not affect whether or not the model 'learns;' this is, instead, determined by the values assigned to the reinforcement matrix.

If all reinforcements are set to zero, then the model does not learn. Learning occurs when, assuming that the simulation has just switched from one action to another, the instantaneous value that the new action assumes is multiplied by the reinforcement value and this product is added to the initial expectancy value, creating a revised expectation. Positive reinforcement values constitute rewards, which increase expectancies and action frequencies, whereas negative reinforcement values function as punishments that decrease action frequencies. Reinforcements set to zero do not change the expectancy values; thus, the model does not learn. I set all reinforcements equal to zero during simulation development and testing.

3.4 Operational Definitions of Personality and Affect

Arguably, the most critical decisions to be made in CTA-RST's development involved defining the independent and dependent personality and affect variables modeled by its

raw tendency and action output. These variables fell into two broad categories; namely, trait variables and state variables. Recall here, too, that the trait variables in the CTA model are its weights (e.g. in the **S** and **C** matrices), whereas CTA's state variables are its outputs; i.e. its tendencies (**t**) and actions (**a**).

3.4.1 Traits

The studies I simulated invoked both familiar personality trait variables like extraversion and ad hoc trait affect variables like trait energetic arousal. In addition, one study considered state personality variables (extraversion and neuroticism). In order to operationally define state personality variables, I followed the same exploratory procedure as I did in constructing state affects. Indeed, as noted in Chapter 5, my operational definitions of state personality were built from components defined in the *States* section, below. Because I decided to assume that states were more alike than traits, regardless of whether they are ascribed to affect or personality, state extraversion and state neuroticism are constructed from the statelike characteristics of tendencies and actions described below. Similarly, traits are more alike than states, regardless of whether they are ascribed to affect or personality; thus, I looked at a slightly different set of components when operationally defining trait personality (e.g. extraversion) or trait affect (e.g. trait energetic arousal).

One way in which to define traits is to use linear combinations of standardized reinforcement sensitivity weights (i.e. BAS_w, BIS_w, $FFFS_w$) and consummatory sensitivity weights (BAS_c, BIS_c, $FFFS_c$); these values are noted in Table 3.3 below the functions used to build state constructs. Again, there are many possible combinations of these that one might reasonably believe could create a working definition of, e.g.,

extraversion (E_T). Consider, for instance,

$$E_T = BAS_w - FFFS_w \tag{3.9}$$

versus

$$E_T = BAS_w - FFFS_w - BIS_w \tag{3.10}$$

Indeed, if high consummatory sensitivity to reward cues (perhaps one definition of 'impulsivity') plays a role, then we could also define extraversion as

$$E_T = BAS_w - FFFS_w - BIS_w + BAS_c \tag{3.11}$$

The fact of the matter is that any or all of these definitions could be correct, and likely *will* be, depending on how a given study measures extraversion and how particular participants interpret those measures. Moreover, linear combinations of sensitivity weights are not the only possible definitions of traits; it is possible that, at least in some instances, self-reported traits are the means of states. Of course, these would be self-assessed means of states that one can remember, given one's current mood – but means might still offer a better way to model traits, sometimes, than weights do.

3.4.2 States

CTA-RST's operational definitions of state affect and personality assume that the latter are functions of several characteristics of either actions or tendencies. In line with the

analogy to Newtonian mechanics on which CTA is based, action-versus-time and tendency-versus-time outputs can be likened to physical position-versus-time data. As such, actions and tendencies have, at any given time t, associated 'position-like,' 'velocity-like,' and 'acceleration-like' quantities. Mathematically, these quantities correspond to the action or tendency function's value, first derivative, and second derivative at t, respectively. Actions' and rising or falling tendencies' durations may also influence affect. Here, I outline the notation to be used in operational definitions of affect and personality states that I constructed for each of the three studies I modeled. The equations that constitute the definitions themselves may be found in the *Assumptions* sections of Chapters 4 through 6.

All state variables were necessarily assessed within a pre-specified time frame (e.g. 'at this moment' or 'within the last thirty minutes'). If a study's state affect measure assessed affect 'at this moment' (or if it did not specify a time frame), the value of the function corresponding to that affect was calculated over several time intervals; the interval that yielded the best results was retained in the definition. If the affect measure specified a time frame (the last hour, the last thirty minutes, etc.), then the affect function's value was calculated over that interval. Modeled instances of 'trait affect' (i.e. trait EA, TA, UA, and PA) were treated somewhat differently, as discussed above.

I also assumed that actions and tendencies were differently related to the types of state-affect or state-personality constructs arising from the activation of a given system. The intensity (i.e. 'position') and duration of a BAS action are positively related to positive affect, while the same variables that characterize a FFFS or BIS action are likely to be positively related to negative affect. Actions proceed in parallel with tendencies' falling velocities, and may be indicative of affects that, in one way or another, address

tendencies' satisfactions or consummations. While actions proceed, however (and while their associated tendencies drop off), other tendencies increase. Rising tendencies, therefore, may reflect activating or motivating (positive or negative) affects (the steeper and longer the rise, the more intense the affect). This assumption is in agreement with Carver's (2009) suggestion that negative affect is a signal that one's current action needs to change.

Tendencies for different systems may rise and fall simultaneously, but approach and avoidance *actions* relative to a particular goal are mutually exclusive (i.e. one cannot simultaneously approach and avoid a goal object; however, one may avoid one goal object while approaching another), which is to say that BAS and FFFS actions relative to a single goal cannot overlap (the only actions allowed to overlap other systems' actions within goals are BIS 'actions,' which merely inhibit ongoing overt approach or avoidance actions). Therefore, one guide to whether it is a tendency's or an action's properties that better reflect a given emotion is whether the latter is observed to be positively or negatively correlated with other emotions. Again, operational definitions of each simulated study's psychological constructs will be given in the corresponding chapter, and a summary of all constructs will be given in the concluding chapter. For now, it will be most useful to discuss the notation used throughout the remainder of this thesis.

Actions' Positions

Here, $\overline{BAS}_{Ap}$ is the average intensity (i.e. position) of the current goal's BAS action over the measurement interval.

$$\overline{BAS}_{Ap} \tag{3.12}$$

The notation is similar for other systems; i.e.

$$\overline{FFFS}_{Ap} \tag{3.13}$$

and

$$\overline{BIS}_{Ap} \tag{3.14}$$

Thus, from here on out, I will use only the BAS form to illustrate each of my notation conventions.

Actions' Durations

The average time spent doing an approach action over a time interval t is denoted

$$\overline{BAS}_{At} \tag{3.15}$$

Average time calculations utilize a binary 'counter' function (i.e. 1 is added for every iteration at which the action is nonzero; 0 is added for every iteration at which it is zero); thus, these quantities are always between 0 and 1 (inclusive).

Tendencies' Rising Velocities

The average rising velocity (i.e. first derivative) of the current goal's BAS tendency is given by

$$\overline{BAS}_{Tvr} \tag{3.16}$$

Again, when a tendency is rising, that means that it has not yet been realized in action (i.e. as soon as the associated action begins, the associated tendency begins to fall). Intuitively, therefore, any affect associated with rising tendencies ought to have an 'unfulfilled,' if not a purely negative quality. One hypothesis, then, is that rising BAS tendencies, for example, reflect what we call 'wanting' (or restlessness, or dissatisfaction). Furthermore, the rising slopes of a system's tendencies are strongly determined by that system's instigating cue sensitivity, as well as by corresponding cues in the environment.

Tendencies' Falling Velocities

The falling velocity of the current goal's BAS tendency is denoted

$$\overline{BAS}_{Tvf} \tag{3.17}$$

Falling tendencies, which accompany action initiation (and therefore, ultimately, goal fulfillment), seem intuitively as though they ought to be associated with affective states that reflect consummation of or progress toward (or away from) a goal. Again, these affective states will depend for their valence upon which system generates them (e.g. if BAS, one might imagine that valence would be strongly positive at first, but becoming weaker as satiation increases; if FFFS, valence would be strongly negative at first, and becoming weaker as avoidance becomes more successful). Because actions are reflected in falling tendencies, affects calculated based on the latter ought to exhibit properties

similar to those calculated based on the former (with one exception; i.e. a goal's associated approach and avoidance actions are mutually exclusive, whereas the corresponding tendencies are not).

Tendencies' Rising and Falling Durations

A tendency's average rising duration is given by

$$\overline{BAS}_{Ttr} \tag{3.18}$$

whereas its falling duration is given by

$$\overline{BAS}_{Ttf} \tag{3.19}$$

Durations, whether for tendencies or actions, provide additional information about the latter that are not reflected in intensities and velocities; thus, it should sometimes be useful to consider their role in affective experience, as well. The functions discussed above are listed in Table 3.3 for easy reference.

Defining Affective and Personality States

Obviously, there are many possible ways in which to construct affects and personality states from the list of basic components above (and this is leaving aside the fact that I have not listed other possible basic components; for instance, tendencies' accelerations, rising and falling action velocities, etc.). Moreover, as stated, broadly-defined emotions like general positive and negative affect could be easily created by summing (or

multiplying, etc.) certain sets of the terms defined above. In fact, I followed a loose, intuitive, highly exploratory procedure in order to create the operational definitions of affect outlined in the chapters to come. Essentially, I explored the parameter space of possible affect definitions while running each study; the definitions I ultimately selected for a given study were those that (a) yielded results that agreed closely with those of the original study, (b) yielded results that agreed closely with other original studies' results when the same constructs were used in more than one study, and (c) seemed simplest and most intuitive.

Table 3.3: Notation for (BAS) Functions Used in Operational Definitions of Traits and States

Notation for BIS and FFFS functions is similar.

BAS function	Notation
Average position of goal's action over time interval t	$\overline{BAS}_{Ap}$
Average time spent doing goal-directed action over t	$\overline{BAS}_{At}$
Average rising velocity of goal's tendency over t	$\overline{BAS}_{Tvr}$
Average falling velocity of goal's tendency over t	$\overline{BAS}_{Tvf}$
Average time goal's tendency spends rising over t	$\overline{BAS}_{Ttr}$
Average time goal's tendency spends falling over t	$\overline{BAS}_{Ttf}$
Instigating cue sensitivity	BAS_w
Consummatory cue sensitivity	BAS_c

Chapter 4

Study 1 – Varieties of Positive Affect

It's not the kill. It's the thrill of the chase.

Deep Purple, 1984 ('Knocking at Your Back Door')

4.1 Background

In five different studies, Smillie et al. (2012) examined the effect of trait extraversion on mood manipulations' efficacy. For simplicity's sake, I have chosen only to model their Experiments 1 and 4, because these are the two manipulations that best and most concisely capture the story that the authors intended to tell. Experiment 1 related extraversion and positive and negative affect changes to positive, negative, and neutral mood induction procedures, while Experiments 2 and 3 replicated these results using different mood-inducing tasks. Experiment 4 replicated and extended Experiments 1-3; now, the affect-altering tasks were intended to produce appetitive, pleasant, or neutral moods. Experiment 5 conceptually replicated Experiment 4. The relevant details of Experiments 1 and 4 are reviewed below.

4.1.1 Experiment 1

Experiment 1's participant sample was comprised of 129 undergraduate psychology students randomly assigned to each of the three mood induction groups (N = 43 per condition). Participants' extraversion and baseline positive and negative affect were assessed prior to the manipulation, which involved written vignettes adapted from Larsen and Ketelaar (1991) and Mayer, Allen, and Beauregard (1995). Each participant read two vignettes intended to induce the same mood (positive, negative, or neutral, depending on condition), spending four minutes on each. Instructions advised participants to imagine the scenarios as vividly as possible, and music consonant with each mood type played in the background throughout the task. Afterward, participants again self-reported positive and negative affect.

Extraversion was assessed using the 23 'yes-no' items from the Eysenck Personality Questionnaire–Revised Extraversion scale (EPQ-R, H. J. Eysenck & Eysenck, 1991) and the 10 Extraversion items (scored on a 5-point scale) from the International Personality Item Pool–Five Factor Model (IPIP-FFM Goldberg, 1999). EPQ-R Extraversion is comprised in large measure of sociability items (Rocklin & Revelle, 1981); however, it contains at least four impulsivity items (i.e. 61, 63, 67, and 69; S. B. Eysenck, Eysenck, & Barrett, 1985, p. 28). IPIP Extraversion, as a shorter measure consisting of items like ' Make friends easily' and (reverse-keyed) 'Don't talk a lot,' seems to be a measure exclusively of sociability rather than of other constructs, like impulsivity and positive affect, with which extraversion is frequently associated. Both extraversion measures achieved acceptable alpha reliabilities (EPQ, .80; IPIP, .79).

The Positive and Negative Affect Scale (PANAS Watson, Clark, & Tellegen, 1988) was used to assess pre- and post-manipulation levels of affect. The PANAS instructs

respondents to report the extent to which, on a 5-point scale, they are currently experiencing the emotions that its items describe. Specifically, the PANAS Positive Affect (PA) scale consists of 10 items that seem to measure activated positive affect or energetic arousal: *interested, excited, strong, enthusiastic, alert, inspired, determined, attentive, active*, and *proud*. The PANAS Negative Affect (NA) scale has the same number of items and is similarly scored; however, its content is less uniform: *distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery*, and *afraid*. The PA and NA scales were reliable both before (α(PA) = .91, α(NA) = .82) and after (α(PA) = .93, α(NA) = .83) the procedure.

The mood manipulations themselves deserve some attention here, as well. Positive mood vignettes were drawn from the Mayer et al. (1995) 'Happy' mood-induction set, which consists of eight scenarios. Briefly, these describe the following events: getting a good new job, enjoying a warm Saturday in winter, buying a winning lottery ticket, having a nice dinner with a friend, leaving work early to get ice cream with friends, spending a relaxing day in a natural setting, deepening a friendship during a coffee date, and enjoying a surprise birthday party. Participants in the positive mood condition listened to the Mazurka from *Coppelia*, by Delibes. The neutral mood induction procedure, musically accompanied by Dvorak's *The New World Symphony*, seems to have drawn its vignettes from a pool of only two; i.e. those included in the procedure described by Larsen and Ketelaar (1991). One of these describes shopping in a grocery store for one's dinner, while the other instructs one to imagine that one is in a car on the highway.

Although Smillie et al. (2012) do not make it entirely clear, negative mood vignettes seem to have been adapted from Mayer et al.'s (1995) 'Sad' scenarios. This set also contains descriptions of eight events; of these, three involve death (one of a favorite teacher, two of

beloved pets) and two involve terminal illness (of a relative or a young friend). The remaining three vignettes encouraged one to imagine that one has endured a romantic break-up, that one's best friend is moving far away, or that everybody one knows has forgotten one's birthday. Meanwhile, a half-speed rendition of Prokofiev's *Russia under the Mongolian Yoke* played in the background.

4.1.2 Experiment 4

Experiment 4 recruited 107 undergraduate psychology students to participate; these were randomly assigned to appetitive (N = 33), pleasant (N = 33), and neutral (N = 41) mood induction conditions. Again, Extraversion (this time, only the EPQ-R extraversion measure was used; $\alpha = .82$) and positive affect levels were assessed at pretest. This time, however, positive affect was broken into two components; namely, pleasant affect and activated positive affect. During the manipulation procedure, subjects received two written vignettes intended to induce the same mood and were instructed to visualize each for four minutes. Only six scenarios in total were constructed. Musical accompaniment was provided during the visualization task, and pleasant and appetitive positive affect were measured a second time afterward.

Both pleasant and activated positive affect were measured using scales from the UWIST Mood Adjective Checklist (UMACL, Matthews, Jones, & Chamberlain, 1990), all of which ask how the participant feels 'at this moment' and are scored on a 4-point scale. Specifically, pleasant affect was operationalized using the UMACL's Hedonic Tone (HT) scale, consisting of eight items: *cheerful, happy, contented, satisfied* and (reversed) *dissatisfied, depressed, sad*, and *sorry*. Note that the items *happy* and *cheerful* also loaded on the UMACL Energetic Arousal (EA) factor, which was used to assess activated

positive affect. Alpha reliabilities for the HT scale were good at both pre-test (α = .86) and post-test (α = .89).

The UMACL EA scale contains the same number of items as the HT scale; these are as follows: *alert, vigorous, active, energetic* and (reversed) *unenterprising, sluggish, tired*, and *passive*. Of these, *tired* also loaded on the HT scale. Alphas for this scale were slightly lower than for the HT scale at pre-test (α= .80) and post-test (α = .82).

As stated, each of the mood induction conditions was comprised of the same two scenarios. The Pleasant condition included one vignette regarding relaxing in the sun during a beach vacation; the second vignette was not described (and, since none of Mayer et al.'s 'Happy' mood scenarios included a beach vacation, it is possible that the other 'Pleasant' vignette was also not one of Mayer et al.'s). Holst's *Venus* from *The Planets* was played for participants in the Pleasant condition.

In the Appetitive and Neutral conditions, mood induction visualizations were (either probably or definitely, respectively) subsets of those used in Experiment 1. In particular, the Appetitive condition utilized an adapted version of Mayer et al.'s instant-win lottery ticket scenario (1995); the undescribed second scenario may have also been culled from Mayer et al.'s 'Happy' stories. The *Waltz of the Flowers* from Tchaikovsky's *Nutcracker Suite* accompanied this condition. Participants in the Neutral condition, on the other hand, listened to the Largo movement from Dvorak's *The New World Symphony* (presumably the same 'Neutral' music as in Experiment 1) and visualized (just as in Experiment 1's Neutral condition) grocery shopping for their dinners and driving in a car on the highway (Larsen & Ketelaar, 1991).

4.2 Assumptions

4.2.1 Participants

A simulation affords certain advantages over a real-life experiment; I availed myself of at least one of these in allotting the same number of participants to each of my experimental groups, for the sake of statistical power. I selected a number of participants-per-condition (N = 40) that was not greatly different from the average number in Experiments 1 and 4. In addition, I ignored participants' ages and genders, as CTA-RST as yet includes no means of simulating these constructs.

4.2.2 Time-scale

The choice of the number of iterations to assign to a simulated 'minute' is, of course, an arbitrary one. Ultimately, I chose CTA-RST's time-scale based on both pragmatic and empirical considerations. In particular, I let 900 iterations = 30 minutes (1 iteration = 2 seconds) because this was, to the nearest hundred, the largest number of iterations that (a) was divisible by 30 and (b) was less than 1000, creating a medium-sized subject-by-experience-by-timestep matrix within the program. Empirically, this choice of timescale did not seem to cause many problems in simulating my three chosen studies' most important findings.

4.2.3 Extraversion

As discussed, the original Experiments 1 and 4 employed two measures of extraversion; these produced rather different results when pitted against one another in Experiment 1. I, on the other hand, simulated only one measure of trait extraversion. Although

determining which CTA-RST settings conform best to each extant measure of extraversion is in itself of interest, to do so is beyond the scope of the current project. Because both Experiment 1 and Experiment 4 used EPQ extraversion, it seemed wisest to equate that measure with my simulated trait extraversion (E_T).

With that in mind, recall that IPIP Extraversion items focus solely on sociability. EPQ Extraversion, on the other hand, also contains several items that assess impulsivity. Because impulsivity is an idea that refers less to an individual's sensitivity to rewards (or insensitivity to punishments) than it does to the quickness with which individuals act on their thoughts, feelings, or desires (or cease to act once they have begun), it is probably an outcome of consummatory sensitivities, which affect action frequency (Revelle & Condon, 2015), rather than of instigating cue sensitivities. Specifically, impulsivity is most likely positively correlated with reward-consummatory sensitivity (BAS_c); it may also be positively correlated with the punishment systems' consummatory sensitivities ($FFFS_c$ and BIS_c). Therefore, I chose to define simulated trait extraversion as

$$E_T = BAS_c + BIS_c + FFFS_c \tag{4.1}$$

Because this choice is itself somewhat arbitrary, I ran six additional simulations, in which extraversion was defined as BAS_c, BIS_c, or $FFFS_c$ alone, or as the sum of each possible pair of the three consummatory weights.

4.2.4 Affect

Experiment 1's measure of 'positive affect' (PANAS PA) is actually a measure of activated affect; thus, I used what I will refer to as my measure of state energetic affect

(EA_S) for the pre- and post-test measures of positive affect in Experiment 1 as well as for the pre- and post-test measures of activated positive affect in Experiment 4. Specifically, EA_S relative to a particular goal at a particular time is the average falling velocity of the BAS tendency for that goal during that time period (in this case, over the 240-iteration, or 8 minute, measurement interval). Because this study experimentally manipulated mood, I looked only at EA_S that pertained to a single goal; i.e. to the goal targeted by the manipulation.

$$EA_S = \overline{BAS}_{Tvf} \tag{4.2}$$

It is difficult to say how much Experiment 1's PANAS measure of negative affect reflects more specific feelings (e.g. tense as opposed to unpleasant affect); therefore, I chose to define state negative affect (NA_S) as the average intensity of FFFS actions over the measurement interval for the target goal.

$$NA_S = \overline{FFFS}_{Ap} \tag{4.3}$$

Note also that this definition later appears in the guises of state neuroticism (in Chapter 5) and state unpleasant affect (in Chapter 6).

Finally, the simulated Experiment 4 measured pleasant positive affect using the average intensity of BAS actions over the measurement interval for the target goal. I refer to this operational definition elsewhere as state positive or pleasant affect (PA_S, Chapter 6), and it will also be used to measure state extraversion (E_S, Chapter 5).

$$PA_S = \overline{BAS}_{Ap} \tag{4.4}$$

4.3 Method: Parameter Settings

In total, 240 simulated subjects (i.e. 40 individuals in each of the 6 conditions) were randomly (seed = 2000) allotted normally-distributed instigating and consummatory cue sensitivities for the BAS, BIS, and FFFS. Within each of the two simulated experiences generated in this study (see below for additional detail), each participant confronted two simulated goals, each of which had its own consummatory and instigating cue settings. Goal 1 was a foreground or 'target' goal and represented the visualization tasks in the original Experiments 1 and 4. Goal 2 served as a background or 'distractor' goal meant to subsume all additional processing, in much the same way as all 'alternative' (i.e. non-target) causes in a causal Bayes net are subsumed in a single node (Fernbach, Darlow, & Sloman, 2011). State affect in this simulation was calculated relative to the foreground goal only. In this regard, this simulation was unlike simulation studies described in later chapters, which calculated state affect over both goals.

4.3.1 Pre- and Post-Test Conditions

All simulated participants encountered two 'experiences;' the first of these corresponded to pre-test (baseline) and the second to post-test. Pre-test experiences were generated randomly; post-test experiences conformed to one of six settings, each of which corresponded to a mood manipulation condition in Experiments 1 or 4. In each experimental condition, instigating and consummatory cues were set to constant values that seemed reasonable in the light of (a) the kind of mood each manipulation was

intended to evoke and (b) the empirical results the manipulations originally produced. Table 4.1 and Figure 4.1 contain all Experience 2 cue parameter settings' exact values; the Figure contains the same information as the Table, but with color-coding that may make the relationships among different conditions easier to grasp. Note here that it proved necessary in disentangling the effects of extraversion on changes in activated (EA_S) and non-activated (PA_S) positive affect to control both the instigating and the consummatory cues in the simulated versions of Experiments 1 (E1) and 4 (E4).

Table 4.1: Instigating and Consummatory Cue Settings for Simulated Mood Manipulations

c_{BAS} = Instigating BAS cue, $c_{c.BAS}$ = Consummatory BAS cue

	Goal 1 (Target)				Goal 2			
Condition	c_{BAS}	$c_{c.BAS}$	c_{FFFS}	$c_{c.FFFS}$	c_{BAS}	$c_{c.BAS}$	c_{FFFS}	$c_{c.FFFS}$
E1: Positive	.95	.50	.05	.05	.50	.50	.50	.50
E1: Negative	.05	.05	.95	.50	.50	.50	.50	.50
E1: Neutral	.05	.50	.05	.50	.50	.50	.50	.50
E4: Pleasant	.75	.05	.05	.50	.50	.05	.50	.05
E4: Appetitive	.95	.25	.05	.95	.05	.95	.05	.95
E4: Neutral	.05	.50	.05	.50	.50	.50	.50	.50

Simulated Experiment 1, Experience 2

Goal 1				Goal 2				
C_BAS	C_C.BAS	C_FFFS	C_C.FFFS	C_BAS	C_C.BAS	C_FFFS	C_C.FFFS	Condition
0.95	0.5	0.05	0.05	0.5	0.5	0.5	0.5	Pos
0.05	0.05	0.95	0.5	0.5	0.5	0.5	0.5	Neg
0.05	0.5	0.05	0.5	0.5	0.5	0.5	0.5	Neu

Simulated Experiment 4, Experience 2

Goal 1				Goal 2				
C_BAS	C_C.BAS	C_FFFS	C_C.FFFS	C_BAS	C_C.BAS	C_FFFS	C_C.FFFS	Condition
0.75	0.05	0.05	0.5	0.5	0.05	0.5	0.05	PI
0.95	0.25	0.05	0.95	0.05	0.95	0.05	0.95	Ap
0.05	0.5	0.05	0.5	0.5	0.5	0.5	0.5	Neu

Figure 4.1: Color-Coded Mood Manipulation Parameters

4.4 Method: Analyses

As was the case in the original studies, both simulated experiments employed the same three sets of analyses. First, correlations among baseline affects and extraversion were obtained. Second, a series of three ANOVAs were run in order to verify that the mood inductions had worked as expected. Third and finally, a three-step moderated multiple regression analysis determined which factors were most predictive of post-manipulation measures of positive affect (either EA_S or PA_S).

Of the three ANOVAs conducted for each experiment, a $2 \times 3 \times 2$ (time by condition by affect) mixed ANOVA first sought to determine whether the conditions differed. Next, a series of six one-way within-subjects ANOVAs revealed which affects changed from pretest to post-test in which conditions. Finally, two between-subjects ANOVAs were run in order to ascertain whether post-test positive and negative affects differed by condition. Tukey follow-up tests were conducted in the event that post-test EA_S, PA_S, or NA_S differed significantly across conditions.

Each of the three steps in the moderated multiple regression (MMR) analysis for each post-test measure of positive affect (EA_S or PA_S) built upon its predecessor; all continuous predictor variables were standardized before modeling. In Step 1, pre-test affect predicted post-test affect. In Step 2, extraversion and contrast dummy variables for the two non-neutral conditions were added to the Step 1 model. Lastly, in Step 3, interactions between extraversion and each contrast variable were tacked onto the model from Step 2. Each additional model was compared to the previous one in order to determine whether its predictions represented an improvement. Additionally, analyses of simple slopes looked at extraversion as the sole predictor of post-test EA_S or PA_S across

conditions. All analysis code for this simulation may be found in Appendix E.

4.5 Results

4.5.1 Descriptive Statistics

Effect sizes (Pearson's r values) based on the original studies' sample sizes, means, and standard deviations are given in Table 4.2; similar effect sizes for the simulation studies' data are given in Table 4.3. Participants in neither the original nor simulated Experiments 1 and 4 differed ($p < .05$) in Extraversion across conditions. All significant pre-post differences in affect within the original experimental conditions were replicated in the simulation. In addition, the pre-post difference in activated positive affect (EA_S) in the simulated Experiment 4's Pleasant condition was significantly different from (lower than) other pre-post differences in EA_S across conditions. Figures 4.4a and 4.4b show correlations between the real and simulated effect sizes of pre-post differences in affect for Experiments 1 ($r = .73$) and 4 ($r = .82$), respectively.

Table 4.2: Trait Descriptives, Effect Sizes of Pre-Post Differences in Affect, Original Experiments

* = differs significantly from others in that row for that experiment

Condition	E1: Negative (43)	E1: Neutral (43)	E1: Positive (43)	E4: Neutral (41)	E4: Pleasant (33)	E4: Appetitive (33)
$E_{T.EPQ}$ Mean:	16.74	16.37	15.83	13.87	14.58	15.27
Standard Deviation:	3.92	4.42	4.43	4.95	6.18	4.43
EA r:	-0.19	-0.05	0.44*	-0.09	0.01	0.66*
95% CI:	-0.47, 0.12	-0.35, 0.25	0.16, 0.65	-0.39, 0.22	-0.33, 0.36	0.41, 0.82
NA r:	0.49*	-0.07	-0.10			
95% CI:	0.22, 0.66	-0.37, 0.24	-0.40, 0.21			
PlA r:				-0.37	0.65*	0.57*
95% CI:				-0.61, -0.07	0.39, 0.81	0.28, 0.76

Table 4.3: Trait Descriptives, Effect Sizes of Pre-Post Differences in Affect, Simulated Experiments

* = differs significantly from others in that row for that experiment

Condition	E1: Negative (40)	E1: Neutral (40)	E1: Positive (40)	E4: Neutral (40)	E4: Pleasant (40)	E4: Appetitive (40)
$E_{T.EPQ}$ Mean:	-0.08	-0.62	0.10	-0.21	0.01	-0.69
Standard Deviation:	1.79	1.45	1.82	1.86	1.50	1.70
EA r:	0.15	-0.07	0.34*	0.29	-0.24*	0.36*
95% CI:	-0.17, 0.44	-0.38, 0.24	0.04, 0.59	-0.03, 0.55	-0.52, 0.07	0.06, 0.61
NA r:	0.46*	-0.23	-0.18			
95% CI:	0.17, 0.67	-0.50, 0.09	-0.47, 0.14			
PlA r:				-0.34	0.50*	0.34*
95%CI:				-0.59, -0.03	0.23, 0.70	0.04, 0.59

4.5.2 Inferential Statistics: Experiment 1

Inferential statistics (i.e. correlations, ANOVAs, and moderated multiple regressions) for original and simulated Experiment 1 data are given in Table 4.4. In the original publication, the inferential statistics reported were commensurate with the types of tests employed; i.e. Pearson's r for correlations, β values for regressions, F and t statistics for ANOVAs, R^2_{ch} statistics for comparisons between ANOVAs, and p values for everything. In order to facilitate comparisons between original and simulated results in this chapter's tables, I have converted all statistics to effect sizes (expressed in Pearson's r values). Asterisks allow for speedy visual apprehension of effects' significance.

Indeed, a quick glance at Table 4.4 reveals that the original and simulated data are highly similar; in fact, calculating the correlation between the table's two columns of r values (and taking the highest possible values whenever r is less than a certain value; i.e. .09 when $r < .09$) yields a strong, positive relationship ($r = .65$), as shown in Figure 4.4c. Comparing effect size results for the six aforementioned alternative definitions of extraversion showed that, when considered across all analyses, results did not depend heavily on the manner in which extraversion was defined (Figure 4.2a shows that correlations ranged from $r = .61$ to $r = .72$). However, if effect sizes were compared only for those analyses that included extraversion (thus removing analyses whose results did

not depend on the trait's definition and therefore were the same across all seven simulations), some definitions appeared to do better than others (see Figure 4.2b). Here, the range of correlations was broader (from $r = .49$ to $r = .80$), and setting E_{EPQ} equal to BIS_c seemed to produce the best results for this experiment (albeit not across experiments). Different definitions of extraversion were somewhat less highly correlated with each other when extraversion-independent analyses were excluded, as well. Nevertheless, these correlations, ranging from $r = .64$ to $r = .98$, indicated that different definitions of the trait still produced highly similar results.

Correlations among simulated pre-test affect and personality were largely in agreement with the original data. Although E_{EPQ} was not significantly positively related to pre-test positive affect in the original study, both the original and simulated relationships are positive, and the correlations' confidence intervals overlap. Similarly, although pre-test positive and negative affect were not significantly negatively related in the original study, both the original and simulated correlations are negative, and their confidence intervals overlap.

Taken together, the original study's ANOVA results for Experiment 1 argued strongly in favor of a successful mood manipulation procedure. The simulated mood manipulation for this experiment was equally successful. The $2 \times 3 \times 2$ mixed ANOVA's three-way interaction was significant, $F(1, 117) = 26.54$, $p < .001$, which implied that both condition and type of affect were relevant in determining pre-post changes in affect. Follow-up ANOVAs that examined post-test affect in each condition separately indicated that, as desired, positive affect (EA_S in the simulation) increased from pre- to post-test in the Positive condition, $F(1, 39) = 5.21$, $p < .05$, while negative affect (simulated by NA_S) increased from pre- to post-test in the Negative condition, $F(1, 39) = 10.43$, $p <$

.01. Neither positive, $F(1, 39) = 0.22$, $p = .64$, nor negative affect $F(1, 39) = 2.10$, $p = .15$, changed significantly from pre- to post-test in the Neutral condition.

Not only did pre- to post-test affects change as expected, affects differed across conditions at post-test in the expected manner, as well. The one-way ANOVA that predicted post-test positive affect from condition was significant, $F(2, 117) = 28.57$, $p < .001$. Tukey post-hoc tests of significance showed that simulated post-test PA was higher among subjects in the Positive than the Negative ($p < .001$) and Neutral ($p < .001$) conditions. Similarly, the one-way ANOVA predicting post-test negative affect from condition was significant, $F(2, 117) = 67.71$, $p < .001$, with simulated subjects in the Negative condition exhibiting greater post-test NA than those in the Neutral ($p < .001$) or Positive ($p < .001$) conditions.

The original moderated multiple regression (MMR) analyses for Experiment 1 indicated that, generally speaking, the effect of mood induction on post-test PA was not moderated by EPQ Extraversion; MMR analyses of simulated data confirmed these findings. At Step 1 of the hierarchical multiple regression (S1 in Table 4.4), post-test positive affect was predicted by pre-test positive affect (a standardized continuous predictor variable) alone. This model was significant, $R^2 = .05$, $F(1, 118) = 6.12$, $p < .05$, and pre-test PA was a significant, positive predictor of post-test PA ($\beta = .002$, $p = .01$).

Simulated EPQ Extraversion and the condition-contrast dummy variables were added to the MMR model at Step 2 (S2 in Table 4.4). Using an ANOVA to compare the Step 1 and Step 2 models showed that the latter significantly improved on the former, $R^2_{ch} = .32$, $F_{ch}(3, 115) = 19.55$, $p < .001$. The Step 2 model itself was, of course, significant, $R^2 = .37$, $F(4, 115) = 16.91$, $p < .001$. As was the case in the original Experiment 1, simulated Extraversion was not a significant predictor of post-test positive affect ($\beta =$ -.001 , $t(115)$

= -1.31, $p = .19$), and neither was the Negative condition contrast (β = -.001 , $t(115)$ = -0.94, $p = .35$); however, the Positive condition contrast did positively predict simulated post-test PA $\beta = .01$, $t(115) = 7.01$, $p < .001$).

Finally, at Step 3, interactions between extraversion and each condition-contrast variable were entered in the model. Like its real-life counterpart, the simulation's Step 3 regression, while significant ($R^2 = .37$, $F(6, 113) = 11.27$, $p < .001$), did not improve upon the Step 2 regression ($R^2_{ch} < .01$, $F_{ch} < 1$, $p = .70$). Specifically, neither the extraversion-by-Positive (β = -.001 , $t(113)$ = -0.73, $p = .47$) nor the extraversion-by-Negative ($\beta = .0001$, $t(113) = 0.11$, $p = .92$) interactions was significant, suggesting that the relationship between E_{EPQ} and post-test PA did not depend upon mood manipulation condition. Analyses of simple slopes looked more specifically at the relationship between extraversion and post-test PA in each condition; results for the simulation showed, as did the original results, that extraversion was not a significant predictor of post-test PA in the Positive (β = -.001 , $t(38)$ = -0.73, $p = .47$), Negative (β = -.0001 , $t(38)$ = -0.24, $p = .81$), or Neutral ($\beta = .0001$, $t(38) = 0.28$, $p = .78$) conditions.

4.5.3 Inferential Statistics: Experiment 4

Inferential statistics for original and simulated Experiment 4 data are given in Table 4.5. Here again, I have converted all statistics to effect sizes (Pearson's r values, with asterisks denoting statistical significance). Again, the original and simulated data are, for the most part, similar in terms of the effects' direction and significance; however, the correlation between the two columns of effect sizes is somewhat lower than it was in Experiment 1; the correlation for all effect sizes, $r = .55$, is shown in Figure 4.4d.

Table 4.4: Results for Study 1, Experiment 1

Correlations, ANOVAs, and Moderated Multiple Regressions (MMRs); * $p < .05$

Analysis	Original Effect Size	95% CI	Sim Effect Size	95% CI
Correlations				
$r(E_{EPQ}$, Pre-NA)	$r = .15$	-.02, .31	$r = .05$	-.13, .23
$r(E_{EPQ}$, Pre-PA)	$r = .10$	-.07, .27	$r = .21$*	.03, .38
r(Pre-NA, Pre-PA)	$r = -.05$	-.22, .12	$r = -.21$*	-.38, -.03
ANOVAs				
3-way interaction, full ANOVA	$r = .31$*	.15, .46	$r = .43$*	.27, .57
Positive: Pre-PA → Post-PA	$r = .44$*	.16, .65	$r = .34$*	.04, .59
Negative: Pre-NA → Post-NA	$r = .49$*	.22, .69	$r = .46$*	.17, .67
Condition → Post-PA	$r = .24$*	.07, .39	$r = .44$*	.29, .58
Post-PA: Positive v. Negative	>*		>*	
Post-PA: Positive v. Neutral	>*		>*	
Condition → Post-NA	$r = .31$*	.14, .46	$r = .61$*	.48, .71
Post-NA: Negative v. Positive	>*		>*	
Post-NA: Negative v. Neutral	>*		>*	
MMR				
S1: Pre-PA → Post-PA	$r = .61$*	.49, .71	$r = .22$*	.04, .39
S1 v. S2	$r = .33$*	.17, .48	$r = .38$*	.22, .52
S2: E_{EPQ} → Post-PA	$r \approx .09$	-.08, .26	$r = -.12$	-.29, .06
S2: Positive → Post-PA	$r = .33$*	.16, .47	$r = .55$*	.41, .66
S2: Negative → Post-PA	$r = .11$	-.06, .28	$r = .09$	-.09, .26
S2 vs. S3	$r \approx .10$	-.07, .27	$r = .06$	-.12, .23
Simple Slopes				
Positive: E_{EPQ} → Post-PA	$r = .16$	-.15, .44	$r = .12$	-.20, .41
Negative: E_{EPQ} → Post-PA	$r \approx -.15$	-.43, .16	$r = .04$	-.28, .35
Neutral: E_{EPQ} → Post-PA	$r \approx .15$	-.16, .43	$r = .05$	-.27, .35

Figure 4.3 shows the comparison of effect size results across the seven alternative definitions of extraversion; as in Experiment 1, these produced rather similar results when all analyses were compared (correlations ranged from $r = .49$ to $r = .54$, as shown in Figure 4.3a). However, unlike the analogous results for Experiment 1, Experiment 4's extraversion-dependent analyses did not differ greatly across the seven alternative operational definitions of the trait (ranging, as may be seen in Figure 4.3b, from $r = .65$ to $r = .75$). Interestingly, correlations among simulated data created by the different definitions did not seem to depend on whether extraversion-independent effect sizes were included or excluded; in the former instance, correlations among alternative models ranged from $r = .69$ to .99, whereas in the latter, they ranged from $r = .61$ to .99. Overall, Figure 4.3 reveals that, for Experiment 4, the 'CTA-RST' definition of E_{EPQ} ($BAS_c + BIS_c + FFFS_c$) performed as well as or better than alternative definitions of the trait.

Correlations among simulated pre-test affect and personality were rather different here than in the original data. While the relationship between E_{EPQ} and pre-test activated affect (EA_S) was much the same in the simulated and real-life data, the confidence intervals for the simulated and real relationships between baseline EA and baseline pleasant affect (denoted 'PlA' in the table, and computed using the PA_S function in the simulation) did not overlap, despite being in the same direction. More peculiarly, the direction of the simulated relationship between extraversion and baseline PlA was negative (albeit nonsignificant), whereas its original counterpart was significant and positive.

Although these difficulties are to some extent the result of the peculiarities of the 120-simulated-person sample used in the Experiment 4 analyses (for instance, the

correlation between pre-test pleasant and activated affect is $r = .30$ across all 240 simulated subjects, with a confidence interval of .18 to .41, which is closer to but still not quite overlapping with the original confidence interval), that's not the whole story. The negative correlation between pre-test PlA and extraversion is not only present in the full 240-person sample, it's significant (r = -.16, 95% CI = -.28, -.03). I will return to this topic with greater thoroughness in this chapter's discussion section.

The original Experiment 4 ANOVA results pointed to an effective mood manipulation procedure; in this case, however, the simulated results were slightly more complicated. Both, however, suggested that Appetitive mood inductions create both EA and PlA, whereas Pleasant mood inductions only create PlA. The 2 × 3 × 2 mixed ANOVA's three-way interaction was significant, $F(1, 117) = 8.06$, $p < .001$, as was the original's. Follow-up ANOVAs testing for significant pre-post changes in affect by condition showed that, as expected, activated affect (EA_S in the simulation) increased in the Appetitive condition, $F(1, 39) = 5.94$, $p < .05$, while pleasant affect (simulated by PA_S) increased in the Pleasant condition, $F(1, 39) = 13.18$, $p < .001$. Additionally, as was the case in the real Experiment 4, pleasant affect increased in the Appetitive condition ($F(1, 39) = 5.24$, $p < .05$), but activated affect didn't increase in the Pleasant condition ($F(1, 39) = 2.47$, $p = .12$). Simulated results also reflected the original's in the Neutral condition, which failed to change activated affect at the $p = .05$ level ($F(1, 39) = 3.51$, $p = .06$), but managed to significantly decrease pleasant affect ($F(1, 39) = 5.10$, $p < .05$).

One-way ANOVAs examining post-test affects across conditions revealed that there were significant differences in both activated ($F(2, 117) = 61.11$, $p < .001$) and pleasant ($F(2, 117) = 52.20$, $p < .001$) affect by condition, which is hardly surprising given the level of control over conditions' parameters afforded by a computational model. Tukey post-hoc

tests' results for the simulated data mirrored the real ones, indicating that simulated post-test EA_S was higher in the Appetitive than in either the Pleasant ($p < .001$) or the Neutral ($p < .001$) conditions, whereas post-test PA_S was only higher in the Pleasant condition when the latter was compared to the Neutral condition ($p < .001$). PA_S in the Pleasant and Appetitive conditions did not differ significantly ($p = .33$).

Original MMR analyses for Experiment 4 demonstrated the interesting finding that extraversion moderates the effect of mood induction on post-test positive affect only when the latter is 'activated' and the former is 'Appetitive.' In other words, extraverts are happier than introverts when they are pursuing rewards, but extraverts and introverts are equally happy when they are consuming rewards. The simulated Experiment 4 results demonstrated a similar interaction; however, there were several notable differences between the real and computationally modeled data.

Two MMRs were conducted for Experiment 4; the first of these predicted post-test activated affect (EA_S), while the second predicted post-test pleasant affect (PA_S). The Step 1 hierarchical multiple regression model predicting post-test EA_S included standardized pre-test EA_S only. Although the original Step 1 model was significant, the simulated model was not, $R^2 = .003$, $F(1, 118) = .38$, $p = .54$, and pre-test EA was not a significant predictor of post-test EA (β = -.001, $p = .54$).

Simulated EPQ Extraversion and the condition-contrast dummy variables were added to the MMR model at Step 2. The Step 2 model could hardly help significantly improving on the Step 1 model, $R^2_{ch} = .55$, $F_{ch}(3, 115) = 46.63$, $p < .001$. Naturally, the Step 2 model was itself significant, $R^2 = .55$, $F(4, 115) = 35.18$, $p < .001$. As in the original Experiment 4, simulated Extraversion was a significant predictor of post-test activated affect ($\beta = .002$, $t(115) = 3.15$, $p < .01$), as was the Appetitive condition contrast ($\beta =$

.01 , $t(115) = 5.37$, $p < .001$); however, the Pleasant condition contrast negatively predicted simulated post-test EA (β = -.01 , $t(115)$ = -6.54, $p < .001$), whereas it was a nonsignificant predictor in the original study.

In Step 3, interactions between extraversion and the condition-contrast variables were added to the MMR model. Like the original, the Step 3 regression was significant (R^2 = .56, $F(6, 113) = 26.63$, $p < .001$) and improved upon the Step 2 regression ($R^2_{ch} = .04$, $F_{ch} = 4.84$, $p < .01$). The coveted interaction effects were also recovered in the simulated data; namely, the extraversion-by-Appetitive contrast was significant and positive (β = .005 , $t(113) = 3.11$, $p < .01$), while the extraversion-by-Pleasant contrast was nonsignificant (β = .002 , $t(113) = 1.35$, $p = .18$). Thus, the relationship between E_{EPQ} and post-test EA_S didn't depend on the pleasant mood manipulation, but it did depend on the appetitive mood manipulation. Analyses of simple slopes confirmed these results, showing that simulated extraversion was not a significant predictor of post-test EA in the Pleasant (β = .002 , $t(38) = 1.88$, $p = .07$) or Neutral (β < .0001 , $t(38)$ = -0.08, $p = .94$) conditions, but was a significant predictor in the Appetitive condition (β = .005 , $t(38)$ = 3.05, $p < .01$).

When similar MMR analyses were conducted that predicted post-test pleasant (PlA, simulated by PA_S) rather than appetitive affect, simulated and original results were once again in agreement. The Step 1 model (post-PlA predicted by pre-PlA) was significant, $R^2 = .08$, $F(1, 118) = 10.17$, $p < .01$, and pre-test PlA significantly positively predicted post-test PlA ($\beta = .20$, $p < .01$). The Step 2 model was significant ($R^2 = .53$, $F(4, 115)$ = 31.86, $p < .001$) and improved on ($R^2_{ch} = .45$, $F_{ch} = 36.08$, $p < .001$) the Step 1 model. Although Extraversion at Step 2 was not a significant predictor of post-test PlA (β = -.03 , $t(113)$ = -.68, $p = .50$), both the Pleasant (β = 1.10 , $t(113) = 9.79$, $p < .001$) and

Appetitive conditions' (β = .88 , $t(113) = 7.84$, $p < .001$) contrast variables were; in all three cases, the same was true of these variables in the original data. Finally, the Step 3 model, which included both Extraversion-by-condition interactions, did not significantly improve on the Step 3 model ($R^2_{ch} = .01$, $F_{ch} = 1.59$, $p = .21$), though it was itself significant ($R^2 = .54$, $F(1, 118) = 21.99$, $p < .001$). Analyses of simple slopes showed that, although extraversion came close to significantly predicting post-test PlA in the Appetitive condition (β = -.18 , $t(38)$ = -2.00, $p = .053$), it was a highly nonsignificant predictor of post-PlA in the Pleasant (β = -.03 , $t(38)$ = -.21, $p = .83$) and Neutral (β = .002 , $t(38) = .89$, $p = .38$) conditions.

4.6 Discussion

For the most part, the findings discussed above are quite promising; the simulated Experiment 1, in particular, produced results that were highly similar to those produced by the original data. Even in the Experiment 4 simulation, the most important results (i.e. the interactions between extraversion and condition in predicting different types of positive affect) were replicated, and all results were largely commensurate in direction and significance with the Smillie et al. (2012) findings.

For Experiment 1, none of the simulation's results disagreed with the original's inasmuch as (a) all significant effects in the original research were reproduced in the simulation, (b) original and simulated effects that were both significant were also in the same direction, and (c) simulated results that were significant where the original results weren't nevertheless had overlapping confidence intervals. In general, Experiment 1 was much easier to simulate than Experiment 4, as might be expected given the greater subtlety of the difference between affects tested by the latter relative to the former.

Table 4.5: Results for Study 1, Experiment 4

Correlations, ANOVAs, and Moderated Multiple Regressions (MMRs); * $p < .05$

Analysis	Original Effect Size	95% CI	Sim Effect Size	95% CI
Correlations				
$r(E_{EPQ}$, Pre-PlA)	$r = .23$*	.04, .40	$r = -.23$*	-.39, -.05
$r(E_{EPQ}$, Pre-EA)	$r = .10$	-.09, .28	$r = .07$	-.11, .25
r(Pre-PlA, Pre-EA)	$r = .61$*	.47, .72	$r = .11$	-.07, .28
ANOVAs				
3-way interaction, full ANOVA	$r = .27$*	.09, .44	$r = .25$*	.08, .41
Appetitive: Pre-EA → Post-EA	$r = .66$*	.41, .82	$r = .36$*	.06, .61
Appetitive: Pre-PlA → Post-PlA	$r = .57$*	.28, .76	$r = .34$*	.04, .59
Pleasant: Pre-PlA → Post-PlA	$r = .65$*	.39, .81	$r = .50$*	.23, .70
Neutral: Pre-PlA → Post-PlA	$r = .37$*	.07, .61	$r = .34$*	.03, .59
Condition → Post-EA	$r = .31$*	.13, .47	$r = .59$*	.45, .69
Post-EA: Appetitive v. Pleasant	>*		>*	
Post-EA: Appetitive v. Neutral	>*		>*	
Condition → Post-PlA	$r = .27$*	.08, .44	$r = .56$*	.42, .67
Post-PlA: Pleasant v. Appetitive	=		=	
Post-PlA: Pleasant v. Neutral	>*		>*	
MMRs				
S1a: Pre-EA → Post-EA	$r = .57$*	.43, .69	$r = -.06$	-.23, .12
S1a v. S2a	$r = .46$*	.29, .60	$r = .54$*	.40, .65
S2a: E_{EPQ} → Post-EA	$r = .21$*	.02, .38	$r = .28$*	.11, .44
S2a: Appetitive → Post-EA	$r = .46$*	.30, .60	$r = .45$*	.29, .58
S2a: Pleasant → Post-EA	$r \approx -.10$	-.28, .09	$r = -.52$*	-.64, -.38
S2a v. S3a	$r = .15$	-.04, .33	$r = .20$*	.02, .37
S3a: E_{EPQ} x Appetitive → Post-EA	$r = .22$*	.03, .39	$r = .28$*	.11, .44
S1b: Pre-PlA → Post-PlA	$r = .59$*	.45, .70	$r = .28$*	.11, .44
S1b v. S2b	$r = .36$*	.18, .52	$r = .49$*	.34, .61
S2b: E_{EPQ} → Post-PlA	$r = .06$	-.13, .25	$r = -.06$	-.24, .12
S2b: Appetitive → Post-PlA	$r = .43$*	.27, .58	$r = .59$*	.46, .70
S2b: Pleasant → Post-PlA	$r = .33$*	.16, .49	$r = .67$*	.56, .76
S2b v. S3b	$r = .12$	-.07, .30	$r = .12$	-.06, .29
Simple Slopes				
Appetitive: E_{EPQ} → Post-EA	$r = .45$*	.13, .69	$r = .44$*	.15, .66
Pleasant: E_{EPQ} → Post-EA	$r = .24$	-.11, .54	$r = .29$	-.02, .55
Neutral: E_{EPQ} → Post-EA	$r \approx -.16$	-.45, .16	$r = -.01$	-.32, .30
Appetitive: E_{EPQ} → Post-PlA	$r \approx .17$	-.18, .49	$r = -.31$	-.57, .002
Pleasant: E_{EPQ} → Post-PlA	$r \approx -.17$	-.49, .18	$r = -.03$	-.34, .28
Neutral: E_{EPQ} → Post-PlA	$r = .23$	-.08, .50	$r = .14$	-.18, .43

Experiment 1's common-sensical cue parameter settings can be partly attributed the fact that I didn't have to explore the simulation's parameter space too exhaustively before I hit on a combination that worked well. My parameter space exploration procedure consisted, roughly, in starting with those parameter settings and operational definitions that I found most intuitively obvious, then making changes from there. I also gave intuitive operational definitions (that were in turn consistent across the studies I simulated in this and the following two chapters) priority over intuitive parameter settings. It is for this reason that the Experiment 4 parameter settings look a little odd, upon close inspection.

However, the simulated Experiment 4's parameter settings are not, I would wager, the only reasons why its results differ more substantially than Experiment 1's from those uncovered in the original research. To be precise, the Experiment 4 simulation's results depart from the real results in four places. In ascending order of troublingness, these departures are that (1) the simulated Pleasant condition contrast significantly negatively predicted post-test activated affect, while the real contrast was non-significantly negative; (2) the simulated baseline pleasant and activated affect are nonsignificantly positively correlated, while the real affects are significantly positively correlated; (3) the simulation's Step 1 MMR revealed that pre-test activated affect did not significantly predict post-test activated affect, while the original relationship was significant and positive; and (4) the simulation's baseline pleasant affect was negatively correlated with its measure of extraversion, while the real study's E_{EPQ} was positively correlated with baseline PlA.

Problem (1) is not very problematic – it is the first of many similar examples that will crop up throughout this dissertation, and is probably attributable to the cleanliness of simulated data relative to real data. Clean data are much more powerful than noisy data,

and one is much more likely to find effects when they are present in the former than in the latter. Problem (2) is also relatively easy to dismiss, although here we must work a little harder because it is the original result that is significant while the simulation's isn't. However, because the original and simulated results are in the same direction, it seems reasonable to suppose that if we were to meddle a bit with the simulation's time-scale (see Chapter 6 for a further discussion of the manner in which calculations made over fewer iterations tend to produce more negative correlations than those made over more iterations), then we might easily turn up a correlation more comparable to (and that would have a confidence interval overlapping that of) the original. As for Problem (3), I'll be discussing it further in Chapter 7.

Problem (4) is the most vexing and, I would argue, the only problem that seriously supports the contention that my operational definition of simulated EPQ Extraversion is incorrect. As noted previously, I chose to define E_{EPQ} as

$$E_T = BAS_c + BIS_c + FFFS_c \tag{4.5}$$

This function is constructed entirely from consummatory rather than instigating cue sensitivities, which one might quite legitimately argue flies in the face of the common wisdom that extraversion is a measure of an individual's sensitivity to (instigating) cues for reward (i.e., of BAS_w, not BAS_c). If nothing else, and even if EPQ extraversion contained more than four (out of 23) items assessing impulsivity (which I claim is related more strongly than sociability to consummatory rather than instigating cue sensitivities) as opposed to sociability, any CTA-RST measure of extraversion ought to include a positive BAS_w term on top of whatever else it might include, right?

To this query, I would respond with three counterarguments. First (and least convincing), the E_T defined above worked better in recreating Experiment 1 and 4's original results than did *any* model of extraversion I tried that incorporated a positive dependence on BAS-instigating-cue sensitivity (including a model that consisted of nothing but BAS_w). In particular, models of E_{EPQ} that included a positive BAS_w term seemed more likely in Experiment 1 to yield a significant increment in prediction when moving from the Step 2 to the Step 3 MMR, which seemed largely to be due to the fact that the interaction between extraversion and the positive condition contrast had become significant and positive (the original result was nonsignificant and positive).

In Experiment 4, the differences between observed and simulated results were more numerous when BAS_w was added to the definition of E_{EPQ}, especially in the MMR predicting pleasant affect. Six changes in particular were notable. First, the analysis of simple slopes predicting post-test activated affect in the Pleasant condition became significant and negative (the original Experiment 4 result was nonsignificantly positive). Second, the increment in prediction between Steps 2 and 3 of the MMR predicting post-test pleasant affect became significant; both interaction terms in the Step 3 model also became significant and positive. The analyses of simple slopes predicting post-test pleasant affect in the Pleasant and Appetitive conditions also became significant and positive (analogous relationships in the original Experiment 4 were nonsignificantly negative and positive, respectively).

Had all the significant simulated effects generated by extraversion defined as a positive function of BAS_w been in the same direction as their real, nonsignificant counterparts, I would have felt more comfortable allowing the term to remain in the function (in Chapter 6, for instance, I retain a set of simulated state and trait constructs that produce more

significant effects than were found in the original data; however, with only one exception, the simulated results are in the same direction as the real, nonsignificant findings). Here, only about half of the differing results are in the same direction as the originals, which suggests that making simulated E_{EPQ} depend on BAS_w does more than simply increasing the power to detect effects.

That said, my second and third arguments in favor of advancing a definition of E_{EPQ} bereft of BAS_w are less redolent of opportunistic dust-bowl empiricism, though the third is stronger than the second. The second reason is simply this: who is to say that what we mean when we say 'reward cue sensitivity' in the context of self-reported sociability and impulsivity is necessarily the same as the BAS_w parameter in the CTA-RST model? Yes, it is intuitively appealing to identify the two, but, ultimately, whether the identity is correct is an empirical question. Of course, the weakness in this argument is, here as elsewhere in this thesis, that I almost certainly *did not* exhaust the parameter space, despite the fact that my exploration of it felt thorough (if not exhaustive, particularly for this chapter's simulations), and as a result I may simply have failed to encounter the correct combination of cue parameters and operational definitions of affect and extraversion that would, in fact, show that construing E_{EPQ} as a function of BAS_w can produce a faithful replication of the real Experiments 1 and 4.

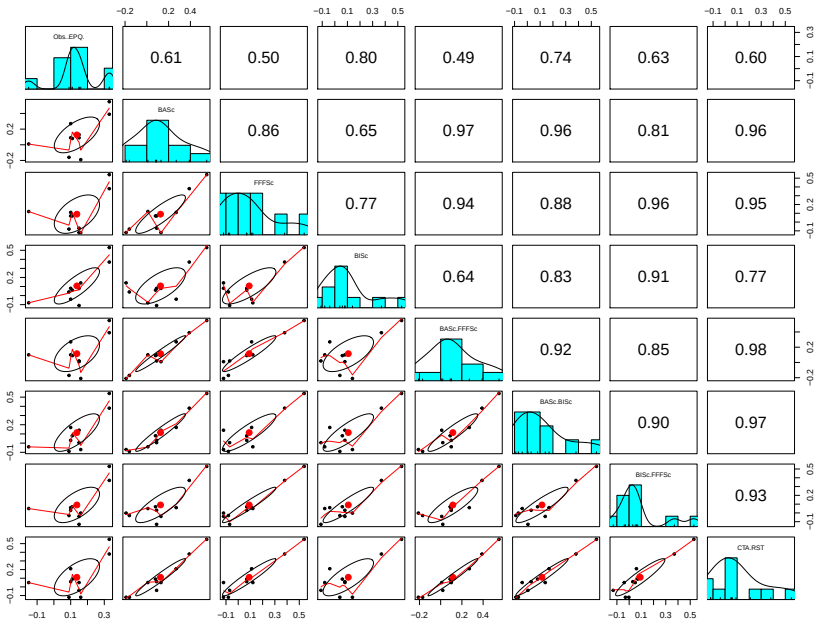
The third argument I would offer in favor of identifying E_{EPQ} with E_T is that the latter is a good-enough, but not perfect, model of the former (this idea will recur in later chapters). That is to say, E_{EPQ} and other self-report measures of Extraversion may indeed depend on something like my BAS_w, but in differing degrees that are not immediately obvious on the basis of the relative numbers, say, of impulsivity and sociability items in the various measures. Assuming that BAS_w is more closely related to

sociability than BAS_c, and vice versa for impulsivity, and although E_{EPQ} contains a greater number of sociability items than impulsivity items, it is possible that the sociability items are less *informative* than the impulsivity items, which would in turn imply that simulated functions of E_{EPQ} should depend more on BAS_c than on BAS_w. This observation brings us back to the question of what functions best represent what measures and, because I have deemed this question beyond the scope of the current research (preferring, as I do, to use unitary weights in all operational definitions of traits and states), I am forced to leave the issue as a subject for future investigations.

Nevertheless, the comparisons made in this chapter between CTA-RST's E_T and six alternative definitions of E_{EPQ} containing only consummatory weight terms (i.e. BAS_c, $FFFS_c$, BIS_c, $BAS_c + FFFS_c$, $BAS_c + BIS_c$, and $BIS_c + FFFS_c$) go some little way toward answering the question of how sensitive CTA-RST and this chapter's Experiments 1 and 4 are to differing, yet similar, models of weight-based personality traits. Combining effect size results for correlational, ANOVA, MMR, and simple slopes analyses for the observed and simulated Experiments 1 and 4, I was able to obtain the correlation matrices shown in Figure 4.5. Specifically, Figure 4.5a shows the correlations among all such analyses, while Figure Figure 4.5b excludes those analyses that were extraversion-independent. Correlations between observed and simulated data are higher in the latter instance than in the former (ranging, respectively, from $r = .65$ to .72 as opposed to $r = .53$ to .59); however, correlations between original and simulated results do not vary much across the seven operational definitions of E_T. The final version of E_T, $BAS_c + BIS_c + FFFS_c$, was the best performer within extraversion-dependent analyses across studies ($r = .72$), but at least one alternative ($E_T = BIS_c$) is not far behind ($r = .71$) and, indeed, this simpler definition produces a slightly higher correlation ($r = .59 > r = .58$) when all analyses are considered together.

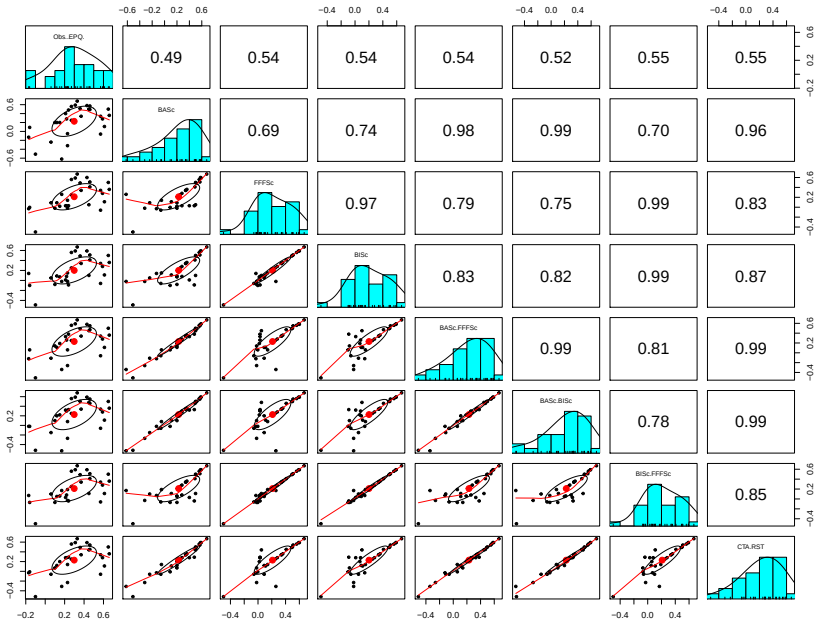
In any case, the fact that (EPQ) extraversion can be modeled so well by consummatory, as opposed to instigating, sensitivities is potentially a very important result for the future of personality theory. In the past, those who have attempted to explain traits in terms of reward and punishment sensitivities have neglected to consider the possibility that there might be more than one kind of biologically-based individual difference in reinforcement sensitivities; the focus has, for the most part, been on the excitatory value of the rewarding or punishing cues that (in the language of the CTA model) instigate approach or avoidance tendencies' growth. As discussed in Chapter 3 and as supported by the evidence presented above, however, individual differences in the satiating effects of actions on tendencies (i.e. consummatory sensitivities) are also good candidates for personality traits' biological underpinnings. Indeed, this chapter's results imply that distinguishing between instigating and consummatory reinforcement sensitivities could provide the key to unraveling several longstanding mysteries within personality psychology, including the problem of how impulsivity is related to extraversion, neuroticism, and other 'big' structural traits.

(a) Correlation between real and simulated effect sizes of all ANOVA, MMR, and simple slopes results.

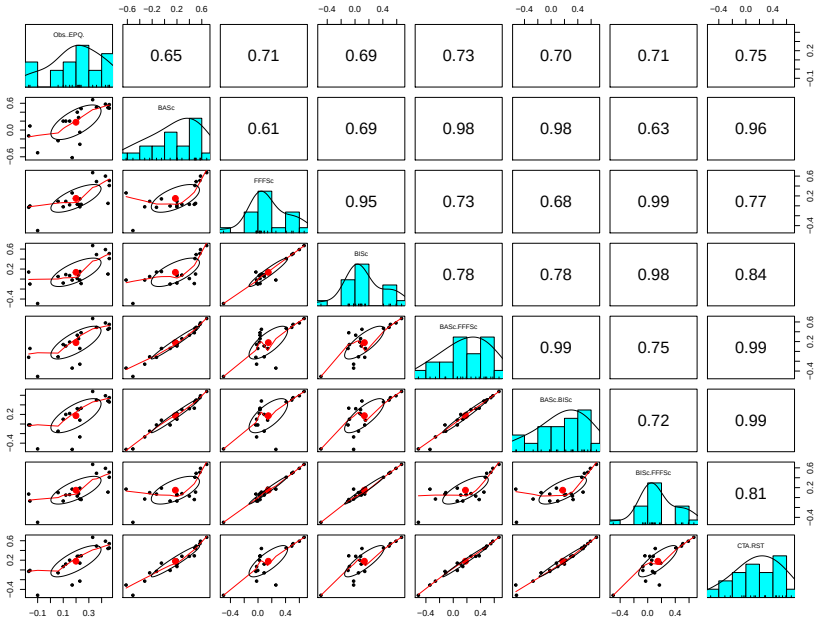


(b) Correlation between real and simulated effect sizes of only those ANOVA, MMR, and simple slopes results that include extraversion.

Figure 4.2: Experiment 1: Comparison of effect sizes of original and simulated data employing different definitions of extraversion.

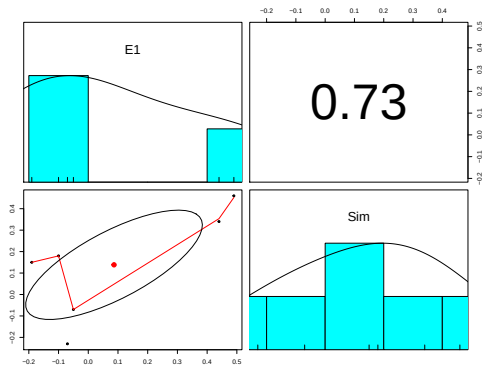


(a) Correlation between real and simulated effect sizes of all ANOVA, MMR, and simple slopes results.

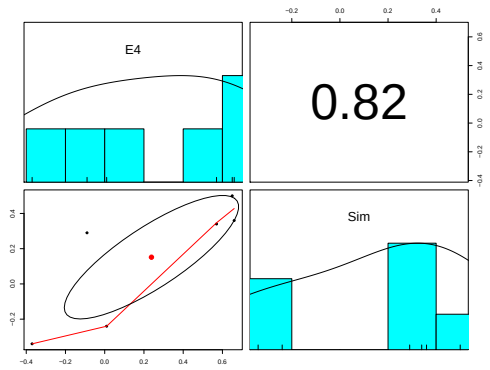


(b) Correlation between real and simulated effect sizes of only those ANOVA, MMR, and simple slopes results that include extraversion.

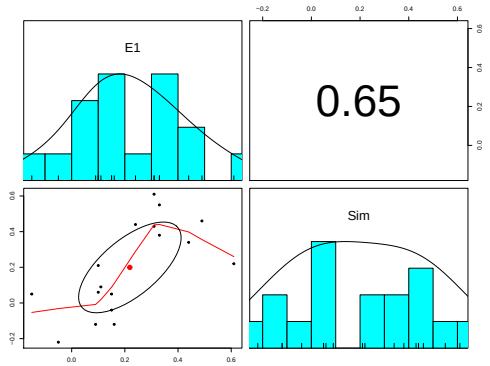
Figure 4.3: Experiment 4: Comparison of effect sizes of original and simulated data employing different definitions of extraversion.



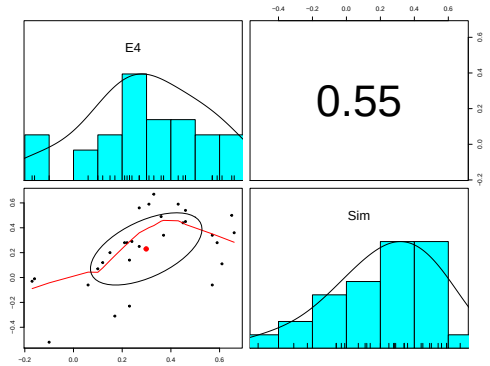
(a) Correlation between real and simulated effect sizes of differences between pre-test and post-test affect for all affects and all conditions in Experiment 1



(b) Correlation between real and simulated effect sizes of differences between pre-test and post-test affect for all affects and all conditions in Experiment 4

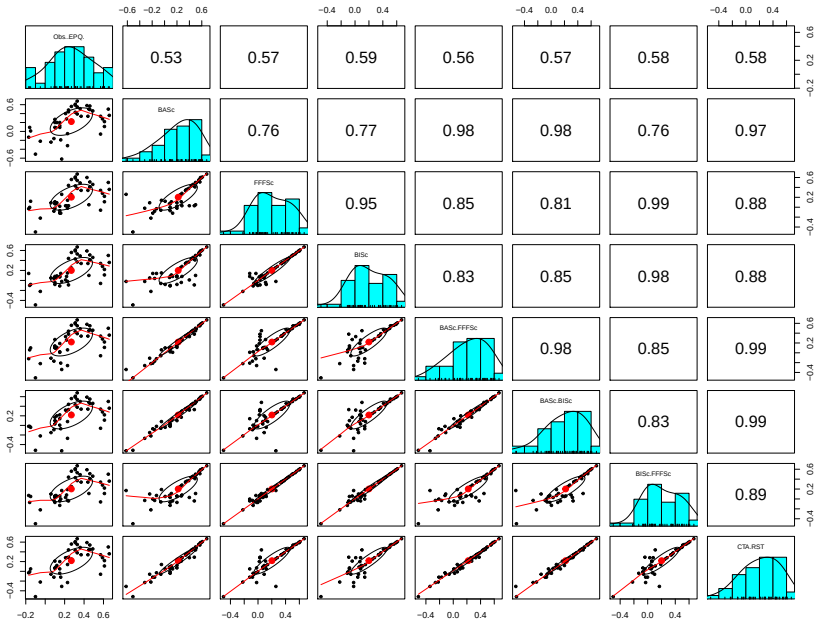


(c) Correlation between real and simulated effect sizes of the correlations, ANOVAs, and moderated multiple regressions (MMRs) for Experiment 1

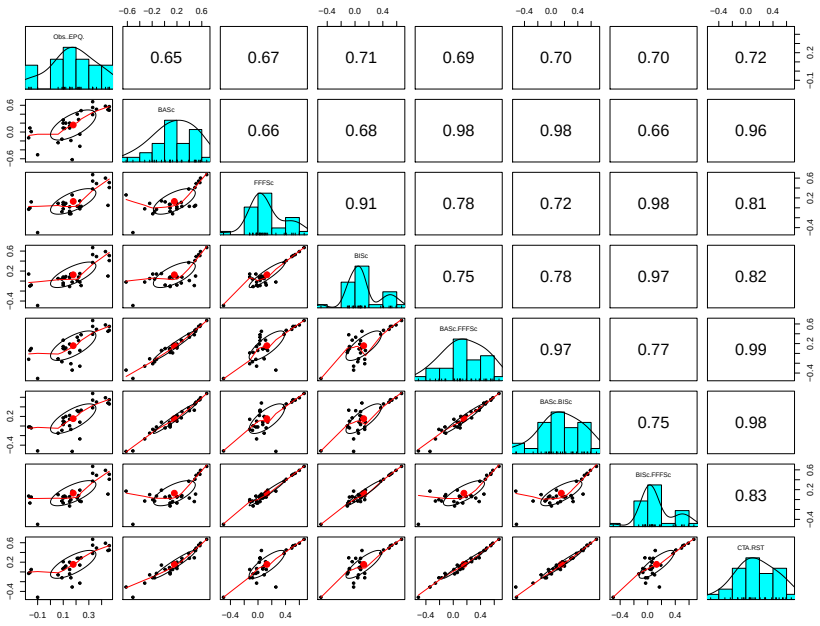


(d) Correlation between real and simulated effect sizes of the correlations, ANOVAs, and MMRs for Experiment 4

Figure 4.4: Correlations between effect sizes of simulated and original results.



(a) Correlation between real and simulated effect sizes of all ANOVA, MMR, and simple slopes results.



(b) Correlation between real and simulated effect sizes of only those ANOVA, MMR, and simple slopes results that include extraversion.

Figure 4.5: Experiments 1 and 4: Comparison of effect sizes of original and simulated data employing different definitions of extraversion.

Chapter 5

Study 2 – Act Velocity and Affect

> When you are courting a nice girl an hour seems like a second. When you sit on a red-hot cinder a second seems like an hour. That's relativity.
>
> Albert Einstein

5.1 Background

Wilt et al. (2016) used mediation analyses on experience-sampling data to show that differences in the rate of progress toward goals can explain the relationship between reinforcement-sensitivity-related personality states and affects. Specifically, increases in velocity toward an approach goal partially accounted for the state Extraversion-state positive affect link, whereas decreases in velocity toward an approach goal partially accounted for the state Neuroticism-state negative affect link.

Although Wilt et al. (2016) conducted two separate studies, the second was essentially a replication of the first, with similar methods and results. Therefore, I will describe these

studies' details together rather than separately. Both studies included 40 participants in analyses; about half of Study 1's participants were college students and half were community adults. All of Study 2's participants were college students. In each of these studies, subjects received text messages six times a day at three-hour intervals (9 AM, 12 PM, 3 PM, 6 PM, 9 PM, 12 AM) for two weeks; thus, each participant received a total of 84 text messages. The mean number of responses provided by each participant in Study 1 was 55.7 (SD = 23.7), whereas the mean response rate in Study 2 was 69.7 (SD = 24.4). Both studies had sufficient power to detect moderate within-person effect sizes in multilevel models, but insufficient power to detect between-persons effects (Scherbaum & Ferreter, 2009).

The text messages in each study prompted participants for data regarding state personality, state affect, goal velocity, goal importance, and goal proximity. In addition, Study 2 collected data on the approach/avoidance nature of the goal. In all cases, prompts referred to affect in the present moment and behavior within the past 30 minutes. Goal-related items instructed participants to reflect on and respond relative to the the single goal that they had most been trying to achieve in the preceding half-hour. Goal Importance and Goal Proximity were assessed using 6-point scales; i.e. from 1 (not at all important) to 6 (very important) and from 1 (far away/gave up) to 6 (very close/achieved), respectively.

Goal Velocity, which Wilt et al. (2016) showed was related to affect even when controlling for importance and proximity, was rated slightly differently in Studies 1 and 2. In Study 1, participants ranked three items on a 1 (disagree) to 6 (agree) scale: 'I was moving quickly toward the goal,' 'I was moving slowly toward the goal,' 'I was doing better than expected.' In Study 2, participants indicated velocity by responding to a single item,

ranking their perceived progress toward the goal on a scale from 1 (more slowly than expected) to 6 (more quickly than expected). Furthermore, in Study 2, the approach/avoidance nature of the goal was rated on a scale of 1 to 2 in response to the prompt 'Would the goal be best described as pursuing a positive outcome (1) or avoiding a negative outcome (2)?'

Personality states (within the past 30 minutes) and state affect (in the present moment) were scored on 6-point scales; in all cases, (1) was low and (6) was high. State Extraversion was described using the adjectives *bold, quiet* (reversed), *talkative* (in Study 1), *assertive, withdrawn* (reversed), and *unrestrained* (in Study 2). The words used to assess Neuroticism states were *touchy, temperamental, insecure* (in Study 1), and *steady* (reversed), *anxious*, and *emotional* (in Study 2).

State Positive and Negative Affect (i.e. PA and NA in each of the two studies) were measured using similar three-to-four-item scales. In Study 1, PA was assessed using the words *alert, happy, attentive*, and *strong*; in Study 2, it was assessed using the words *happy, cheerful*, and *pleased*. Study 1's NA items were *anxious, irritable, intense*, and *upset*; Study 2's were *grouchy, irritable*, and *gloomy*.

5.2 Assumptions

5.2.1 Participants and Experiences

As in the simulations discussed in the previous chapter, participant age and gender were ignored. Each simulated participant 'reported on' the same number of experiences: in order to obtain this number, I averaged the mean response rates for the original Studies 1 and 2, then rounded that figure to the nearest integer. Thus, each participant was

subjected to 63 experiences, all of which were created using random-normally distributed instigating and consummatory cue values.

5.2.2 State and Goal Measures

State extraversion and neuroticism were operationally defined as the average intensities of BAS and FFFS actions, respectively, over the measurement interval.

$$E_S = \overline{BAS}_{Ap} \tag{5.1}$$

$$N_S = \overline{FFFS}_{Ap} \tag{5.2}$$

State positive and negative affect were operationally defined as the average falling velocities of BAS and FFFS tendencies, respectively, over the measurement interval.

$$PA_S = \overline{BAS}_{Tvf} \tag{5.3}$$

$$NA_S = \overline{FFFS}_{Tvf} \tag{5.4}$$

As in the Chapter 4 simulation, each participant at each experience was confronted by two simulated goals. However, because all cues were randomly-generated (making it difficult to label one goal the 'target' and the other a 'distractor,' as in Chapter 4's experimentally-controlled Experience 2), state and goal data were calculated by summing

the appropriate averages for both goals.

Most of the participants in the Wilt et al. (2016) Study 2 reported pursuing more approach than avoidance goals. To reiterate, Approach was coded as '1' and Avoidance as '2.' Because the mean of all participants' Approach/Avoidance means was 1.16 (SD = .32), it can be inferred that 84% of reports described approach goals rather than avoidance goals.

For this reason, and because both goal-directed approach (BAS) and avoidance (FFFS/BIS) actions occur relatively frequently for any given CTA-RST goal, I elected to drop the original Study 2's measure of Approach/Avoidance and focus solely on target-goal-directed approach actions' velocities. Somewhat surprisingly, approach goal velocity was ultimately operationally defined as the average duration of BAS actions over the measurement interval.

$$V_{BAS} = \overline{BAS}_{At} \tag{5.5}$$

CTA-RST is not at present sophisticated enough to simulate goal importance or proximity; therefore, these variables also were not included in the simulated data.

5.3 Method: Parameter Settings

Each of eighty simulated subjects (40 for those in the original Study 1 and 40 for the original Study 2) confronted two simulated goals in each of 63 simulated experiences. A different set of participants was generated for this simulation by altering the random 'seed' setting (seed = 2500). Each experience in turn was run for 30 simulated minutes

(900 iterations). Actions and tendencies for each goal in each experience started from the constant initial values described in Chapter 3. Measurements were made over the last 300 of the 900 iterations. These settings were chosen because all affect prompts asked participants how they felt at the present moment (as opposed to within the last half-hour) about the primary goal they'd been pursuing in the past 30 minutes, and because 300 was similar to the number of iterations, 240, employed in Chapter 4. Instigating and consummatory cues (for each goal and experience) and weights (for each subject over all experiences) were randomly drawn and normally distributed. All analysis code may be found in Appendix F.

5.4 Method: Analyses

As in the original Studies 1 and 2, descriptive statistics, bivariate multilevel models (MLMs), and path-analytic mediation MLMs were calculated from the simulated data. In particular, descriptive statistics included the average individual's means, within-person standard deviations, and within-person alphas of states.

Six bivariate MLMs were conducted using R's `nlme` package (Pinheiro, Bates, DebRoy, Sarkar, & R Development Core Team, 2016): (1) Extraversion predicts Velocity, (2) Neuroticism predicts Velocity, (3) Velocity predicts PA, (4) Velocity predicts NA, (5) Extraversion predicts PA, and (6) Neuroticism predicts NA. Data for each experience were nested within subjects. The procedure for each MLM entailed, first, that predictor variables be centered around each subject's mean and that dummy variables for persons and observations be created. When null and alternative models were built for each bivariate MLM, alternative models included within-subjects effects as random effects.

The bivariate MLMs paved the way in the original studies for the 1-1-1 path-analytic mediation MLMs. Whereas the former models each relationship of interest separately, the latter models all relationships simultaneously. The `MplusAutomation` package in R and a demonstration copy of *MPlus* (Muthén & Muthén, 2005) were used to construct and evaluate these models.

5.5 Results

5.5.1 Descriptive Statistics

Means, within-person standard deviations, and average within-person alphas for the original and simulated studies are shown in Table 5.1. Because all simulation constructs were computed from scaled scores or sums of scaled scores, all simulation means were equal to 0. Simulated within-person standard deviations, which ranged from .85 to 1.12, were on average slightly higher than but otherwise comparable to those in the original studies, which ranged from .66 to 1.01 in Study 1 and from .66 to 1.08 in Study 2. In all cases, state Neuroticism had the lowest within-person SD. Simulated within-person alpha statistics were much higher than the original studies', which is to be expected given that simulated constructs represent true scores rather than relatively error-ridden, self-reported scores obtained from measures consisting of a few adjectives per scale. Nevertheless, as Wilt et al. (2016) noted, the original alphas' relatively low values likely had more to do with the scales' brevity than with the strength of their items' relationships.

Table 5.1: Descriptive Statistics for Wilt et al. (2016) S1 and S2; Simulation

SD = Within-person standard deviation; α = Within-person alpha

Variable	M_{S1}	M_{S2}	M_{Sim}	SD_{S1}	SD_{S2}	SD_{Sim}	α_{S1}	α_{S2}	α_{Sim}
E_S	2.65	3.36	0	1.01	0.74	0.89	0.60	0.44	0.98
N_S	1.79	2.62	0	0.66	0.66	0.85	0.54	0.39	0.98
V	3.60	3.90	0	0.98	1.08	0.96	0.78	–	0.96
PA_S	3.27	3.86	0	0.75	0.90	1.12	0.62	0.83	0.89
NA_S	1.94	1.84	0	0.66	0.76	1.03	0.59	0.83	0.91

5.5.2 Inferential Statistics

Betas, p-values, and confidence intervals from the original and simulated studies' bivariate MLMs are given in Table 5.2; inferential statistics for the original and simulated studies' 1-1-1 mediation MLMs are given in Table 5.3 and Table 5.4. In particular, the bivariate models' success is a prerequisite to the success of the 1-1-1 path analyses. In the original studies by Wilt et al. (2016), it was expected that these analyses would show that (a) state extraversion was related to state positive affect by way of increased approach goal velocity and (b) state neuroticism was related to state negative affect by way of decreased approach goal velocity. Each of these path analyses contained three key bivariate linkages: (1) state personality to state affect, (2) state personality to velocity, and (3) velocity to state affect. Significant bivariate multilevel models of these relationships would testify to the plausibility of conducting more stringent multilevel path analyses of moderation.

The last three columns of Table 5.2 show that, as was the case in the original research, all hypothesized bivariate links were supported. That is, higher approach goal velocity was positively predicted by state extraversion and negatively by state neuroticism; in turn, higher velocity was positively related to state positive and negatively related to state negative affect. Also as expected, state extraversion positively predicted state positive

affect, while state neuroticism positively predicted state negative affect. In general, these relationships, which ranged in absolute value from $|\beta| = .51$ to $|\beta| = .88$, were stronger in the simulated results than in the original ones, which ranged from $|\beta| = .13$ to $|\beta| = .53$. Overall, the simulated bivariate MLMs' beta values were highly correlated with those in the original Study 1 ($r = .85$) and Study 2 ($r = .96$), as shown in Figure 5.1.

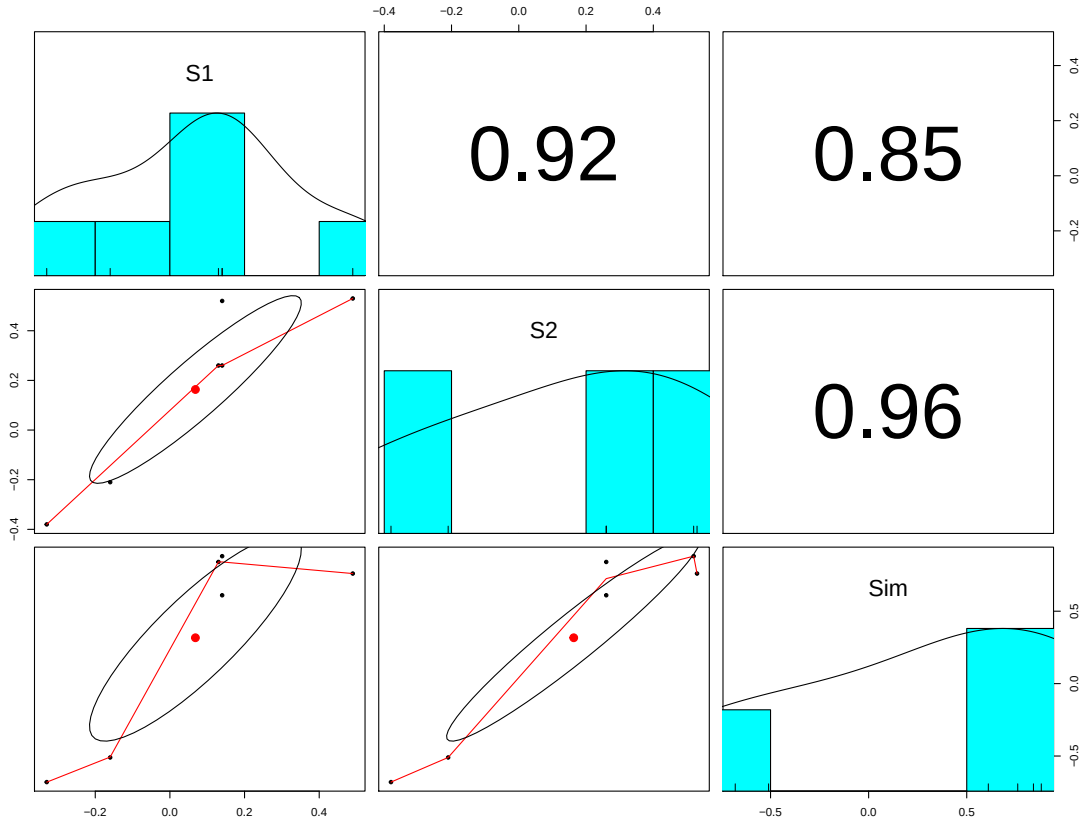


Figure 5.1: Correlations between bivariate MLMs' beta weights for simulated data (Sim) and data from the original Study 1 (S1) and Study 2 (S2)

Note that the multilevel path analyses in Wilt et al.'s (2016) Study 2 originally included a link predicting state extraversion (or neuroticism) from a categorical approach/avoidance goal variable; results for these links have been left out of Tables 5.3

and 5.4 because I have assumed for the sake of simplicity that all simulated goals are approach goals. Out of concern for consistency as well as interpretability, I have negatively scored velocity in Table 5.4, just as Wilt et al. (2016) did.

Once again, results for both path models were in agreement with those in the original research; direct effect results reflected those of the bivariate multilevel models and supported all hypothesized linkages among state personality, state affect, and approach goal velocity. Importantly, the indirect effects' significance support the hypothesis that goal velocity mediates the relationships between state personality and state affect. In other words, one of the ways in which state extraversion (neuroticism) gives rise to state positive (negative) affect is by means of increased (decreased) progress toward approach goals.

Table 5.2: Results of Wilt et al. (2016) and Simulated Bivariate MLMs

Predictor	Outcome	b_{S1}	95% CI_{S2}	p_{S1}	b_{S2}	95% CI_{S2}	p_{S2}	b_{Sim}	99% CI_{Sim}	p_{Sim}
E_S	V	0.13	[.06, .21]	$< .001$	.26	[.18, .35]	$< .001$	0.84	[.73, .94]	$< .001$
N_S	V	-0.33	[-.44, -.23]	$< .001$	-.38	[-.43, -.32]	$< .001$	-0.68	[-.74, -.62]	$< .001$
V	PA_S	0.14	[.08, .20]	$< .001$	.26	[.19, .34]	$< .001$	0.61	[.54, .68]	$< .001$
V	NA_S	-0.16	[-.22, -.10]	$< .001$	-.21	[-.27, -.15]	$< .001$	-0.51	[-.60, -.42]	$< .001$
E_S	PA_S	0.14	[.08, .20]	$< .001$	.52	[.43, .62]	$< .001$	0.88	[.75, 1.01]	$< .001$
N_S	NA_S	0.49	[.40, .59]	$< .001$	.53	[.43, .63]	$< .001$	0.76	[.61, .88]	$< .001$

Table 5.3: Results of Original and Model 1-1-1 MLMs Relating State Extraversion to State PA via Velocity

Path	b_{S1}	95% CI_{S2}	p_{S1}	b_{S2}	95% CI_{S2}	p_{S2}	b_{Sim}	99% CI_{Sim}	p_{Sim}
Direct effects									
V ON E_S	.19	[.13, .24]	$<.001$	.28	[.20, .35]	$<.001$	0.51	[.40, .63]	$< .001$
PA_S ON V	.10	[.03, .17]	.004	.24	[.18, .30]	$<.001$	0.52	[.41, .62]	$< .001$
PA_S ON E_S	.14	[.07, .20]	$<.001$	.44	[.33, .54]	$<.001$	0.24	[.11, .37]	$< .001$
Indirect effect									
E_S to V to PA_S	.02	[.01, .03]	.010	.07	[.04, .09]	$<.001$	0.27	[.19, .34]	$< .001$
Total effect									
E_S to PA_S	.16	[.09, .22]	$<.001$	.50	[.40, .61]	$<.001$	0.51	[.34, .68]	$< .001$

Table 5.4: Results of Original and Model 1-1-1 MLMs Relating State Neuroticism to State NA via Velocity

Path	b_{S1}	95% CI_{S2}	p_{S1}	b_{S2}	95% CI_{S2}	p_{S2}	b_{Sim}	99% CI_{Sim}	p_{Sim}
Direct effects									
V (-) ON N_S	.27	[.17, .38]	$< .001$	.38	[.29, .47]	$< .001$	.52	[.44, .59]	$< .001$
NA_S ON V (-)	.09	[.05, .12]	$< .001$	.13	[.08, .18]	$< .001$	.35	[.27, .42]	$< .001$
NA_S ON N_S	.43	[.35, .50]	$< .001$	.54	[.43, .65]	$< .001$	.25	[.12, .39]	$< .001$
Indirect effect									
N_S to V (-) to NA	.12	[.06, .17]	$< .001$	.20	[.13, .27]	$< .001$	.18	[.13, .23]	$< .001$
Total effect									
N_S to NA_S	.54	[.43, .65]	$< .001$	.74	[.57, .91]	$< .001$	.43	[.28, .58]	$< .001$

5.6 Discussion

In Chapter 4, the CTA-RST simulation's results suggested that different kinds of positive affect are predicated on different characteristics of the functions describing tendencies and actions as they evolve over time. In turn, trait personality need not have the same effect on or relationship to all positive affects. In using the same operational definitions for state personality (E_S, N_S) as Chapter 4 used for pleasant and negative state affect (PA_S, NA_S), respectively, Chapter 5 shows not only that personality states can be described in terms of actions' characteristics, but also that state personality and state affect are not necessarily such very different things. Of course, this is by no means a revolutionary idea, as numerous other researchers have suggested that extraversion (trait or state) is, at its core, strongly related if not identical to positive affect, while the essence of neuroticism (trait or state) is negative affect. Trivially, Chapter 5's results also underscore the common wisdom that the same or similar psychological constructs often go by a variety of different names (e.g., the same construct as was called activated affect or EA_S in Chapter 4 is called positive affect or PA_S in Chapter 5, etc.).

What is perhaps more interesting as regards this simulation's operational definitions is

that I was unable, as it were, to easily model 'velocity *as* velocity.' That is, when I defined goal velocity as a function of the average rising velocity of BAS actions over the measurement interval, then (among other difficulties) goal velocity was much more likely to be positively than negatively related to state negative affect. Of course, given the large number of researcher degrees of freedom in any modeling endeavor, there is a very real possibility that I simply have not stumbled upon the correct function of BAS action velocity or the correct combination of state affect and personality variables that would, when used in tandem with such a function, yield the expected results. It is safer at this juncture to say that it is *easier* to model 'approach goal velocity' as 'approach action duration,' given the operational definitions I have chosen to adopt for other state variables, than it is to say that it is *more correct* to do so. Nevertheless, it seems highly likely that, in much the same way as it is easier to obtain a 'good-enough' solution to a set of differential equations by means of any one of a large number of quite different techniques than it is to obtain an exact 'best' solution, it is probably also easier to obtain a 'good-enough' than a 'best' model of psychological phenomena by means of any one of a large number of different possible combinations of operational definitions. That being the case, it is if nothing else interesting to note that approach goal action duration is a 'good-enough' model of perceived approach goal velocity.

The results of this second simulated study were particularly promising inasmuch as the precise pattern and direction of the original studies' results were replicated in full, using only simple, intuitive operational definitions of affect that mirrored those used in other simulations. Nevertheless, the modeled study's effects are somewhat larger than in the original. To some extent, this can be attributed to the fact that simulations are inherently less noisy than real studies; thus, we would expect all effects to be stronger in the former than in the latter. There are, in fact, several ways in which the model's data

are necessarily cleaner than real data would be; for instance, not only are the model's psychological constructs more akin to true scores than self-reported approximations, but the model's participants' tendencies and actions also (unrealistically) start at the same value for any given goal and experience. The problem to be more closely examined in future research is, of course, exactly *how much* stronger we ought to expect the model's effects to be, given the model's constraints.

Chapter 6

Study 3 – RST and Affective Synchrony

> Poetry might be defined as the clear expression of mixed feelings.
>
> W. H. Auden, *New Year Letter*

6.1 Background

Rafaeli, Rogers, and Revelle (2007) noted that individuals tend to vary from mild affective synchrony ($r > 0$) to mild affective desynchrony ($r < 0$), with a majority falling in the asynchronous range. These authors also made an initial foray into explanation when they investigated several personality and social-cognitive variables that might predict differences in affective synchrony. Somewhat unexpectedly, affective traits like Extraversion and Neuroticism were uncorrelated with affective synchrony.

Wilt et al. (2011) replicated the Rafaeli et al. (2007) findings that individuals differ in affective synchrony and that the latter is temporally stable within persons. Using

experience-sampling methods, Wilt et al. (2011) extended the original research by examining the relationships between not only state energetic (EA_S) and tense arousal (TA_S), but also state pleasant (PA_S) and unpleasant affect (UA_S) and threat and challenge appraisals. The authors found that individual differences in EA_S-TA_S relationships were predicted by interactions between corresponding affective traits (i.e. EA_T, TA_T, UA_T, and PA_T); Rafaeli et al. (2007) failed to find effects of extraversion and neuroticism in predicting affective synchrony because they only looked for main effects. Specifically, EA_S-TA_S correlations were higher for individuals low in TA_T (or UA_T) as EA_T (or PA_T) decreased, while EA_S-TA_S correlations were higher for individuals high in TA_T (or UA_T) as EA_T (or PA_T) increased. In essence, when affective traits' (EA_T-TA_T or PA_T-UA_T) strengths were more evenly matched, the tendency to experience synchronous activated and tense emotions was stronger.

As was the case in Chapter 5, the original research consisted of two separate studies, where the second replicated and extended the first. Studies 1 and 2 included 42 and 40 participants, respectively. The experience-sampling procedure was the same as in Chapter 5; participants text-messaged self-reports of state affect in response to prompts received six times a day at three-hour intervals (9 AM, 12 PM, 3 PM, 6 PM, 9 PM, 12 AM) for two weeks. The mean number of responses provided by each participant in Study 1 was 69.7 (SD = 24.4); the mean number of responses for Study 2 was 55.7 (SD = 23.7).

Studies 1 and 2 measured state affect and situational appraisals; in addition, Study 2 assessed trait affect. Notably, trait affect measures consisted of the same words as state affect measures did, but with an altered prompt: whereas state affects were responses to the question 'How are you feeling right now?', trait affects were completions of the sentence 'In general, I feel...' All affective adjectives were drawn from the Motivational

State Questionnaire (MSQ, Revelle & Anderson, 1996), which is scored on a 6-point scale (1: not at all; 6: very well) and may be decomposed into three factors (EA, TA, and (un)pleasantness or valence Rafaeli & Revelle, 2006).

In particular, (state and trait) EA, TA, UA, and PA measures each consisted of two or three adjectives. EA_S was represented in Study 1 by *energetic, alert, sluggish* (R); in Study 2, EA_S and EA_T were framed as *excited, lively, full-of-pep*, and *vigorous*. The words *calm* (R), *relaxed* (R), and *tense* stood in for TA_S in Study 1; *distressed, jittery, nervous*, and *stirred-up* measured TA_S and TA_T in Study 2. Study 1 PA_S was *confident, cheerful, pleased*, while Study 2 PA_S and PA_T were *happy* and *strong*. Finally, Study 1's UA_S comprised *grouchy, irritable*, and *gloomy*, but Study 2's UA_S and UA_T were *irritable* and *upset*.

Situational appraisal items prompted participants to describe (again, on a 1 to 6 scale representing descriptiveness from 'not at all' to 'very well,' respectively) the situations they'd experienced in the past half-hour. Study 1 operationalized appraisals in terms only of threat (i.e. with the adjectives *threatening, risky*, and *negative*) and reward (*challenging, rewarding, positive*), whereas Study 2 used three single-item measures to examine threat (*threatening*), challenge (*challenging*), and pleasing (*pleasing*) appraisals.

6.2 Assumptions

6.2.1 Participants and Experiences

Again, all subject demographic characteristics (i.e. gender, age, etc.) were ignored. The number of simulated experiences for each participant was calculated in the same manner as in Chapter 5, and (because mean response rates for the two studies were the same as

in the original Chapter 5 study) yielded the same result (i.e. 63 experiences). In fact, as the original study's lead author confirms (Wilt, personal communication), the subjects in the original Studies 1 and 2 for this chapter were the same as for the original Studies 2 and 1, respectively, in the preceding chapter. Nevertheless, I simulated different groups of participants for this chapter and the previous one.

Although no situational appraisal measures were included in the simulation (see below), each experience was allowed to run for 900 iterations (30 simulated minutes) because prompting participants to think of the situations they'd experienced in the past half-hour likely influenced the nature of the affect they reported. All participant data were analyzed together, rather than in two separate replications.

6.2.2 State, Trait, and Appraisal Measures

I chose not to include measures of threatening, rewarding, challenging, or pleasing situational appraisals in my simulation because CTA-RST doesn't yet model the cognitive aspects of emotional experience in any obvious way. As in the second simulated study, participants were presented with two simulated goals, and state measures were summed over both goals. Here, EA_S and TA_S were operationalized using the same definitions as were used for PA_S and NA_S in Chapter 5; that is, EA_S and TA_S were represented by the average falling velocities of BAS and FFFS tendencies, respectively, over a 300-iteration measurement interval (as in Chapter 5).

$$EA_S = \overline{BAS}_{Tvf} \tag{6.1}$$

$$TA_S = \overline{FFFS}_{Tvf} \tag{6.2}$$

State pleasant and unpleasant affect, on the other hand, were defined by modified versions of the equations used to compute state extraversion and state neuroticism in Chapter 5. Here, PA_S and UA_S are defined in terms of BAS and FFFS action intensities, respectively, and averaged over 300 iterations before summing across goals.

$$PA_S = \overline{BAS}_{Ap} \tag{6.3}$$

$$UA_S = \overline{FFFS}_{Ap} \tag{6.4}$$

This simulation's measures of trait affect are given by the following instigating or consummatory weights, or linear combinations thereof:

$$EA_T = BAS_w \tag{6.5}$$

$$TA_T = BIS_w + FFFS_w - BIS_c \tag{6.6}$$

$$PA_T = -BAS_c \tag{6.7}$$

$$UA_T = -FFFS_c \tag{6.8}$$

6.3 Method: Parameter Settings

Eighty-two subjects (42 for Study 1 and 40 for Study 2) were run (seed = 1500), confronting two simulated goals in each of 63 simulated experiences for 30 simulated minutes. Instigating and consummatory sensitivities for each subject over all experiences were random-normally distributed, as were instigating and consummatory cues for each goal and experience.

6.4 Method: Analyses

Descriptive statistics, bivariate multilevel models (MLMs), and moderated MLMs examining the main effects and interactions of trait affect in predicting affective synchrony were calculated using all simulated data. All analyses' code may be found in Appendix G.

6.4.1 Descriptive Statistics

These included the average individual's mean (i.e., each variable's aggregated item mean across all subjects and experiences); the standard deviation across all reports, which represents the sum of between- and within-person variability; the between-person variability (i.e., variation over participants' means); and the within-person variability (i.e., variation within each participant's responses). Means and standard deviations of affective traits were found, as well.

6.4.2 Bivariate MLMs of State Affect

Two bivariate MLMs (TA_S predicts EA_S and UA_S predicts PA_S) were conducted using the `nlme` package (Pinheiro et al., 2016). Predictor variables (TA_S, UA_S) were mean-centered, dummy variables for persons and observations were created, and data were stacked before creating the data frame for all bivariate MLMs.

Null (fixed effects) and alternative (random effects) models were built for both bivariate MLMs, and these were compared (specifically, the likelihood ratios were examined to determine whether the assumption of random effects improved on the original models). If the alternative model significantly improved on the null model, it was safe to assume that there were individual differences in these affects' synchrony. The average within-persons relationship between the affects in the model was given by the beta value, and the 95% confidence interval for that beta provides a measure of the relationship's variability. The Pearson correlation between the average EA_S-TA_S and UA_S-PA_S relationships was also found.

6.4.3 Moderated MLMs: State and Trait Affect Predict State Affect

Here, trait as well as state affects served as predictor variables; thus, these were mean-centered and stacked before being combined with person, observation, and other variables in data frames for each of the four trait-interaction MLMs.

The four trait-interaction models included (1) TA_S, EA_T, and TA_T predict EA_S; (2) TA_S, PA_T, and UA_T predict EA_S; (3) UA_S, PA_T, and UA_T predict PA_S; and (4) UA_S, EA_T, and TA_T predict PA_S. As usual, if the alternative model improved significantly on the null, betas and their significances were reported.

6.5 Results

6.5.1 Descriptive Statistics

Means and standard deviations for state affect in the simulation and in the original Studies 1 and 2 are shown in Table 6.1, as are the means and standard deviations of trait affect in the simulation and in the original Study 2. Alpha reliabilities for the original measures of trait affect were .88 (EA), .63 (TA), .73 (PA), and .40 (UA). Note that both the mean of all reports and the mean across subjects were calculated for the original Studies 1 and 2, but only the mean of all reports was calculated for the simulation. This was because the two means would have been identical for the simulation, in which all 'subjects' contributed complete data. The standard deviation across all reports (SD) is a measure of the total variability within the data, which is itself a result of both between-person (traitlike) and within-person (statelike) variability, here represented by SD_{BP} and SD_{WP}, respectively.

As was the case in Chapter 5, simulated dependent variables were constructed from scaled scores, resulting in means of zero. Total, between-person, and within-person variability were higher for all variables in the model than in the original research. Although original values for SD_{BP} were relatively lower than original values for SD_{WP}, this trend appeared to be reversed in the simulation's results.

6.5.2 Inferential Statistics

Results for original and simulated bivariate multilevel models (MLMs) may be found in Table 6.2, while results for the original and simulated trait-interaction MLMs (significant predictors only) are shown in Table 6.3. As in Chapter 5, bivariate MLMs are preliminary

Table 6.1: Descriptive Statistics for Wilt et al. 2011, Studies 1 and 2 (S2), and Simulation (Sim)

X_{BP} = between subjects variable, X_{WP} = within subjects variable

Variable	All	Reports					Across	Subjects						
	M	(S2)	(Sim)	SD	(S2)	(Sim)	M	(S2)	SD_{BP}	(S2)	(Sim)	SD_{WP}	(S2)	(Sim)
EA_S	3.99	2.43	0	1.08	1.10	1.26	3.96	2.48	0.51	0.77	1.26	0.98	0.80	1.11
TA_S	2.55	2.23	0	1.04	0.89	1.21	2.62	2.26	0.66	0.58	1.21	0.82	0.69	1.03
PA_S	3.91	3.32	0	1.09	1.14	1.31	3.86	3.42	0.68	0.80	1.32	0.87	0.80	0.93
UA_S	1.85	1.85	0	0.98	0.99	1.32	1.88	1.87	0.59	0.50	1.32	0.81	0.87	1.00
EA		3.95	0.02		1.07	1.07								
TA		2.49	-0.08		0.70	1.64								
PA		4.47	0.19		0.91	0.99								
UA		2.48	0.21		0.84	0.91								

analyses that reveal the feasibility of moving on to more stringent tests of moderation. In Table 6.2, L-ratios (L_{S1}, L_{S2}, L_{Sim}) are the log likelihood values for each model; when these are significant, the ('random effects') model in which slopes relating the predictor to the outcome are allowed to vary across individuals is superior to the ('fixed effects') model in which the slope is held constant (fixed) across all individuals. Table 6.2 shows data for random effects models only. Here, the average relationships between predictor and outcome are given by unstandardized beta coefficients (b_{S1}, b_{S2}, b_{Sim}), while the variability in these relationships is represented by the range of beta coefficients observed ($Range_{S1}$, etc.). All beta values reported in Table 6.2 are significant at the $p < .001$ level.

Because their likelihood ratios were significant, all simulated bivariate MLMs supported the hypothesis that relationships between EA_S and TA_S (or between PA_S and UA_S) varied across participants, replicating the findings in the original Studies 1 and 2. Furthermore, the simulation's results were similar in magnitude and direction to those in the original studies. The bivariate MLM predicting EA_S from TA_S showed that the EA_S-TA_S relationship ranged from strong desynchrony (β = -.95) to weak synchrony (β = .15), as it did in original research by both Wilt et al. (2011) and Rafaeli et al. (2007). Similarly, simulated results mirrored the original's in showing the PA_S-UA_S relationship to range from strong desynchrony (β = -1.75) to a value near zero (β = .04), if not to

weak desynchrony (indeed, the original Study 2 exhibited similar results).

Nevertheless, the range of simulated EA_S-TA_S relationships was slightly compressed relative to the original and, though the model mimicked real life inasmuch as the average β linking simulated EA_S and TA_S was less negative (-.21) than the average β between simulated PA_S and UA_S (-.26), this difference was not as pronounced as the ones reported by Wilt et al. (2011). Analyses in the original research also found the correlation between average EA-TA and UA-PA relationships; in Study 1, this was r = -.27, and in Study 2 it was r = -.05, whereas simulated results yielded a correlation of r = .06. This relationship may also owe its weakly positive value to the relatively more negative range of simulated EA_S-TA_S relationships.

The correlation between simulated and original beta values was reliable and positive (r = .53), as shown in Figure 6.1. Nevertheless, results of the simulation's moderated multilevel models also disagreed in some ways with those featured in Wilt et al. (2011). One of the two critical positive three-way trait-by-trait-by-state interactions were non-significant in the simulation, and one was negative rather than positive (indicating that more nearly matching levels of EA_T and TA_T, or UA_T and PA_T, predicted greater simulated *desynchrony* between EA_S and TA_S).

Table 6.2: Results of Wilt et al. (2011) and Simulated Bivariate MLMs

$p < .001$ for all L-ratios

Predictor	Outcome	b_{S1}	$Range_{S1}$	L_{S1}	b_{S2}	$Range_{S2}$	L_{S2}	b_{Sim}	$Range_{Sim}$	L_{Sim}
TA_S	EA_S	-.26	(-.85, .26)	87.84	-.10	(-.99, .45)	141.67	-.21	(-.95, .15)	61.37
UA_S	PA_S	-.64	(-1.17,-.17)	60.90	-.27	(-1.09, .22)	187.87	-.26	(-1.75, .04)	472.72

Table 6.3: Results of Wilt et al. (2011) and Simulated Moderated MLMs

Model	1: $TA_S * EA_T * TA_T \rightarrow EA_S$					
	R2			Sim		
Predictor	β	95% CI	p	β	95% CI	p
TA_S	-0.01	(-.17, .15)	.88	-0.22	(-.27, -.16)	$< .001$
EA_T	0.41	(.19, .62)	$< .001$	0.28	(.22, .34)	$< .001$
TA_T	0.42	(.10, .74)	$< .05$	.04	(.004, .08)	$< .05$
$TA_S \times EA_T$	-0.16	(-.32, -.004)	$< .05$	-0.04	(-.09, .01)	.14
$TA_S \times TA_T$	0.15	(-.07, .38)	.19	0.03	(-.01, .06)	.13
$EA_T \times TA_T$	-0.34	(-.67, -.01)	$< .05$	0.02	(-.03, .07)	.41
$TA_S \times EA_T \times TA_T$	0.39	(.16, .63)	$< .001$	-0.02	(-.06, .01)	.22
	2: $TA_S * PA_T * UA_T \rightarrow EA_S$					
TA_S	-0.02	(-.20, .15)	.78	-0.20	(-.25, -.16)	$< .001$
PA_T	0.08	(-.20, .37)	.56	-0.03	(-.12, .05)	.45
UA_T	0.05	(-.26, .37)	.73	0.02	(-.07, .12)	.60
$TA_S \times PA_T$	-0.09	(-.28, .10)	.37	-0.05	(-.09, -.002)	$< .05$
$TA_S \times UA_T$	0.12	(-.09, .34)	.25	-0.16	(-.21, -.11)	$< .001$
$PA_T \times UA_T$	-0.31	(-.61,-.005)	$< .05$	-0.22	(-.34, -.11)	$< .001$
$TA_S \times PA_T \times UA_T$	0.31	(.10, .51)	$< .01$	0.04	(-.02, .10)	.24
	3: $UA_S * PA_T * UA_T \rightarrow PA_S$					
UA_S	-0.28	(-.38, -.17)	$< .001$	-0.24	(-.29, -.19)	$< .001$
PA_T	0.33	(.09, .62)	$< .05$	0.68	(.54, .82)	$< .001$
UA_T	0.28	(-.04, .60)	.09	-0.08	(-.23, .07)	.32
$UA_S \times PA_T$	-0.05	(-.17, .07)	.40	-0.27	(-.32, -.22)	$< .001$
$UA_S \times UA_T$	0.07	(-.06, .20)	.30	0.0004	(-.05, .05)	.99
$PA_T \times UA_T$	-0.07	(-.38,.23)	.63	-0.32	(-.51, -.14)	$< .01$
$UA_S \times PA_T \times UA_T$	-0.01	(-.13, .12)	.93	0.13	(.06, .19)	$< .001$
	4: $UA_S * EA_T * TA_T \rightarrow PA_S$					
UA_S	-0.26	(-.36, -.15)	.38	-0.26	(-.33, -.19)	$< .001$
EA_T	0.37	(.13, .60)	$< .01$	0.39	(.22, .57)	$< .001$
TA_T	0.44	(.09, .78)	$< .05$	-0.06	(-.17, .06)	.32
$UA_S \times EA_T$	-0.05	(-.15, .05)	.33	-0.10	(-.16, -.03)	$< .001$
$UA_S \times TA_T$	0.07	(-.08, .22)	.37	0.03	(-.02, .07)	$< .05$
$EA_T \times TA_T$	-0.28	(-.64, .08)	.12	0.03	(-.10, .17)	.64
$UA_S \times EA_T \times TA_T$	0.07	(-.08, .22)	.37	-0.02	(-.07, .03)	.40

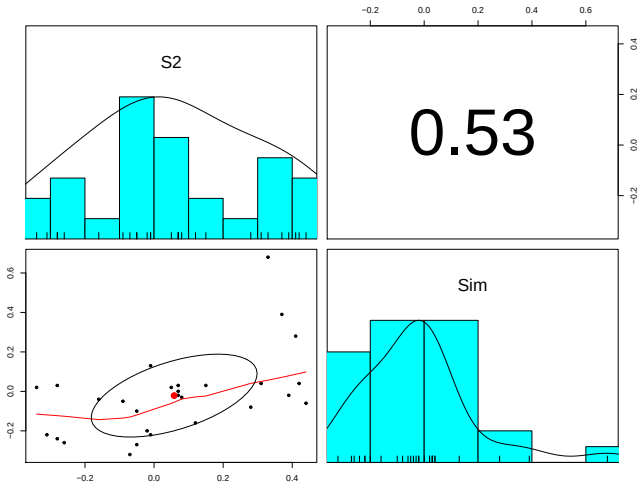


Figure 6.1: Correlation between beta values for original and simulated moderated MLMs

6.6 Discussion

The current simulation's operational definitions were, by and large, consistent with those used throughout this paper. Nevertheless, and although I was able to replicate many important findings regarding affective synchrony, Chapter 6's results are less similar to the originals than other simulations' have been.

6.6.1 Bivariate Multilevel Models

The dissimilarity between the range of EA_S-TA_S associations observed in the simulated and real studies' bivariate multilevel models raises interesting empirical questions. Consider that I have made essentially arbitrary decisions here and elsewhere regarding the number of iterations over which to evaluate the functions determining dependent variables' values. While I was test-running this simulation, I noticed that the range of observed affective associations (EA_S-TA_S or PA_S-UA_S) was quite sensitive to the

number of iterations over which state affects were calculated. I noticed that, in general, both (action-based) measures of PA_S and UA_S and (tendency-based) measures of EA_S and TA_S were more negatively correlated the shorter the time interval over which they were evaluated (that is, the ranges of observed β coefficients contained fewer positive values).

Moreover, it seems intuitively obvious that individual differences are in play here, too; some people will no doubt reflect on feelings that have occurred over a longer interval preceding the prompt than others will, to say nothing of the differences in the *types* of feelings that each emphasizes most within his or her preferred span. If individual differences in evaluation period exist, then issuing all participants a constant number of iterations as a proxy for that period would be tantamount to assuming that said constant was the (weighted) average evaluation period for that participant sample. When I ran simulations in which evaluation time periods varied between persons, the results were similar to those reported above. Ultimately, I chose not to use randomly-drawn evaluation intervals out of concern for both simplicity and consistency. I elected to make the state affects' evaluation interval consistent with the constant 300-iteration interval used in Chapter 5 (where the same constructs were instead labeled PA_S and NA_S, respectively).

6.6.2 Moderated Multilevel Models

Some of the difference between the real and simulated moderated multilevel models can, as usual, be chalked up to the fact that simulated data are cleaner than real data. Other dissimilarities, however, require more careful consideration. Overall, there are two types of discrepancies worthy of discussion here: first, results that are non-significant in the original but significant in the simulated data; and second, results that are non-significant

in the simulation but significant in the original data. Regarding the first of these, the 'clean simulation data' explanation is not the only one; the 'number of iterations' explanation from the previous section could apply as well.

The second of the two difficulties is both more serious and more difficult to explain. On the bright side, only five effects out of 28 across all four models exhibit this problem, so it is possible that they are nothing more than random errors. In three of the cases (i.e. the three-way and trait-by-trait two-way interactions in Model 1 and the main effect of TA_T in Model 4), the beta values are in different directions; however, the simulation's statistics are very close to zero, which is less problematic than it would be if they were stronger. Nevertheless, the problem remains. My intuition is that its probable causes are similar to the ones that I believe caused the difficulties with the range of EA_S-TA_S relationships in the bivariate MLMs and with the strong main effects of state affect in the moderated MLMs. In other words, this simulation appears to be much more sensitively dependent than others were on the number of iterations over which I chose to reduce data in order to create states and traits.

6.6.3 Summary

Despite the difficulties I had with this simulation, it at the very least still contains the seeds of several promising findings. Not only was I able to replicate the basic results that individuals differ in affective synchrony and that different kinds of affects are differently synchronous, I was also able to show, in many cases, the simulated states and traits produced results concordant with those found in real data.

Chapter 7

Conclusion

Knowledge is power.

Sir Francis Bacon, 1597 (Meditationes Sacrae and Human Philosophy)

7.1 General Discussion

Despite its professed aim to describe and explain individual differences in the patterning, across time and situations, of affect, behavior, cognition, and desire, personality psychology has continued to struggle with one of those two critical goals. Because it is ethically impossible to randomly assign human beings' personalities, mechanistic explanation of the individual differences that scholars have over the past century, with abundant success, observed and factor-analyzed has, at best, failed to thrive. In accordance with its role as descriptive personality psychology's vestigial twin, explanatory personality psychology has largely been either overlooked or spurned within the academic community. All hope, however, is not lost; recent developments in computational modeling have afforded investigators an alternative means of experimentally manipulating

individual differences and, as such, have bestowed new legitimacy on the problem of how to explain not only psychological differences between persons, but also their time- and cue-dependent changes within persons.

In this thesis, I have introduced CTA-RST, a new computational model that simulates within- and between-persons individual differences in affect and behavior. Although CTA-RST is not the only model of its kind, I have gone further with it than any of my predecessors have with their simulations.

Other models of RST and related theories (e.g. Pickering, 1997; Read et al., 2010; Smillie, Pickering, & Jackson, 2006) have demonstrated their capacity for producing output that is in agreement with theoretical prediction (i.e. that is face-valid), but have not endeavored to simulate real research in any great detail. Indeed, the only computational model of which I am aware that *does* attempt to simulate real research studies is the Quek and Moskowitz (2007) 'Workplace Roles' neural network, and the latter is quite different from CTA-RST, both because it is a standard non-deterministic model that 'trains' connections between nodes using a portion of the to-be-modeled data, and because the individual differences it models are less general (e.g. 'agreeable' or 'quarrelsome' behavior) and more social-cognitive (e.g. 'male' or 'female,' 'supervisor' or 'supervisee') than those that CTA-RST models.

Here, I have shown that CTA-RST is capable of modeling state and trait affect and personality data from real studies. In Chapters 4 through 6, I correlated the beta values or effect sizes of comparable analyses whenever there were at least six numbers that allowed me to compare simulated and real data. As Figures 4.4, 5.1, and 6.1 show, these correlations ranged from $r = .53$ to $r = .96$, with a median correlation of $r = .73$. Pooling all the results across studies in three columns (where 'Res1' refers to the original results

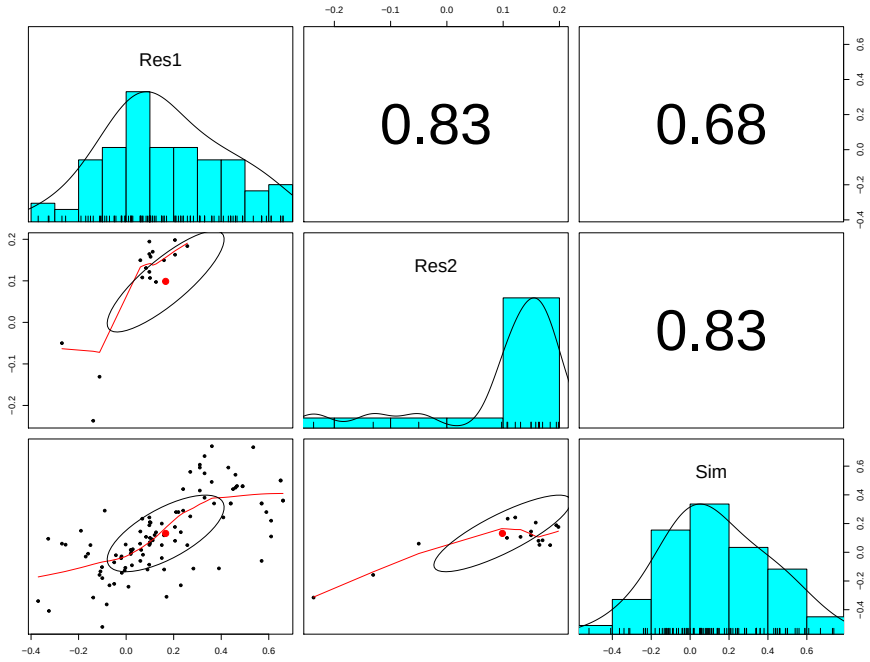


Figure 7.1: Correlations between original and simulated statistics across studies

in Chapters 4, 6, and 'Study 1' in Chapter 5 and 'Res2' refers to the original results in Chapter 5's 'Study 2') and converting all of them to effect sizes (Pearson's r) yields a correlation of $r = .68$ between the simulation data all 'Res1' original data (see Figure 7.1).

Taken together, the results of the simulations discussed in the previous three chapters attest to CTA-RST's validity as a computational model of personality and affect. The CTA-RST functions used in each simulation's operational definitions of states and traits are summarized in Table 7.1. Note here that there is consistency across functions insofar as BAS- and FFFS-related functions always correspond to positively and negatively valenced constructs, respectively. As much as was possible, I tried to use the same functions to define affective states with the same names (e.g. energetic positive affect, EA_S, in Chapter 6 and in Chapter 4's Experiment 4) and, barring that (as was the case

when affects had relatively uninformative names or variable content; e.g. 'Negative Affect'), I tried to use the same functions to define constructs whose self-report measures had the most similar item content (e.g. using EA_S in Chapter 4's Experiment 1 and Chapter 5 for the constructs originally labeled 'Positive Affect'). Although I tried a number of different possible measurement intervals on each function in each chapter, I also did my best to keep these consistent across studies.

Table 7.1: Functions Used in Operational Definitions of Traits and States, by Chapter

Chapter	4: Smillie et al., 2012 (240 iterations)			5: Wilt et al., 2016 (900 iterations)			6: Wilt et al., 2011 (900 iterations)		
Function	State	t	G	State	t	G	State	t	G
$\overline{BAS}_{Ap}$	PA_S	240	1	E_S	300	1, 2	PA_S	300	1, 2
$\overline{FFFS}_{Ap}$	NA_S	240	1	N_S	300	1, 2	UA_S	300	1, 2
$\overline{BAS}_{Tvf}$	EA_S	240	1	PA_S	300	1, 2	EA_S	300	1, 2
$\overline{FFFS}_{Tvf}$	–	–	–	NA_S	300	1, 2	TA_S	300	1, 2
$\overline{BAS}_{At}$	–	–	–	V_{BAS}	300	1, 2	–	–	–

Trait	Chapter	Function
Extraversion, E_T	4	$BAS_c + BIS_c + FFFS_c$
Pleasant Affect, PA_T	6	$-BAS_c$
Unpleasant Affect, UA_T	6	$-FFFS_c$
Energetic Affect, EA_T	6	BAS_w
Tense Affect, TA_T	6	$BIS_w + FFFS_w$ - BIS_c

Although BAS- and FFFS-related functions appear where they might be expected to, a few hypotheses with which I began my exploration of CTA-RST's parameter space did not bear fruit. For instance, tendencies' rising velocities proved to be less appropriate measures of activated and tense (i.e. motivated positive and negative) affect than I had expected they would. The key problem with rising tendencies' velocities was that using them in the Chapter 5 study of goal velocity and state affect created a peculiar situation in which the expected relationships in the 1-1-1 multilevel models could only be recovered if I made the counterintuitive assumption that state positive affect corresponded to the

rising velocity of FFFS tendencies, whereas state negative affect corresponded to the rising velocity of BAS tendencies! Noting that FFFS tendencies usually rose while BAS actions were proceeding (and thus while BAS tendencies were falling) and vice versa, I decided to try letting tendencies' falling velocities, rather than their rising velocities, stand in for motivated affect. In other words, because falling tendencies imply that the person in question is *doing* the associated action, it seemed reasonable to assume that falling tendencies would provide a good measure of activated affect. This assumption proved to be a great success, in the Chapter 5 simulation and elsewhere.

Why might it be the case that falling tendencies are better determinants of activated or tense affect than rising tendencies are? After all, in Chapter 3 I noted that it seemed quite natural to assume that, because tendencies rise while they are not being acted upon, they must then correspond to a (necessarily unfulfilled) desire to act in a certain way. Assuming for the moment that falling tendencies' good performance in this thesis as measures of motivated affect is not merely accidental, I would suggest that because rising tendencies are not being acted upon, we may assume that consciousness of them (and therefore affect related to them) is inhibited. It is only when the rising tendency becomes great enough to dominate other tendencies that the action it controls will begin; therefore, the tendency only creates self-reportable affect after it begins creating an action (i.e. after it begins falling).

In other words, the potential objection is that rising tendencies reflect things that a person is not currently doing, and people are only motivated or 'want' to do things that they're not currently doing. Therefore, if 'wanting' and energetic affect, for example, are the same, shouldn't rising BAS tendencies reflect 'wanting' better than falling ones do? Again, the argument in favor of falling tendencies as producers of motivated affect pivots

on consciousness. People only know *upon reflection* at any given moment that they aren't doing everything they *want* to do. However, unless what they *are doing* at a given moment *is reflecting* on all they want to do, they can only 'want to' do what they *are* doing. In other words, motivated affect should be related to what people are doing, not to what they aren't doing.

Having made the conceptual transition from thinking about motivated affect in terms of rising tendencies' slopes to thinking of it in terms of falling tendencies' slopes, I nevertheless did not give up on the idea that what differentiated motivated affects like EA_S and TA_S from ones like pleasant and unpleasant affect (e.g. Chapter 6's PA_S and UA_S) was the extent to which the functions that defined them were capable of synchrony. That is, as noted in the original study of affective synchrony on which Chapter 6 is based, motivated affects of opposite valence are more likely to co-occur than pleasant and unpleasant affect are; indeed, given their consistent negative correlations with one another, the latter appear almost to be mutually exclusive. Because I programmed CTA-RST's action-generating functions in such a way that FFFS-related (avoidance) and BAS-related (approach) actions relative to a given goal could not occur simultaneously, it seemed reasonable to suppose that measures of affect based on actions' characteristics would tend to be more desynchronous than measures based on tendencies' characteristics. This supposition was ultimately supported by the results of the Chapter 6 simulation and, indirectly, by the results of the Chapter 4 simulation.

It will be useful to note here, too, that in defining UA_S as a property of FFFS actions rather than (the absence of) BAS actions, I am departing from the hypotheses advanced by Wilt et al. (2011) in their discussion regarding pleasantness-unpleasantness as a unitary, essentially BAS-related dimension. However, I think I am justified in doing so

because (a) defining UA_S as a negative function of PA_S would have created greater desynchrony between pleasantness and unpleasantness than the original results uncovered (and, I believe, greater than the latter should have led us to expect to observe even in relatively clean simulated data) and (b) CTA-RST's BAS-FFFS action exclusivity within goals *ought* to create something like a unitary dimension (i.e., for a given goal, one is either advancing towards it or not; thus, one is either happy with one's progress or not, and, if the latter, regardless of the reasons, BAS- or FFFS-related, why one is not making progress), even if the latter isn't quite the same as the one that Wilt et al. (2011) proposed. Indeed, we can still recover the two unitary dimensions on which Wilt et al. (2011) speculated (in addition to pleasantness-unpleasantness, the authors deduced that there should be a unitary avoidance-related affective dimension, with fear at the negative pole and relief at the positive pole) if instead of couching the latter in terms of BAS and FFFS we couch them, respectively, in terms of whether progress toward or away from a given goal or set of goals comprises 'good' progress.

Finally, the results of the three simulations discussed in this thesis demonstrate that some assumptions about trait personality and affect cannot be taken for granted, and that there may be more than one kind of biologically-based reinforcement sensitivity underlying observable, self-reportable traits. In constructing the operational definitions of traits employed in Chapters 4 and 6, I assumed that sensitivity weights (or linear combinations thereof) would be better able to model real data than means of states. Whenever it seemed necessary to construct functions of weights to approximate real trait data, I preferred to use sums (not products) of three or fewer sensitivity variables, and I assigned all summed terms equal (unitary) weight. These restrictions led to simply-defined traits, many of which were equal to only one sensitivity variable or its negative.

This simplicity, however, does not imply that those who seek to construct self-report measures of RST are right in developing scales that assume a one-to-one correspondence between RST system sensitivities and observable traits. As Smillie et al. (2006) point out, it is reinforcement sensitivity systems' *functional outcomes* (i.e. their effects on the tendencies and actions that they interact to produce), rather than properties of the systems themselves, that we measure when we ask participants to fill out self-report measures of behavior and affect. Although it may sometimes be possible, for certain simple traits, to use only one system's sensitivity to model trait data, it is equally possible that a trait is best defined as a linear combination of system sensitivities, such that many different (unobservable) combinations of weights lead to the same self-reported level of the (observable) trait. The fact that systems probably possess both instigating and consummatory sensitivities further complicates the matter, and, as stated in Chapter 4, creates all kinds of new possibilities regarding traits' correspondence to the structures and processes that give rise to reinforcement sensitivities.

7.2 Limitations

Obviously, as it is a computational model, CTA-RST is and always will be vulnerable to the objection that its data are *not real* and are therefore irrelevant to any meaningful conversation about the constructs it purports to model. However, to any who would make this objection, I would respond that models by definition do not produce real data, and therein lies their value. It is only in constructing models and observing their agreement or disagreement with real data that we are able to improve our understanding of anything. Our understanding of anything is itself a model, as are any of our attempts to analyze whatever 'real' data we might have collected.

Models are made to be falsified, and CTA-RST is no different. There are improvements to be made to the model's BIS cue input function, action- and tendency-generating functions, and operational definitions of personality and affect; indeed, so much room for improvement that I am sure I won't be able to enumerate all the improvements that there are to be made. That said, I can certainly remark on a few of the most important and on those which I'm most likely to make in the immediate future.

First, and despite the fact that Ashby's (1956) argument for cybernetic isomorphism is a good one, it very nearly goes without saying that it will ultimately be necessary to render CTA-RST's tendency and action functions in greater neurophysiological detail (after all, we only know so much about what kinds of input-output relationships to expect, independent of anatomy and physiology; incorporating what we know about the latter in our models will help us make predictions about as-yet unknown or unremarked phenomena). As it is, CTA-RST still lacks many of the bells and whistles that lend similar models their biological verisimilitude. The Read et al. (2010) neural network model, for instance, incorporates not only positivity offset but also negativity bias; CTA-RST only includes the former. It might also be beneficial to undertake a more complicated project; for example, to modify CTA-RST in such a way as to enable it to simulate frontal-related negativity, as the Pickering and Pesola (2014) model does.

One way in which CTA-RST's parameter space might be made more realistic (but also more complicated) would be to create greater continuity across simulated subjects' experiences by allowing previous experiences' final action and tendency values to carry over into the next experience as initial conditions. In point of fact, I attempted to do this in several instantiations of the simulation described in Chapter 4 (pleasant versus appetitive mood inductions). Recall that in discussing modeled data's results' four

departures from those reported for the original Experiment 4, I mentioned that the simulated Step 1 moderated multiple regression predicting post-test from pre-test EA_S was nonsignificant, whereas the original had been significant and positive.

Although this failure (labeled in Chapter 4, 'Problem (3),' with a promise to discuss it further here) was rather serious in terms of what it seemed to do to the simulation's results, it was also relatively easy to explain. In several of the many pilot versions of this simulation that I ran, I made the aforementioned supposedly simple improvement to continuity, and allowed previous experiences' final action and tendency values to 'wrap around' into the initial values of the next experience. When I used this version of the simulation in pilot-testing, the correlations between baseline and post-test affects were higher, and the correlation between pre-test and post-test activated affect in Experiment 4, in particular, was significant and positive. However, once I made the Chapter 4 timescale consistent with those used in Chapters 5 and 6 (reducing the total time for each experience from an initial 2400 iterations to 240 iterations), the wrap-around simulations had much lower power to detect effects than wrapper-less simulations, and I ultimately chose not to use the wrap-around code in the final simulations. This decision often helped and rarely hurt results' similarity to the original studies'; the one exception being, of course, that I lost the correlation between pre-test and post-test activated affect in Chapter 4's Experiment 4.

Naturally, as it is still in the early stages of development, CTA-RST lacks not only biological and (as in the case of the wrap-around code) some ecological validity, but also the mere ability to model certain constructs. As noted in previous chapters, CTA-RST does not model age, gender, or cognitive processes. In order to allow it to do so, I would need either to equip CTA-RST with more of the features that one traditionally associates

with neural networks (e.g. categorical input and output nodes linked via probabilistic 'hidden layers' that, over a number of trials, would be capable of learning input-output relationships similar to those observed in real data), or to complicate some of its current parameter functions accordingly. Indeed, one might argue that CTA-RST already includes crude cognitive processing capabilities, inasmuch as it reduces data about tendencies and affects over a measurement interval in order to generate simulated self-reported states or traits. It is quite possible that *any* data reduction process that takes place in the brain could be construed as a cognitive process; individual differences in such processing include, but are not limited to, the number of timesteps over which data are reduced, the weights given to different types of data (BAS vs. FFFS, tendency vs. action, velocity vs. position, etc.; note that these may or may not be the same things as systems' instigating or consummatory sensitivities), and the fidelity with which the reduction is made (individual differences in the latter may correspond to what we might term individual differences in 'cognitive ability').

7.3 Possibilities for Future Research

One the best opportunities as regards future research using CTA-RST has already been discussed at length; that is, I am sure that improving upon the model's current limitations will itself take up a good deal of my time (and that of other researchers who choose to adopt it). As the CTA-RST simulation is improved, I expect it to do an ever-better job of demystifying the relationships between affect and behavior; however, its applications don't end there. Ultimately, I hope it will prove helpful both in generating novel hypotheses and in settling scholars' longstanding RST-related disagreements. Naturally, such exercises will also inevitably contribute to CTA-RST's refinement. Below,

I will sketch a few of the empirical questions that the model could be used to investigate.

One possibility is that, once CTA-RST reaches a certain level of biological sophistication, it could be used to probe the details of the neurophysiological mechanisms that give rise to the functional outcomes that I've thus far been referring to as 'goal activations' or 'tendencies.' Are certain sets of neurons' integrated firing rates the means by which tendencies are neurally instantiated (Rolls, McCabe, & Redoute, 2008)? If so, this simple identity could be used to corroborate and add precision to extant knowledge regarding which cortical and subcortical structures handle which goals, and how, thus providing a window on such RST-relevant phenomena as response prediction error (Cooper, Duke, Pickering, & Smillie, 2015; Pickering & Pesola, 2014; Smillie, Cooper, & Pickering, 2011) and arousal (Anderson & Revelle, 1994; Matthews & Gilliland, 1999).

A second broad category of fruitful future research may lie in work coupling elaborated computational models, like those developed by Fua et al. (2010), and comparative psychological data. These could prove beneficial in understanding how goals' outputs are expressed and experienced at different levels of cognitive complexity (reactive, routine, and reflective; Ortony et al., 2005). For example, CTA-RST simulations of play and hunting behavior within and across species might provide a clearer picture of the dynamics involved in learning complex behaviors. Similarly, simulations of competition for mates or territory might illuminate the dynamics of aggression (RST's *Fight* construct), and simulations of kin or group altruism might better explain the difference between impulsive and defensive approach.

CTA-RST is by no means limited to the role of 'faithful draftsman,' good for nothing but reproducing archival data. Again, once it has reached a certain level of development, CTA-RST will become an invaluable tool for generating novel hypotheses, which may

then be tested by collecting new (real) data. Self-report data, for one, could be used to test CTA-RST-generated hypotheses regarding (a) what individual differences determine which behavior, approach or avoidance, will occur following BIS activation and (b) what RST predicts with regard to decisions made among multiple non-conflicting, yet complementary goals (not to mention the affects and cognitions that accompany or follow such decisions). Briefly, I tend to believe that (a) whether BIS activation is followed by (defensive) approach or (passive) avoidance is a function of how BAS-reactive the individual is and (b) non-conflicting goals can be 'grafted' together to form goal combinations or hybrids (as when an individual avoids one goal – e.g. doing homework – by approaching another – e.g. playing video games). Indeed, Read, Smith, Droutman, and Miller (2016) have already begun to make strides in this direction, and I hope to collaborate with them in the coming years.

Finally, future projects could use CTA-RST or similar computational models to simulate the manner in which RST gives rise to or interacts with other kinds of individual differences. For instance, impulsivity is an RST-related trait with an identity crisis (Matthews, 2008); and, as suggested in Chapter 4, simulations could finally put to bed the question of where impulsivity belongs, structurally speaking, in the pantheon of human traits (personally, I suspect that acknowledgment of the role of consummatory sensitivities, in addition to that of instigating sensitivities, in trait creation will go a long way toward disentangling the sociability and impulsivity components of extraversion). This idea is in turn linked with that of discovering which combinations of CTA-RST functions create the best representations of each of the well-known top-down trait measures, both within (EPQ Extraversion, IPIP Extraversion, etc.) and across (Extraversion, Neuroticism, etc.) recognized traits.

Of course, there are probably a number of individual differences that are unrelated to, yet capable of interacting with, individual differences in reinforcement sensitivities; intelligence is one such (see *Limitations*, above, for a note regarding possible candidates for cognitive capacities in CTA-RST as it stands). Given the reciprocal interplay between motivation, affect, and cognition, understanding the relationships between intelligence- and RST-related individual differences could be quite important in predicting all kinds of pertinent criteria. In fact, all the potential avenues for future research outlined above are stepping stones along the path to improved understanding of the processes that give rise to individual differences and, as such, are keys to unlocking the principles that will, in time, allow us to gain control of them via effective interventions.

Judicious control over our environments and ourselves is, as always, the ultimate goal; knowledge is power, is it not? With that in mind, it is worth repeating one last time that good computational models are one of the most powerful weapons in the theoretical psychologist's arsenal. Building and testing such models allows us to construct precise operational definitions of individual differences and untangle the time-dependent, often-complicated mechanisms by which they interact. Only once we have gained such understanding will we be able not only to make strong claims about the missing conceptual links between reinforcement sensitivities and the traits to which they give rise, but also to make changes for the better, as human nature compels us to, in what has hitherto appeared to be in the hands of fate.

References

Allport, G. W. (1937). *Personality: A psychological interpretation.* New York: Holt.

Anderson, K. J., & Revelle, W. (1994). Impulsivity and time of day: Is rate of change in arousal a function of impulsivity? *Journal of Personality and Social Psychology*, *67*(2), 334–344.

Aristotle. (2014). In R. Crisp (Ed.), *Nicomachean ethics.* Cambridge, UK: Cambridge University Press. (Original work published c. 350 BCE)

Ashby, W. R., et al. (1956). *An introduction to cybernetics.* London: Chapman and Hail Ltd.

Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality*, *15*(5), 327–353.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*(2), 150–166.

Atkinson, J. W., & Birch, D. (1970). *The dynamics of action.* New York, NY: John Wiley.

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.

Blanchard, R. J., & Blanchard, D. C. (1990a). Anti-predator defense as models of animal fear and anxiety. In P. F. Brain, S. Parmigiani, R. J. Blanchard, & D. Mainardi (Eds.), *Fear and defense* (p. 89-108). Harwood Academic Publishers.

Blanchard, R. J., & Blanchard, D. C. (1990b). An ethoexperimental analysis of defense, fear, and anxiety. In N. McNaughton & G. Andrews (Eds.), *Anxiety* (pp. 124–133). Dunedin: Otago University Press.

Broadhurst, P. (1975). The Maudsley reactive and nonreactive strains of rats: A survey. *Behavioral Genetics*, *5*(4), 299–319.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space.

Personality and Social Psychology Review, *1*(1), 3–25.

Carver, C. S. (2006). Approach, avoidance, and the self-regulation of affect and action. *Motivation and emotion*, *30*(2), 105–110.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, *135*(2), 183–204.

Cervone, D. (2004). The architecture of personality. *Psychological Review*, *111*(1), 183–204.

Clark, L. A., & Watson, D. (1999). Temperament: A new paradigm for trait psychology. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research, 2nd edition* (pp. 399–423). New York, NY: Guilford Press.

Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A proposal. *Archives of General Psychiatry*, *44*(6), 573–588.

Cole, M., Hood, L., & McDermott, R. (1994). *Ecological niche picking: Ecological invalidity as an axiom of experimental cognitive psychology.* Rockefeller University: Laboratory of Comparative Human Cognition.

Collins, M. D., Jackson, C. J., Walker, B. R., O'connor, P. J., & Gardiner, E. (2017). Integrating the context-appropriate balanced attention model and reinforcement sensitivity theory: Towards a domain-general personality process model. *Psychological Bulletin*, *143*(1), 91–106.

Condon, D. M. (2014). *An organizational framework for the psychological individual differences: Integrating the affective, cognitive, and conative domains* (Unpublished doctoral dissertation). Northwestern University.

Confucius. (2005).
In A. Waley (Ed.), *The analects of Confucius.* New York, NY: Routledge. (Original work published c. 435 BCE)

Cooper, A. J., Duke, É., Pickering, A. D., & Smillie, L. D. (2015). Individual differences in reward prediction error: Contrasting relations between feedback-related negativity and trait measures of reward sensitivity, impulsivity and extraversion. *Dopaminergic Foundations of Personality and Individual Differences*, *8*, 40–50.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience & Biobehavioral Reviews*, *28*(3), 317–332.

Corr, P. J., DeYoung, C. G., & McNaughton, N. (2013). Motivation and personality: A neuropsychological perspective. *Social and Personality Psychology Compass*, *7*(3), 158–175.

Corr, P. J., & McNaughton, N. (2008). Reinforcement sensitivity theory and personality. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 155–187). Cambridge: Cambridge University Press.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual.* Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J., & Gleser, G. C. (1965). Psychological tests and personnel decisions, 2nd edition.

DeYoung, C. G. (2015). Cybernetic big five theory. *Journal of Research in Personality*, *56*, 33–58.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880.

Digman, J. M. (1997). Higher-order factors of the big five. *Journal of personality and social psychology*, *73*(6), 1246.

Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, *82*(5), 804–818.

Eysenck, H. J. (1981). *A model for personality*. Cambridge University Press.

Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, *73*(6), 1224–1237.

Eysenck, H. J. (1999). Biological dimensions of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (p. 244-276). Elsevier.

Eysenck, H. J., & Eysenck, S. B. (1967). On the unitary nature of extraversion. *Acta Psychologica*, *26*, 383–390.

Eysenck, H. J., & Eysenck, S. B. (1991). *The Eysenck personality questionnaire – revised*. Sevenoaks, England: Hodder and Stoughton.

Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, *6*(1), 21–29.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2011). Asymmetries in predictive and diagnostic reasoning. *Journal of Experimental Psychology: General*, *140*(2), 168–185.

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689–723.

Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, *13*(2), 83–87.

Fleeson, W. (2007a). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, *75*(4), 825–862.

Fleeson, W. (2007b). Using experience sampling and multilevel modeling to study person-situation interactionist approaches to positive psychology. In A. D. Ong & M. H. M. Van Dulmen (Eds.), *The Oxford handbook of methods in positive psychology* (p. 501-514). Oxford University Press.

Freud, S. (1924). *Beyond the pleasure principle*. New York: Boni and Liveright.

Fua, K., Revelle, W., & Ortony, A. (2010). Modeling personality and individual differences: The approach-avoid-conflict triad. In *Annual meeting of the cognitive science society (cogsci 2010)* (Vol. 31, pp. 25–30). Portland, OR.

Galton, F. (1884). The measurement of character. *Fortnightly Review*, *42*, 179-185.

Goldberg, L. R. (1990). An alternative 'description of personality': the Big-Five factor

structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in europe* (Vol. 7, pp. 7–28). Tilburg, Holland: Tilburg University Press.

Gosling, S. D., & John, O. P. (1999). Personality dimensions in nonhuman animals: A cross-species review. *Current Directions in Psychological Science*, *8*(3), 69–75.

Gray, J. A. (1970). The psychophysiological basis of introversion-extraversion. *Behaviour research and therapy*, *8*(3), 249–266.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system.* New York: Oxford University Press.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system, 2nd edition.* New York: Oxford University Press.

Gray, J. A., & Smith, P. T. (1969). An arousal-decision model for partial reinforcement and discrimination learning. In R. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 243–272). London: Academic Press.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*(3), 121–134.

Hall, C. S. (1934). Emotional behavior in the rat. i. defecation and urination as measures of individual differences in emotionality. *Journal of Comparative Psychology*, *18*(3), 385–403.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not weird. *Nature*, *466*(7302), 29–29.

Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, *16*(2), 107–117.

Hippocrates, & Galen. (1846). In J. R. Coxe (Ed.), *The writings of Hippocrates and Galen.* Philadelphia, PA: Lindsay and Blakiston. (Original work published c. 460 BCE)

Hull, C. L. (1932). The goal-gradient hypothesis and maze learning. *Psychological Review*, *39*(1), 25–43.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, *91*(2), 153 - 184.

Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review*, *114*(2), 423–440.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–292.

Kuhl, J. (2000). A functional-design approach to motivation and self-regulation: The dynamics of personality systems and interactions. In M. Boekaerts, P. R. Pintrich,

& M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 111–169). San Diego, CA: Academic Press.

Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, *61*(1), 132–140.

Leszko, M., Elleman, L. G., Bastarache, E. D., Graham, E. K., & Mroczek, D. K. (2016). Future directions in the study of personality in adulthood and older age. *Gerontology*, *62*(2), 210–215.

Matthews, G. (2008). Reinforcement sensitivity theory: A critique from cognitive science. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 482–507). Cambridge: Cambridge University Press.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits*. Cambridge University Press.

Matthews, G., & Gilliland, K. (1999). The personality theories of H. J. Eysenck and J. A. Gray: A comparative review. *Personality and Individual Differences*, *26*(4), 583–626.

Matthews, G., Jones, D. M., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST mood adjective checklist. *British Journal of Psychology*, *81*(1), 17–42.

Mayer, J. D., Allen, I. P., & Beauregard, K. (1995). Mood inductions for four specific moods: A procedure employing guided imagery. *Journal of Mental imagery*, *19*(1-2), 133–150.

McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, *63*(3), 365-396.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*(3), 204-217.

McNaughton, N., & Corr, P. J. (2008a). Animal cognition and human personality. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 95–119). Cambridge: Cambridge University Press.

McNaughton, N., & Corr, P. J. (2008b). The neuropsychology of fear and anxiety: A foundation for Reinforcement Sensitivity Theory. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 44–94). Cambridge: Cambridge University Press.

Miller, N. E. (1959). Liberalization of basic SR concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science. Study 1* (pp. 198–292). New York, NY: McGraw-Hill.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246–268.

Mitchell, J. T., Kimbrel, N. A., Hundt, N. E., Cobb, A. R., Nelson-Gray, R. O., & Lootens, C. M. (2007). An analysis of reinforcement sensitivity theory and the

Five-Factor Model. *European Journal of Personality*, *21*(7), 869-887.
Morsella, E. (2005). The function of phenomenal states: Supramodular interaction theory. *Psychological Review*, *112*(4), 1000-1021.
Moskowitz, D. S., Suh, E. J., & Desaulniers, J. (1994). Situational influences on gender differences in agency and communion. *Journal of Personality and Social Psychology*, *66*(4), 753–761.
Mroczek, D. K., & Spiro, A. (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *58*(3), 153–165.
Muthén, L. K., & Muthén, B. O. (2005). *Mplus: Statistical analysis with latent variables: User's guide*. Los Angeles, CA: Muthén & Muthén.
Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622.
Nettle, D., & Penke, L. (2010). Personality: Bridging the literatures from human psychology and behavioural ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1560), 4043–4050.
Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
Ortony, A., Norman, D. A., & Revelle, W. (2005). Affect and proto-affect in effective functioning. In J.-M. Fellous & M. A. Arbib (Eds.), *Who needs emotions?: The brain meets the robot* (pp. 173–202). Oxford University Press.
Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: Oxford University Press.
Pearl, J. (2009). *Causality: Models, reasoning, and inference, 2nd edition*. New York: Cambridge University Press.
Pickering, A. D. (1997). The conceptual nervous system and personality: From Pavlov to neural networks. *European Psychologist*, *2*(2), 139–163.
Pickering, A. D. (2004). *A formal model of 'new' reinforcement sensitivity theory.* (12th European Conference on Personality, Groningen, The Netherlands)
Pickering, A. D. (2008). Formal and computational models of RST. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 453–481). Cambridge: Cambridge University Press.
Pickering, A. D., Corr, P. J., Powell, J. H., Kumari, V., Thornton, J. C., & Gray, J. A. (1997). Individual differences in reactions to reinforcing stimuli are neither black nor white: to what extent are they Gray? In H. Nyborg (Ed.), *The scientific study of human nature: Tribute to hans j. eysenck at eighty.* Pergamon/Elsevier Science Inc.
Pickering, A. D., & Pesola, F. (2014). Modeling dopaminergic and other processes involved in learning from reward prediction error: Contributions from an individual differences perspective. *Frontiers in Human Neuroscience*, *8*.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Development Core Team. (2016). nlme: Linear and nonlinear mixed effects models [Computer software manual].

Vienna, Austria.
Poropat, A. E., & Corr, P. J. (2015). Thinking bigger: The Cronbachian paradigm & personality theory integration. *Journal of Research in Personality*, *56*, 59–69.
Poznanski, M., & Thagard, P. (2005). Changing personalities: Towards realistic virtual characters. *Journal of Experimental and Theoretical Artificial Intelligence*, *17*(3), 221–241.
Prut, L., & Belzung, C. (2003). The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: A review. *European Journal of Pharmacology*, *463*(1), 3–33.
Quek, M., & Moskowitz, D. (2007). Testing neural network models of personality. *Journal of Research in Personality*, *41*(3), 700–706.
R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. (ISBN 3-900051-07-0)
Rafaeli, E., & Revelle, W. (2006). A premature consensus: Are happiness and sadness truly opposite affects? *Motivation and Emotion*, *30*(1), 1–12.
Rafaeli, E., Rogers, G. M., & Revelle, W. (2007). Affective synchrony: Individual differences in mixed emotions. *Personality and Social Psychology Bulletin*, *33*, 915–932.
Read, S. J., & Miller, L. C. (1989). Inter-personalism: Toward a goal-based theory of persons in relationships. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 413–472). Hillsdale, NJ: Erlbaum.
Read, S. J., & Miller, L. C. (2002). Virtual personalities: A neural network model of personality. *Personality and Social Psychology Review*, *6*(4), 357–369.
Read, S. J., Miller, L. C., Monroe, B., Brownstein, A., Zachary, W., LeMentec, J.-C., & Iordanov, V. (2006). A neurobiologically inspired model of personality in an intelligent agent. In J. Gratch, M. Young, & R. Aylett (Eds.), *Intelligent virtual agents: 6th international conference, IVA 2006* (pp. 316–328). Berlin, Germany: Springer-Verlag.
Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological Review*, *117*(1), 61–92.
Read, S. J., Smith, B. J., Droutman, V., & Miller, L. C. (2016). Virtual personalities: Using computational modeling to understand within-person variability. *Journal of Research in Personality*.
Revelle, W. (1983). Factors are fictions, and other comments on individuality theory. *Journal of Personality*, *51*(4), 707–714.
Revelle, W. (1986). Motivation and efficiency of cognitive performance. In *Frontiers of motivational psychology* (pp. 107–131). Springer.
Revelle, W. (1995). Personality processes. *Annual Review of Psychology*, *46*(1), 295–328.
Revelle, W. (2008). The contribution of reinforcement sensitivity theory to personality theory. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp.

508–527). Cambridge University Press.

Revelle, W. (2012). Integrating personality, cognition and emotion: Putting the dots together? In M. W. Eysenck, M. Fajkowska, & T. Maruszewski (Eds.), *Personality, cognition and emotion. Warsaw lectures in personality and social psychology* (p. 157-177). New York: Eliot Werner Publications.

Revelle, W. (2016). Psych: Procedures for psychological, psychometric, and personality research. *Northwestern University, Evanston, IL, http://personality-project. org/r/psych. manual. pdf*.

Revelle, W., & Anderson, K. (1996). Time of day and activation states: Interactions with impulsivity and neuroticism. In *European association of personality psychology.* Belgium.

Revelle, W., & Condon, D. (2015). A model for personality at three levels. *Journal of Research in Personality*, *56*, 70–81.

Revelle, W., & Michaels, E. J. (1976). The theory of achievement motivation revisited: The implications of inertial tendencies. *Psychological Review*, *83*(5), 394–404.

Revelle, W., & Wilt, J. A. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, *47*(5), 493–504.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1-25.

Rocklin, T., & Revelle, W. (1981). The measurement of extraversion: A comparison of the Eysenck Personality Inventory and the Eysenck Personality Questionnaire. *British Journal of Social Psychology*, *20*(4), 279–284.

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex*, *18*(3), 652–663.

Rushton, J. P., Bons, T. A., & Hur, Y.-M. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality*, *42*(5), 1173–1185.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367.

Shearer, T. A., & Pruitt, J. N. (2014). Individual differences in boldness positively correlate with heart rate in orb-weaving spiders of genus larinioides. *Current Zoology*, *60*(3), 387–391.

Shoda, Y., & Mischel, W. (1998). Personality as a stable cognitive-affective activation network: Characteristic patterns of behavior variation emerge from a stable personality structure. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior* (pp. 175–208). Mahwah, NJ: Erlbaum.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359-1366.

Smillie, L. D. (2008). What is reinforcement sensitivity? Neuroscience paradigms for approach-avoidance process theories of personality. *European Journal of Personality*, *22*(5), 359–384.

Smillie, L. D., Cooper, A. J., & Pickering, A. D. (2011). Individual differences in reward–prediction–error: Extraversion and feedback-related negativity. *Social Cognitive and Affective Neuroscience*, *6*(5), 646–652.

Smillie, L. D., Cooper, A. J., Wilt, J. A., & Revelle, W. (2012). Do extraverts get more bang for the buck? Refining the affective-reactivity hypothesis of extraversion. *Journal of Personality and Social Psychology*, *103*(2), 306-327.

Smillie, L. D., Geaney, J. T., Wilt, J. A., Cooper, A. J., & Revelle, W. (2013). Aspects of extraversion are unrelated to pleasant affective-reactivity: Further examination of the affective-reactivity hypothesis. *Journal of Research in Personality*, *47*(5), 580–587.

Smillie, L. D., Loxton, N. J., & Avery, R. E. (2011). Reinforcement sensitivity theory, research, applications, and future. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (p. 101-131). John Wiley & Sons.

Smillie, L. D., Pickering, A. D., & Jackson, C. J. (2006). The new reinforcement sensitivity theory: Implications for personality measurement. *Personality and Social Psychology Review*, *10*(4), 320–335.

Suh, E. J., Moskowitz, D., Fournier, M. A., & Zuroff, D. C. (2004). Gender and relationships: Influences on agentic and communal behaviors. *Personal Relationships*, *11*(1), 41–60.

Theophrastus. (1927). In J. M. Edmonds (Ed.), *The characters of Theophrastus.* New York, NY: G. P. Putnam's Sons. (Original work published c. 319 BCE)

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and Individual Differences*, *31*(6), 837–862.

Van Egeren, L. F. (2009). A cybernetic model of global personality traits. *Personality and Social Psychology Review*, *13*(2), 92-108.

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.

Wilson, E. O. (1975). *Sociobiology: The new synthesis.* Cambridge, MA: Harvard University Press.

Wilson, G. D., Gray, J. A., & Barrett, P. T. (1990). A factor analysis of the Gray-Wilson personality questionnaire. *Personality and Individual Differences*, *11*(10),

1037–1044.

Wilt, J. A. (2014). *A new form and function for personality* (Unpublished doctoral dissertation). Northwestern University.

Wilt, J. A., Bleidorn, W., & Revelle, W. (2016). Velocity explains the links between personality states and affect. *Journal of Research in Personality*.

Wilt, J. A., Funkhouser, K., & Revelle, W. (2011). The dynamic relationships of affective synchrony to perceptions of situations. *Journal of Research in Personality*, *45*(3), 309–321.

Yarkoni, T. (2013). Principled failure to detect a five-factor structure in a canonical Big Five measure.

Zinbarg, R. E., & Mohlman, J. (1998). Individual differences in the acquisition of affectively valenced associations. *Journal of personality and social psychology*, *74*(4), 1024–1040.

Zuckerman, M. (2002). Zuckerman-kuhlman personality questionnaire (ZKPQ): An alternative five-factorial model. In B. DeRaad & P. M (Eds.), *Big five assessment* (pp. 376–392). Ashland, OH: Hogrefe and Huber.

Appendices

Appendix A

Original MatLab Code: Pickering's RST

```
clear all;
clc;

rand('seed',2000);
randn('seed',2000);

%wa=sensitivity of the BAS;
%wi=sensitivity of BIS;
%wf=sensitivity of the FFS
%k1 is the activation growth rate constant
%k3 is the activation passive decay rate constant
%k2 is the constant controlling the excitatory (Inhibitory)
   feedback of BIS on FFFS (BAS)

k1a=1;
k1f=k1a;
k1i=k1a;
k3a=0.05;
k3f=k3a;
k3i=k3a;
k2f=0.2;
k2a=k2f;
k4f=0.2;
k4a=k4f;
```

```
maxe=1;
maxebis=1;
maxi=0.5;

%bas = activity and output of BAS
%bis ditto for BIS
%fffs ditto for FFFS
%Rstim is reward (BAS) stimulus
%Fstim is fear (FFS) stimulus
%nsub is number so simulated subjects
%ntest is number of conditions used in the experiment
%nexp is number of simulated experiences per subject
%niter is the number of system activation iterations

nsub=100;
ntest=2;
nexp=200;
niter=100;

stimrandtype=1; %=2 for uniform random variables; =1 for
   normal

if stimrandtype==1;
mn=0.5;
sd=0.2; %0.15;
Rstim = mn + sd.*randn(nsub,nexp+ntest);
Fstim = mn + sd.*randn(nsub,nexp+ntest);
%clip all values to between 0 and 1
for ii=1:nsub;
for jj=1:nexp;
if Rstim(ii,jj)<0;
Rstim(ii,jj)=0;
end;
if Fstim(ii,jj)<0;
Fstim(ii,jj)=0;
end;
if Rstim(ii,jj)>1;
Rstim(ii,jj)=1;
end;
if Fstim(ii,jj)>1;
```

```
Fstim(ii,jj)=1;
end;
end;
end;

else;
Rstim = rand(nsub,nexp+ntest);
Fstim = rand(nsub,nexp+ntest);
end;

Rstim(:,nexp+1)=0.3  ; %low reward condition
Rstim(:,nexp+2)=0.8;   %high reward condition
Fstim(:,nexp+1)=0.3; %low level of fear stimulation
Fstim(:,nexp+2)=0.3; %ditto

wtrandtype=1; %1=normal, 2 =uniform
if wtrandtype==1;
mn=0.5;
sd=0.2; %0.15;
wa= mn + sd.*randn(nsub,1);
wi= mn + sd.*randn(nsub,1);
wf= mn + sd.*randn(nsub,1);

for ii=1:nsub;
if wa(ii)<0;
wa(ii)=0;
end;
if wa(ii)>1;
wa(ii)=1;
end;
if wf(ii)<0;
wf(ii)=0;
end;
if wf(ii)>1;
wf(ii)=1;
end;
if wi(ii)<0;
wi(ii)=0;
end;
if wi(ii)>1;
wi(ii)=1;
```

```
end;
end;

else;
wa= rand(nsub,1);
wi= rand(nsub,1);
wf= rand(nsub,1);
end;

corrcoef(wa,wi)
corrcoef(wa,wf)
corrcoef(wi,wf)
pause;

testit=0; %random
if testit==1; %specific
wa(1:7)=0.5;
wf(1:7)=0.5;
wi(1:7)=0.5;
Fstim(1:7,1)=[0.1 0.2 0.3 0.4 0.5 0.6 0.7]'; %0.6;
Rstim(1:7,1)=Fstim(1:7,1)+0.2;
end;

bas=zeros(nsub,nexp+ntest);
bis=zeros(nsub,nexp+ntest);
fffs=zeros(nsub,nexp+ntest);
basout=bas;
bisout=bis;
fffsout=fffs;

basplot=[];
bisplot=[];
fffsplot=[];

compvers=4;

for i=1:niter;
i
for e=1:nexp+ntest;

if compvers==2;
```

```
dfffs(:,e)=k1f.*(Fstim(:,e).*wf(:)-max(fffs(:,e),0)).*max(
   Fstim(:,e).*wf(:)-k4f.*basout(:,e) + k2f.*bisout(:,e),0)-
   k3f*fffs(:,e);
dbas(:,e)= k1a.*(Rstim(:,e).*wa(:)-max(bas(:,e),0)).*max(Rstim
   (:,e).*wa(:)-k4a.*fffsout(:,e) - k2a.*bisout(:,e),0)-k3a*
   bas(:,e);

elseif compvers==1;
dfffs(:,e)=k1f.*(Fstim(:,e).*wf(:)-max(fffs(:,e),0)).*Fstim(:,
   e).*wf(:)-(max(k4f.*basout(:,e)-k2f.*bisout(:,e),0)+k3f).*
   fffs(:,e);
dbas(:,e)= k1a.*(Rstim(:,e).*wa(:)-max(bas(:,e),0)).*Rstim(:,e
   ).*wa(:)-(k4a.*fffsout(:,e)+k2a.*bisout(:,e)+k3a).*bas(:,e)
   ;

elseif compvers==3; %Grossberg-style
dfffs(:,e)=k1f.*(maxe-max(fffs(:,e),0)).*Fstim(:,e).*wf(:)-(
   maxi+fffs(:,e)).*max(k4f.*basout(:,e)-k2f.*bisout(:,e),0)-
   k3f.*fffs(:,e);
dbas(:,e)= k1a.*(maxe-max(bas(:,e),0)).*Rstim(:,e).*wa(:)-(
   maxi+bas(:,e)).*(k4a.*fffsout(:,e)+k2a.*bisout(:,e))-k3a.*
   bas(:,e);

elseif compvers==4
dfffs(:,e)=k1f.*(maxe-max(fffs(:,e),0)).*(Fstim(:,e).*wf(:)+
   k2f.*bisout(:,e))-(maxi+fffs(:,e)).*k4f.*basout(:,e)-k3f.*
   fffs(:,e);
dbas(:,e)= k1a.*(maxe-max(bas(:,e),0)).*Rstim(:,e).*wa(:)-(
   maxi+bas(:,e)).*(k4a.*fffsout(:,e)+k2a.*bisout(:,e))-k3a.*
   bas(:,e);

else;
end;

bas(:,e)=bas(:,e)+dbas(:,e);
fffs(:,e)=fffs(:,e)+dfffs(:,e);

dbis(:,e)=k1i.*(maxebis-bis(:,e)).*(basout(:,e).*fffsout(:,e)
   .*wi(:))-k3i.*bis(:,e);

bis(:,e)=bis(:,e)+dbis(:,e);
```

```
basout=max(bas,0);
fffsout=max(fffs,0);
bisout=max(bis,0);
end;

meanbas=mean(basout(:,1:nexp)')';
meanfffs=mean(fffsout(:,1:nexp)')';
meanbis=mean(bisout(:,1:nexp)')';
baschg=basout(:,nexp+2)-basout(:,nexp+1);
fffschg=fffsout(:,nexp+2)-fffsout(:,nexp+1);
bischg=bisout(:,nexp+2)-bisout(:,nexp+1);

corrdata=zeros(nsub,6);
corrdata(:,1)=wa;
corrdata(:,2)=wf;
corrdata(:,3)=wi;
corrdata(:,4)=baschg;
corrdata(:,5)=fffschg;
corrdata(:,6)=bischg;

corrcoef(corrdata)

corrdata2=zeros(nsub,6);
corrdata2(:,1)=meanbas;
corrdata2(:,2)=meanfffs;
corrdata2(:,3)=meanbis;
corrdata2(:,4)=baschg;
corrdata2(:,5)=fffschg;
corrdata2(:,6)=bischg;

corrcoef(corrdata2)
```

Appendix B

Original R Code: CTA

```
"cta"<-function(n=3,t=5000,cues=NULL,act=NULL,inhibit=NULL,
   expect=NULL,
consume=NULL,tendency=NULL,tstrength=NULL,type="both",fast=2,
   compare=FALSE,
learn=TRUE,reward=NULL){

##basic model
tf<-function(tendency,cues,step,expect,act,consume){tf<-
   tendency+cues%*%step%*%expect-act%*%step%*%consume}

af<-function(act,tendency,step,tstrength,inhibit){af<-tendency
   %*%step%*%tstrength+act-act%*%step%*%inhibit}

#learning function
ef<-function(expect,act,step,consume,reward){if(learn){
        which.act<-which.max(act)
        if(old.act!=which.act){
                diag(temp)<-act%*%reward
                expect<-expect+temp
                expect<-expect*1/tr(expect)
                old.act<-which.act}
        }
        ef<-expect
}
```

```
temp<-matrix(0,n,n)
if(n>4){colours<-rainbow(n)}else{colours<-c("blue","red","
   black","green")}

stepsize<-.05
tendency.start<-tendency
act.start<-act
expect.start<-expect

if(is.null(cues)){cues<-2^(n-1:n)}
if(is.null(inhibit)){inhibit<-matrix(1,ncol=n,nrow=n)
diag(inhibit)<-.05}
if(is.null(tstrength))tstrength<-diag(1,n)
if(n>1){colnames(inhibit)<-rownames(inhibit)<-paste("A",1:n,
   sep="")}
if(is.null(consume)){consume<-diag(.03,ncol=n,nrow=n)}
step<-diag(stepsize,n)
if(is.null(expect))expect<-diag(1,n)
if(is.null(tendency.start)){tendency<-rep(0,n)}else{tendency<-
   tendency.start}
if(is.null(act.start)){act<-cues}else{act<-act.start}

if(is.null(reward)){reward<-matrix(0,n,n)
diag(reward)<-c(rep(0,n-1),.05)}else{temp1<-reward
reward<-matrix(0,n,n)
diag(reward)<-temp1}

maxact<-minact<-mintendency<-maxtendency<-0
counts<-rep(0,n)
transitions<-matrix(0,ncol=n,nrow=n)
frequency<-matrix(0,ncol=n,nrow=n)
colnames(frequency)<-paste("T",1:n,sep="")
rownames(frequency)<-paste("F",1:n,sep="")
old.act<-which.max(act)

for(iin1:t){
        tendency<-tf(tendency,cues,step,expect,act,consume)
        act<-af(act,tendency,step,tstrength,inhibit)
        act[act<0]<-0
        expect<-ef(expect,act,step,consume,reward)
#STATS
```

```
        maxact<-max(maxact,act)
        minact<-min(minact,act)
        maxtendency<-max(maxtendency,tendency)
        mintendency<-min(mintendency,tendency)
#count
        which.act<-which.max(act)
        counts[which.act]<-counts[which.act]+1
        transitions[old.act,which.act]<-transitions[old.act,
           which.act]+1
        if(old.act!=which.act){
        frequency[old.act,which.act]<-frequency[old.act,which.
           act]+1
        frequency[which.act,which.act]<-frequency[which.act,
           which.act]+1
        }
        old.act<-which.act
}
#PLOTS
plots<-1
action<-FALSE
if(type!="none"){if(type=="state"){
        op<-par(mfrow=c(1,1))
        if(is.null(tendency.start)){tendency<-rep(0,n)}
                else{tendency<-tendency.start}
        if(is.null(act.start)){act<-cues}else{act<-act.start}
        plot(tendency[1],tendency[2],xlim=c(mintendency,
           maxtendency),ylim=c(mintendency,maxtendency),col="
           black",main="Statediagram",xlab="tendency1",ylab="
           tendency2")
        for(iin1:t){
                        tendency<-tf(tendency,cues,step,expect
                           ,act,consume)
                        act<-af(act,tendency,step,tstrength,
                           inhibit)
                        act[act<0]<-0
                        if(!(i%%fast))
                        points(tendency[1],tendency[2],col="
                           black",pch=20,cex=.2)
                }
        }else{
                if(type=="both"){if(compare){op<-par(mfrow=c
```

```
                    (2,2))}
                else{op<-par(mfrow=c(2,1))}
                plots<-2}else{op<-par(mfrow=c(1,1))}
                if(type=="action"){action<-TRUE}
                        else{if(type=="tendencyd")action<-
                           FALSE}
        for(kin1:plots){
                if(is.null(tendency.start)){tendency<-rep(0,n)
                   }
                else{tendency<-tendency.start}
                if(is.null(act.start)){act<-cues}else{act<-act
                   .start}
                if(is.null(expect.start)){expect<-diag(1,n)}
                        else{expect<-expect.start}
                if(action)plot(rep(1,n),act,xlim=c(0,t),ylim=c
                   (minact,maxact),xlab="time",ylab="action",
                   main="Actionsovertime")
                else{plot(rep(1,n),tendency,xlim=c(0,t),ylim=c
                   (mintendency,maxtendency),xlab="time",ylab
                   ="actiontendency",main="
                   Actiontendenciesovertime")}
for(iin1:t){
        tendency<-tf(tendency,cues,step,expect,act,consume)
        act<-af(act,tendency,step,tstrength,inhibit)
        act[act<0]<-0

        maxact<-max(maxact,act)
        minact<-min(minact,act)
        maxtendency<-max(maxtendency,tendency)
        mintendency<-min(mintendency,tendency)
        #count
        which.act<-which.max(act)
        counts[which.act]<-counts[which.act]+1
        transitions[old.act,which.act]<-transitions[old.act,
           which.act]+1
        if(old.act!=which.act){frequency[old.act,which.act]
                <-frequency[old.act,which.act]+1
        expect<-ef(expect,act,step,consume,reward)}
        old.act<-which.act

if(!(i%%fast)){if(action)points(rep(i,n),act,col=colours,cex
```

```
   =.2)
        elsepoints(rep(i,n),tendency,col=colours,cex=.2)}}
action<-TRUE}
}}
results<-list(cues=cues,expectancy=expect,strength=tstrength,
   inihibition=inhibit,
consumation=consume,reinforcement=reward,time=counts,frequency
   =frequency,
tendency=tendency,act=act)

return(results)
```

Appendix C

Core Code: CTA-RST

```
##LIBRARIES/DIRECTORIES
require(psych)

##PARAMETERS
set.seed(2000)
subjects<-240
experiences<-2
maxIter<-240
goals<-2;multsize<-.05
k1BAS=1;k1FFFS=1;k1BIS=1;MaxE=1;MaxEBIS=1
k2BAS<-k2FFFS<-10*0.20
k3BAS<-k3FFFS<-k3BIS<-0.05
k4BAS<-k4FFFS<-10*0.20
MaxI<-0.5

##SENSITIVITIES(INSTIGATING)
source("CTA-RST-Weight-Functions.r")
SensAtT0<-WtDataNormB()
SensArray<-pmax(pmin(SensAtT0,.95),.05)

##CUES(INSTIGATING)
source("CTA-RST-Cue-Functions.R")
CuesAtT0<-CueDataNormB()
CueArray<-pmax(pmin(CuesAtT0,.95),.05)
```

```
##CUES(CONSUMMATION)
source("CTA-RST-Consummation-Functions.r")
ConsQT0<-ConsQNormB()
ConsQArray<-pmax(pmin(ConsQT0,.95),.05)

##SENSITIVITIES(CONSUMMATION)
ConSensT0<-CWtDataNorm()
ConSensArray<-pmax(pmin(ConSensT0,.95),.05)
ConsArray<-ConsQMult()

##TENDENCIES
source("CTA-RST-Tendency-Functions.r")
TendAtT0<-TendArray<-TendPos()
ErrorArray<-EDataNorm()

##ACTS
source("CTA-RST-Action-Functions.r")
ActArray<-ActAt0

##EXPECTANCIES
source("CTA-RST-Expectation-Functions.r")
ExpectArray<-ExpectNull

##REINFORCEMENTS
source("CTA-RST-Reinforcement-Functions.r")
ReinfoArray<-ReDataNull

##EXCITATIONS
source("CTA-RST-Excitation-Functions.r")
ExcitArray<-ExcitConst()

##INHIBITIONS
source("CTA-RST-Inhibition-Functions.r")
InAtT0<-InhibArray<-InhibAP2()


##DATA
#CTA:
maxact<-minact<-mintendency<-maxtendency<-matrix(data=0,nrow=
   subjects,ncol=experiences)
old.act<-which.act<-matrix(data=0,nrow=subjects,ncol=
```

```
   experiences)
spentTime<-array(data=0,c(subjects,experiences,3*goals))
transitions<-frequencies<-array(data=0,c(subjects,experiences
   ,3*goals,3*goals))
tData<-td1Data<-td2Data<-array(data=0,c(3*goals,maxIter,
   subjects,experiences))
aData<-ad1Data<-ad2Data<-array(data=0,c(3*goals,maxIter,
   subjects,experiences))
atData<-ttrData<-ttfData<-array(data=0,c(3*goals,maxIter,
   subjects,experiences))

#EQUATIONS:
ftendency<-function(row,col){ftendency<-TendArray[row,col,]}
ferror<-function(x,y,z){ferror<-ErrorArray[x,y,,z]}
fcueSens<-function(row,col){cues<-CueArray[row,col,]
        Sens<-diag(SensArray[row,],nrow=3*goals,ncol=3*goals)
        fcueSens<-cues%*%Sens}
fexpect<-function(row,col){fexpect<-diag(ExpectArray[row,col
   ,])}
fconsume<-function(row,col)
        {fconsume<-diag(ConsArray[row,col,],nrow=3*goals,ncol
           =3*goals)}
fact<-function(row,col){fact<-ActArray[row,col,]}
fexcite<-function(row,col)
        {fexcite<-matrix(ExcitArray[row,col,],nrow=3*goals,
           ncol=3*goals)}
freinf<-function(row,col)
        {freinf<-diag(ReinfoArray[row,col,],nrow=3*goals, ncol
           =3*goals)}
finhibit<-function(row,col,g=goals){finhibit<-matrix(data=1,
   nrow=3*g,ncol=3*g)
        finh<-matrix(data=InhibArray[row,col,],nrow=3,ncol=3)
        for(n in 1:g){finhibit[c(((3*n)-2):(3*n)),c(((3*n)-2)
           :(3*n))]<-finh}
        return(finhibit)}

mult<-diag(multsize,3*goals)

FTendency<-function(tendency,cueSens,mult,expect,act,consume,
   error){
FTendency<- tendency+cueSens%*%mult%*%expect-act%*%mult%*%
```

```
   consume+error}

FAction<-function(act,tendency,mult,excite,inhibit,error){
FAction<-tendency%*%mult%*%excite+act-act%*%mult%*%inhibit}

FExpect<-function(expect,act,mult,consume,reinf,row,col){
        which.act[row,col]<-which.max(act)
        if(old.act[row,col]!=which.act[row,col]){
                temp<-matrix(0,3*goals,3*goals)
                diag(temp)<-act%*%reinf
                expect<-expect+temp
                Trace<-sum(diag(expect))
                expect<-expect/Trace
                old.act[row,col]<-which.act[row,col]}
        FExpect<-expect}

##PROCESSING:
for(k in 1:subjects)
  {
  for(j in 1:(experiences))
        for (i in 1:maxIter)
          {
                tendency<-ftendency(k,j)
                error<-ferror(k,j,i)
                cueSens<-fcueSens(k,j)
                expect<-fexpect(k,j)
                act<-fact(k,j)
                consume<-fconsume(k,j)
                excite<-fexcite(k,j)
                reinf<-freinf(k,j)
                inhibit <-finhibit(k,j)

#Update:tendencies
TendArray[k,j,]<-tendency<-FTendency(tendency,cueSens,mult,
   expect,act,consume,error)
                maxtendency[k,j]<-max(maxtendency[k,j],
                   TendArray[k,j,])
                mintendency[k,j]<-min(mintendency[k,j],
                   TendArray[k,j,])

#Update:actions
```

```
                act<-FAction(act,tendency,mult,excite,inhibit)
                act[act<0]<-0
                ActArray[k,j,]<-act
                maxact[k,j]<-max(maxact[k,j],ActArray[k,j,])
                minact[k,j]<-min(minact[k,j],ActArray[k,j,])
                which.act[k,j]<-which.max(c(ActArray[k,j,]))
                spentTime[k,j,which.act[k,j]]<-spentTime[k,j,
                   which.act[k,j]]+1
                transitions[k,j,old.act[k,j],which.act[k,j]]
                        <-transitions[k,j,old.act[k,j],which.
                           act[k,j]]+1
                if(old.act[k,j]!=which.act[k,j])
                {frequencies[k,j,old.act[k,j],which.act[k,j]]
                        <-frequencies[k,j,old.act[k,j],which.
                           act[k,j]]+1
                frequencies[k,j,which.act[k,j],which.act[k,j]]
                        <-frequencies[k,j,which.act[k,j],which
                           .act[k,j]]+1}

#Update:expectancies
                expect<-FExpect(expect,act,mult,consume,reinf,
                   k,j)
                ExpectArray[k,j,]<-diag(expect)
                old.act[k,j]<-which.act[k,j]

#Update:(BIS)cues
                for(g in 1:goals)
                {CueArray[k,j,3*g]<-CueArray[k,j,3*g]+(act[(3*
                   g)-2]*act[(3*g)-1])}

#Update:(BIS)(Consummation)cues
        for (g in 1:goals)
        {ConsQArray[k,j,3*g]<-ConsQArray[k,j,3*g]+(act[(3*g)
           -2]*act[(3*g)-1])}
        for(x in 1:goals){ConsArray[k,j,3*x]<-ConsQArray[k,j
           ,3*x]*ConSensArray[k,3*x]}

#Update:Inhibitions
                for(y in 1:(3*goals)){
                if(y%%3 == 1){inhibit[y,y+1]<-k4BAS*(MaxI+
                   tendency[y])
```

```
                inhibit[y,y+2]<-k2BAS*(MaxI+tendency[y])
                }
                if(y%%3 == 2){inhibit[y,y-1]<-k4FFFS*(MaxI+
                   tendency[y])
                inhibit[y,y+1]<-k2FFFS*(tendency[y]-MaxE)}
                }

#UPDATE:DATA
                tData[,i,k,j]<-TendArray[k,j,]
                aData[,i,k,j]<-ActArray[k,j,]
                if(i>=2)
                {td1Data[,i,k,j]<-diag((tData[,i,k,j]-tData[,i
                   -1,k,j])/mult)}
                if(i>=2)
                {ad1Data[,i,k,j]<-diag((aData[,i,k,j]-aData[,i
                   -1,k,j])/mult)}
                if(i>=3)
                {td2Data[,i,k,j]<-diag((td1Data[,i,k,j]-
                   td1Data[,i-1,k,j])/mult)}
                if(i>=3)
                {ad2Data[,i,k,j]<-diag((ad1Data[,i,k,j]-
                   ad1Data[,i-1,k,j])/mult)}
                for(x in 1:(3*goals)){
                if(aData[x,i,k,j]>0){atData[x,i,k,j]<-1}
                if(td1Data[x,i,k,j]>0){ttrData[x,i,k,j]<-1}
                if(td1Data[x,i,k,j]<0){ttfData[x,i,k,j
                   ]<-1}}}}}

##Standardized Instigating Weights
zBAS<-(SensArray[,1]-wBASmean)/wBASsd
zFFFS<-(SensArray[,2]-wFFFSmean)/wFFFSsd
zBIS<-(SensArray[,3]-wBISmean)/wBISsd

##Standardized Consummation Weights
zCBAS<-(ConSensArray[,1]-cwBASmean)/cwBASsd
zCFFFS<-(ConSensArray[,2]-cwFFFSmean)/cwFFFSsd
zCBIS<-(ConSensArray[,3]-cwBISmean)/cwBISsd
```

Appendix D

Function Code: CTA-RST

```
##SENSITIVITIES(INSTIGATING)
#SensAtT0<-WtDataNormB()

##Random normal distributions
wBASmean=.5;  wBASsd=.2
wBISmean=.5;  wBISsd=.2
wFFFSmean=.5; wFFFSsd=.2

WtDataNormB<-function
(s=subjects,g=goals,wBASmu=wBASmean,wBASsg=wBASsd,wBISmu=
   wBISmean,wBISsg=wBISsd,wFFFSmu=wFFFSmean, wFFFSsg=wFFFSsd,
   A1=k1BAS,F1=k1FFFS,I1=k1BIS,ME=MaxE,MB=MaxEBIS)
{WtDataNormalB <- matrix(data=0,nrow=s,ncol=3*g)
 wBAS<-rnorm(s,wBASmu,wBASsg)
 wBIS<-rnorm(s,wBISmu,wBISsg)
 wFFFS<-rnorm(s,wFFFSmu,wFFFSsg)
        for(i in 1:(3*g)){
                if(i%%3 == 0){
                        WtDataNormalB[,i] <- I1*MB*wBIS}
                if(i%%3 == 1){
                        WtDataNormalB[,i] <- A1*ME*wBAS}
                if(i%%3 == 2){
                        WtDataNormalB[,i] <- F1*ME*wFFFS}}
        return(WtDataNormalB)}

##CUES(INSTIGATING)
```

```
#CuesAtT0<-CueDataNormB()

##Random normal distributions for FFFS, BAS cues; BIS starts
   at 0
SBASmean<-.5; SBASsd<-.2; SFFFSmean<-.5; SFFFSsd<-.2

CueDataNormB<-function
(s=subjects,e=experiences,g=goals,SBASmu=SBASmean,SBASsg=
   SBASsd,
SFFFSmu=SFFFSmean,SFFFSsg=SFFFSsd)
{CueDataNorB <- array(data=0,c(s,e,3*g))
        for(i in 1:(3*g)){
                if(i%%3 == 0){
                        tmp <- matrix(data = 0, nrow = s, ncol
                            = e)
                        CueDataNorB[,,i]<-tmp}
                if(i%%3 == 1){
                        tmp <- matrix(data=rnorm(s*e,SBASmu,
                           SBASsg), nrow = s)
                        CueDataNorB[,,i]<-tmp}
                if(i%%3 == 2){
                        tmp <- matrix(data=rnorm(s*e,SFFFSmu,
                           SFFFSsg), nrow = s)
                        CueDataNorB[,,i]<-tmp}}
        return(CueDataNorB)}


##CUES(CONSUMMATION)
#ConsQT0<-ConsQNormB()

##Random normal distributions for FFFS, BAS ConsCues; BIS
   starts at 0
cBASmean<-.5; cBASsd<-.2; cFFFSmean<-.5; cFFFSsd<-.2

ConsQNormB<-function(s=subjects,e=experiences,g=goals,cBASmu=
   cBASmean,cBASsg=cBASsd,cFFFSmu=cFFFSmean,cFFFSsg=cFFFSsd)
{CQNorB <- array(data=0,c(s,e,3*g))
        for(i in 1:(3*g)){
                if(i%%3 == 0){
                        tmp <- matrix(data = 0, nrow = s, ncol
                            = e)
```

```
                        CQNorB[,,i]<-tmp}
                if(i%%3 == 1){
                        tmp <- matrix(data=rnorm(s*e,cBASmu,
                           cBASsg), nrow = s)
                        CQNorB[,,i]<-tmp}
                if(i%%3 == 2){
                        tmp <- matrix(data=rnorm(s*e,cFFFSmu,
                           cFFFSsg), nrow = s)
                        CQNorB[,,i]<-tmp}}
        return(CQNorB)}

##SENSITIVITIES(CONSUMMATION)
#ConSensT0<-CWtDataNorm()

##Random normal distributions for Consummation Weights
cwBASmean=.5;  cwBASsd=.2; cwBISmean=.5;  cwBISsd=.2;
   cwFFFSmean=.5; cwFFFSsd=.2

CWtDataNorm<-function(s=subjects,g=goals,cwBASmu=cwBASmean,
   cwBASsg=cwBASsd,cwBISmu=cwBISmean,cwBISsg=cwBISsd,cwFFFSmu=
   cwFFFSmean,cwFFFSsg=cwFFFSsd,A1=k1BAS,F1=k1FFFS,I1=k1BIS)
{CWtDataN <- matrix(data=0,nrow=s,ncol=3*g)
 cwBAS<-rnorm(s,cwBASmu, cwBASsg)
 cwBIS<-rnorm(s,cwBISmu, cwBISsg)
 cwFFFS<-rnorm(s,cwFFFSmu,cwFFFSsg)
        for(i in 1:(3*g)){
                if(i%%3 == 0){
                        CWtDataN[,i] <- I1*cwBIS}
                if(i%%3 == 1){
                        CWtDataN[,i] <- A1*cwBAS}
                if(i%%3 == 2){
                        CWtDataN[,i] <- F1*cwFFFS}}
        return(CWtDataN)}


##CONSUMMATIONS
#ConsArray<-ConsQMult()

ConsQMult<-function(s=subjects,e=experiences,g=goals,CS=
   ConSensArray,CQ=ConsQArray){
ConsQMlt<-array(data=0,c(s,e,(3*g)))
```

```
        for(i in 1:(3*g)){
                if(i%%3 == 0){ ##BIS
                        tmp <- CQ[,,i]*CS[,i]
                        ConsQMlt[,,i]<-tmp}
                if(i%%3 == 1){ ##BAS
                        tmp <- CQ[,,i]*CS[,i]
                        ConsQMlt[,,i]<-tmp}
                if(i%%3 == 2){ ##FFFS
                        tmp <- CQ[,,i]*CS[,i]
                        ConsQMlt[,,i]<-tmp}}
        return(ConsQMlt)}

##TENDENCIES

#TendAtT0<-TendArray<-TendPos()

PosOff <- 0.05
TendPos<- function(s=subjects,e=experiences,g=goals,po=PosOff)
   {
TP <- array(data=0,c(s,e,3*g))
        for(i in 1:(3*g)){
                if(i%%3 == 1){ #positivity offset
                        TP[,,i]<-po}}
        return(TP)}

#ErrorArray<-EDataNorm()

eBISmean<-0; eBISsd<-.005; eBASmean<-0; eBASsd<-.005;
   eFFFSmean<-0; eFFFSsd<-.005
EDataNorm<-function(s=subjects, e=experiences, g=goals, mi=
   maxIter, eBASmu=eBASmean,eBASsg=eBASsd, eBISmu=eBISmean,
   eBISsg=eBISsd, eFFFSmu=eFFFSmean, eFFFSsg=eFFFSsd)
{EDataNor <- array(data=0,c(s,e,3*g,mi))
  for(x in 1: mi){
        for(i in 1:(3*g)){
                if(i%%3 == 0){
                        tmp <- matrix(data = rnorm(s*e,eBISmu,
                           eBISsg), nrow = s)
                        EDataNor[,,i,x]<-tmp}
                if(i%%3 == 1){
                        tmp <- matrix(data=rnorm(s*e,eBASmu,
```

```
                        eBASsg), nrow = s)
                    EDataNor[,,i,x]<-tmp}
            if(i%%3 == 2){
                    tmp <- matrix(data=rnorm(s*e,eFFFSmu,
                        eFFFSsg), nrow = s)
                    EDataNor[,,i,x]<-tmp}}}
    return(EDataNor)}

##ACTS
#ActArray<-ActAt0

##Start actions at zero:
ActAt0 <- array(data=0,c(subjects,experiences,3*goals))

##EXPECTANCIES
#ExpectArray<-ExpectNull

##Simple diagonal expectancy matrices with 1's on the diagonal
    for everyone
ExpectNull<-array(data=1,c(subjects,experiences,3*goals))

##REINFORCEMENTS
#ReinfoArray<-ReDataNull

##No learning: reinforcements are all 0
ReDataNull<-array(data=0,c(subjects,experiences,3*goals))

##EXCITATIONS
#ExcitArray<-ExcitConst()

##Constant diagonal excitation matrix (diagonals are k5's; all
    other terms are 0)
k5BAS<-k5FFFS<-k5BIS<-1
ExcitConst <- function(s=subjects,e=experiences,g=goals,A5=
   k5BAS,F5=k5FFFS,I5=k5BIS){
ExConst <- array(data=0,c(s,e,(3*g)^2))
        nam<-rep(0,(3*g)^2)
        for(x in 1:((3*g)^2)){nam[x] <- x}
        index<-matrix(nam,3*g,3*g)
        for(i in 1:(3*g)){
                if(i%%3 == 0){ExConst[,,index[i,i]]<-I5}
```

```
                if(i%%3 == 1){ExConst[,,index[i,i]]<-A5}
                if(i%%3 == 2){ExConst[,,index[i,i]]<-F5}}
        return(ExConst)}

##INHIBITIONS
#InAtT0<-InhibArray<-InhibAP2()

BA2BA = k3BAS; BA2FF = k4FFFS*MaxI; BA2BI = 0
FF2BA = k4BAS*MaxI; FF2FF = k3FFFS; FF2BI = 0
BI2BA = k2BAS*MaxI; BI2FF = (-1)*k2FFFS*MaxE; BI2BI = k3BIS
goal2goal = 1 ##These are off-diagonal terms.

InhibAP2<-function(g=goals,s=subjects,e=experiences,a2a=BA2BA,
   f2f=FF2FF,i2i=BI2BI,f2a=FF2BA,i2a=BI2BA,a2f=BA2FF,i2f=BI2FF
   ,a2i=BA2BI,f2i=FF2BI){
        InhibU<-array(data=0,c(s,e,3^2))
        InhibUOnD<-c(a2a,a2f,a2i,f2a,f2f,f2i,i2a,i2f,i2i)
        for(i in 1:(3^2)){InhibU[,,i]<-rep(InhibUOnD[i],s*e)}
        return(InhibU)}

##COMPONENTS (of actions, tendencies)

#DurArray<-Dur() #actions' duration
Dur<- function(at=atData,s=subjects,e=experiences,g=goals,mI=
   maxIter,x=xIter){
        DurData <- array(data=0,c(s,e,3*g))
        for(k in 1:s){
                for(j in 1:e){
                        for(i in 1:(3*g)){
                                DurData[k,j,i]<-(sum(at[i,c((
                                   mI-x):mI),k,j])/(x+1)) }}}
        return(DurData)}

#FVelTArray<-VelTF() #falling velocity of tendencies
VelTF<- function(td1=td1Data,s=subjects,e=experiences,g=goals,
   mI=maxIter,x=xIter){
        FVelTData <- array(data=0,c(s,e,3*g))
        for(k in 1:s){
                for(j in 1:e){
                        for(i in 1:(3*g)){
                                tmp<-0
```

```
                                    for(m in (mI-x+1):mI){
                                    if(td1[i,m,k,j] < 0){
                                    tmp<-tmp+abs(td1[i,m,k,j]) }}
                                    FVelTData[k,j,i]<-(tmp/(x+1))
                                       }}}
        return(FVelTData)}

#PosArray<-Pos()        #actions' intensity/position
Pos<- function(ad=aData,s=subjects,e=experiences,g=goals,mI=
   maxIter,x=xIter,u=multsize){
        PosData <- array(data=0,c(s,e,3*g))
        for(k in 1:s){
                for(j in 1:e){
                        for(i in 1:(3*g)){
                                PosData[k,j,i]<-(sum(ad[i,c((
                                   mI-x):mI),k,j])/(x+1))}}}
        return(PosData)}
```

Appendix E

Parameters and Analyses: Study 1 (Chapter 4)

```
##PARAMETERS
set.seed(2000)
subjects<-240
experiences<-2
maxIter<-240
goals<-2;multsize<-.05
k1BAS=1;k1FFFS=1;k1BIS=1;MaxE=1;MaxEBIS=1
k2BAS<-k2FFFS<-10*0.20
k3BAS<-k3FFFS<-k3BIS<-0.05
k4BAS<-k4FFFS<-10*0.20
MaxI<-0.5

##...Cue and Consummation Matrices Created...##

##EXPERIMENTAL CONDITIONS
##Overwrite Random Instigating Cues, Experience 2:
##E1,Cond1: Positive
CueArray[c(1:40),2,1]<-.95      #Goal 1 BAS
CueArray[c(1:40),2,2]<-.05      #Goal 1 FFFS
CueArray[c(1:40),2,4]<-.5       #Goal 2 BAS
CueArray[c(1:40),2,5]<-.5       #Goal 2 FFFS
##E1,Cond2: Negative
```

```
CueArray[c(41:80),2,1]<-.05     #Goal 1 BAS
CueArray[c(41:80),2,2]<-.95     #Goal 1 FFFS
CueArray[c(41:80),2,4]<-.5      #Goal 2 BAS
CueArray[c(41:80),2,5]<-.5      #Goal 2 FFFS
##E1,Cond3: Neutral
CueArray[c(81:120),2,1]<-.05 #Goal 1 BAS
CueArray[c(81:120),2,2]<-.05 #Goal 1 FFFS
CueArray[c(81:120),2,4]<-.5      #Goal 2 BAS
CueArray[c(81:120),2,5]<-.5      #Goal 2 FFFS

##E4,Cond1: Pleasant
CueArray[c(121:160),2,1]<-.75 #Goal 1 BAS
CueArray[c(121:160),2,2]<-.05 #Goal 1 FFFS
CueArray[c(121:160),2,4]<-.5  #Goal 2 BAS
CueArray[c(121:160),2,5]<-.5  #Goal 2 FFFS
##E4,Cond2: Appetitive
CueArray[c(161:200),2,1]<-.95 #Goal 1 BAS
CueArray[c(161:200),2,2]<-.05 #Goal 1 FFFS
CueArray[c(161:200),2,4]<-.05 #Goal 2 BAS
CueArray[c(161:200),2,5]<-.05 #Goal 2 FFFS
##E4,Cond3: Neutral
CueArray[c(201:240),2,1]<-.05 #Goal 1 BAS
CueArray[c(201:240),2,2]<-.05 #Goal 1 FFFS
CueArray[c(201:240),2,4]<-.5  #Goal 2 BAS
CueArray[c(201:240),2,5]<-.5  #Goal 2 FFFS

##Overwrite Random Consummatory Cues, Experience 2:
##E1,Cond1: Positive
ConsQArray[c(1:40),2,1]<-.5  #Goal 1 BAS
ConsQArray[c(1:40),2,2]<-.05 #Goal 1 FFFS
ConsQArray[c(1:40),2,4]<-.5  #Goal 2 BAS
ConsQArray[c(1:40),2,5]<-.5  #Goal 2 FFFS
##E1,Cond2: Negative
ConsQArray[c(41:80),2,1]<-.05 #Goal 1 BAS
ConsQArray[c(41:80),2,2]<-.5  #Goal 1 FFFS
ConsQArray[c(41:80),2,4]<-.5  #Goal 2 BAS
ConsQArray[c(41:80),2,5]<-.5  #Goal 2 FFFS
##E1,Cond3: Neutral
ConsQArray[c(81:120),2,1]<-.5 #Goal 1 BAS
ConsQArray[c(81:120),2,2]<-.5 #Goal 1 FFFS
ConsQArray[c(81:120),2,4]<-.5 #Goal 2 BAS
```

```
ConsQArray[c(81:120),2,5]<-.5 #Goal 2 FFFS

##E4,Cond1: Pleasant
ConsQArray[c(121:160),2,1]<-.05  #Goal 1 BAS
ConsQArray[c(121:160),2,2]<-.5   #Goal 1 FFFS
ConsQArray[c(121:160),2,4]<-.05  #Goal 2 BAS
ConsQArray[c(121:160),2,5]<-.05  #Goal 2 FFFS
##E4,Cond2: Appetitive
ConsQArray[c(161:200),2,1]<-.25 #Goal 1 BAS
ConsQArray[c(161:200),2,2]<-.95 #Goal 1 FFFS
ConsQArray[c(161:200),2,4]<-.95 #Goal 2 BAS
ConsQArray[c(161:200),2,5]<-.95 #Goal 2 FFFS
##E4,Cond3: Neutral
ConsQArray[c(201:240),2,1]<-.5  #Goal 1 BAS
ConsQArray[c(201:240),2,2]<-.5  #Goal 1 FFFS
ConsQArray[c(201:240),2,4]<-.5  #Goal 2 BAS
ConsQArray[c(201:240),2,5]<-.5  #Goal 2 FFFS

##...CTA-RST Processing...##

##MEASUREMENT-INTERVAL
xIter<-240

##COMPONENTS
source("CTA-RST-Component-Functions.r")
FVelTArray<-VelTF()
PosArray<-Pos()
PoArray<-PosArray+2.225074e-308

##TRAITS
Extra<-function(subj){zCBIS[subj]+zCFFFS[subj]+zCBAS[subj]}

##STATES
PlFarray<-FVelTArray[,,1]
ulPlarray<-PoArray[,,1]
ulUarray<-PoArray[,,2]

##S01:PRE-TEST
Pre.EA<-PlFarray[,1]
Pre.PA<-ulPlarray[,1]
Pre.NA<-ulUarray[,1]
```

```
##S01:POST-TEST
Post.EA<-PlFarray[,2]
Post.PA<-ulPlarray[,2]
Post.NA<-ulUarray[,2]

##S01:ANALYSES

#t-tests:Extraversion
Exdf<-data.frame(Extra(c(1:40)),Extra(c(41:80)),Extra(c
   (81:120)),
        Extra(c(121:160)),Extra(c(161:200)),Extra(c(201:240)))
#Ex1
with(Exdf,t.test(Exdf[,1],Exdf[,2],equal.var=TRUE)) #Pos v.
   Neg, p-value = 0.6516
with(Exdf,t.test(Exdf[,1],Exdf[,3],equal.var=TRUE)) #Pos v.
   Neu, p-value = 0.05248
with(Exdf,t.test(Exdf[,3],Exdf[,2],equal.var=TRUE)) #Neu v.
   Neg,  p-value = 0.1403
#Ex4
with(Exdf,t.test(Exdf[,4],Exdf[,5],equal.var=TRUE)) #Pl v. Ap,
    p-value = 0.05316
with(Exdf,t.test(Exdf[,4],Exdf[,6],equal.var=TRUE)) #Pl v. Neu
   , p-value = 0.5597
with(Exdf,t.test(Exdf[,5],Exdf[,6],equal.var=TRUE)) #Ap v. Neu
   , p-value = 0.2296

##Regressions-E1
E1<-data.frame(Pre.EA[c(1:120)],Post.EA[c(1:120)],Pre.NA[c
   (1:120)],Post.NA[c(1:120)],
Extra(c(1:120)))
R01<-lm(E1[,1]~E1[,5],E1)
R02<-lm(E1[,3]~E1[,5],E1)
R03<-lm(E1[,1]~E1[,3],E1)

##Regressions-E4
E4<-data.frame(Pre.EA[c(121:240)],Post.EA[c(121:240)],Pre.PA[c
   (121:240)],
Post.PA[c(121:240)],Extra(c(121:240)))
R04<-lm(E4[,1]~E4[,5],E4)
R05<-lm(E4[,3]~E4[,5],E4)
```

```
R06<-lm(E4[,1]~E4[,3],E4)

##ANOVA-E1.DATA
E1a<-data.frame
(Pre.EA[c(1:40)],Post.EA[c(1:40)],Pre.NA[c(1:40)],Post.NA[c
   (1:40)],
Pre.EA[c(41:80)],Post.EA[c(41:80)],Pre.NA[c(41:80)],Post.NA[c
   (41:80)],
Pre.EA[c(81:120)],Post.EA[c(81:120)],Pre.NA[c(81:120)],Post.NA
   [c(81:120)])

#Means
E1am<-colMeans(E1a)
#Differences
D01<-E1am[2]-E1am[1]
D01b<-E1am[4]-E1am[3]
D02<-E1am[8]-E1am[7]
D02b<-E1am[6]-E1am[5]
D03<-E1am[10]-E1am[9]
D04<-E1am[12]-E1am[11]

#Dummies
condi<-c(rep("Pos",160),rep("Neg",160),rep("Neu",160))
affe<-c(rep("PA",80),rep("NA",80),rep("PA",80),rep("NA",80),
   rep("PA",80),rep("NA",80))

pre.po<-c(rep("Pre",40),rep("Post",40),rep("Pre",40),rep("Post
   ",40),rep("Pre",40),
rep("Post",40),rep("Pre",40),rep("Post",40),rep("Pre",40),rep
   ("Post",40),rep("Pre",40),
rep("Post",40))

subj1<-c(rep(c(1:40),4),rep(c(1:40),4),rep(c(1:40),4))
subj2<-c(rep(c(1:40),4),rep(c(41:80),4),rep(c(81:120),4))

#Data.Frame
S1E1<-
data.frame(value=stack(E1a)$value,subj=subj1,sub3w=subj2,cond=
   condi,aff=affe,pp=pre.po)
##E1-ANOVA
A01<-aov(value~pp*cond*aff+Error(sub3w/(pp*aff))+cond,data=
```

```
   S1E1)
A02<-aov(value~pp+Error(subj/pp),data=S1E1[c(1:80),])
A03<-aov(value~pp+Error(subj/pp),data=S1E1[c(241:320),])
A04<-aov(value~pp+Error(subj/pp),data=S1E1[c(321:400),])
A05<-aov(value~pp+Error(subj/pp),data=S1E1[c(401:480),])
A06<-aov(value~cond,data=S1E1[c(41:80,201:240,361:400),])
A07<-aov(value~cond,data=S1E1[c(121:160,281:320,441:480),])
A17<-aov(value~pp+Error(subj/pp),data=S1E1[c(81:160),])
A18<-aov(value~pp+Error(subj/pp),data=S1E1[c(161:240),])

##ANOVA-E4.DATA
E4a<-data.frame
(Pre.EA[c(121:160)],Post.EA[c(121:160)],Pre.PA[c(121:160)],
   Post.PA[c(121:160)],
Pre.EA[c(161:200)],Post.EA[c(161:200)],Pre.PA[c(161:200)],Post
   .PA[c(161:200)],
Pre.EA[c(201:240)],Post.EA[c(201:240)],Pre.PA[c(201:240)],Post
   .PA[c(201:240)])

#Means
E4am<-colMeans(E4a)

#Differences
D05<-E4am[2]-E4am[1]
D06<-E4am[4]-E4am[3]
D07<-E4am[6]-E4am[5]
D08<-E4am[8]-E4am[7]
D09<-E4am[10]-E4am[9]
D10<-E4am[12]-E4am[11]

#Dummies
cond4<-c(rep("Pos",160),rep("App",160),rep("Neut",160))
aff4<-c(rep("EA",80),rep("PlA",80),rep("EA",80),rep("PlA",80),
   rep("EA",80),rep("PlA",80))

pre.po4<-c(rep("Pre",40),rep("Post",40),rep("Pre",40),rep("
   Post",40),rep("Pre",40),
rep("Post",40),rep("Pre",40),rep("Post",40),rep("Pre",40),rep
   ("Post",40),rep("Pre",40),
rep("Post",40))
```

```
subj3<-c(rep(c(1:40),4),rep(c(1:40),4),rep(c(1:40),4))
subj4<-c(rep(c(1:40),4),rep(c(41:80),4),rep(c(81:120),4))
subj5<-c(rep(c(121:160),4),rep(c(161:200),4),rep(c(201:240),4)
   )

#Data.Frame
S1E4<-data.frame(value=stack(E4a)$value,subj=subj3,sub3w=subj4
   ,sub3w2=subj5,
cond=cond4,aff=aff4,pp=pre.po4)

##E4-ANOVA
A08<-aov(value~pp*cond*aff+Error(sub3w/(pp*aff))+cond,data=
   S1E4)
A09<-aov(value~pp+Error(subj/pp),data=S1E4[c(1:80),])
A10<-aov(value~pp+Error(subj/pp),data=S1E4[c(81:160),])
A11<-aov(value~pp+Error(subj/pp),data=S1E4[c(161:240),])
A12<-aov(value~pp+Error(subj/pp),data=S1E4[c(241:320),])
A13<-aov(value~pp+Error(subj/pp),data=S1E4[c(321:400),])
A14<-aov(value~pp+Error(subj/pp),data=S1E4[c(401:480),])
A15<-aov(value~cond,data=S1E4[c(41:80,201:240,361:400),])
A16<-aov(value~cond,data=S1E4[c(121:160,281:320,441:480),])

##MMR-E1.DATA
Pre.EA.z1<-scale(E1[,1],center=TRUE,scale=TRUE)
Extra.z1<-scale(E1[,5],center=TRUE,scale=TRUE)
Post.EA.s<-stack(data.frame(Post.EA[c(1:40)],Post.EA[c(41:80)
   ],Post.EA[c(81:120)]))$value
Pcontr<-c(rep("Pos",40),rep("NPos",80))
Ncontr<-c(rep("NNeg",40),rep("Neg",40),rep("NNeg",40))
E1m<-data.frame(Post.EA.s,Pre.EA.z1,Extra.z1,Pcontr,Ncontr)

##E1-MMR
M01<-lm(E1m[,1]~E1m[,2],E1m)
M02<-lm(E1m[,1]~E1m[,2]+E1m[,3]+E1m[,4]+E1m[,5],E1m)
M03<-lm(E1m[,1]~E1m[,2]+E1m[,3]+E1m[,4]+E1m[,5]+E1m[,3]*E1m
   [,4]+E1m[,3]*E1m[,5],E1m)
M04<-lm(E1m[c(1:40),1]~E1m[c(1:40),3],E1m[c(1:40),])
M05<-lm(E1m[c(41:80),1]~E1m[c(41:80),3],E1m[c(41:80),])
M06<-lm(E1m[c(81:120),1]~E1m[c(81:120),3],E1m[c(81:120),])

##MMR-E4.DATA
```

```
Pre.EA.z4<-scale(E4[,1],center=TRUE,scale=TRUE)
Pre.PlA.z4<-scale(E4[,3],center=TRUE,scale=TRUE)
Extra.z4<-scale(E4[,5],center=TRUE,scale=TRUE)

Post.EA.s4<-stack(data.frame(Post.EA[c(121:160)],Post.EA[c
   (161:200)],
Post.EA[c(201:240)]))$value

Post.PlA.s<-stack(data.frame(Post.PA[c(121:160)],Post.PA[c
   (161:200)],
Post.PA[c(201:240)]))$value

Plcontr<-c(rep("Pl",40),rep("NPl",80))
Econtr<-c(rep("NEA",40),rep("EA",40),rep("NEA",40))
E4m<-data.frame(Post.EA.s4,Post.PlA.s,Pre.EA.z4,Pre.PlA.z4,
   Extra.z4,Plcontr,Econtr)

##E4-MMR
M07<-lm(E4m[,1]~E4m[,3],E4m)
M08<-lm(E4m[,1]~E4m[,3]+E4m[,5]+E4m[,6]+E4m[,7],E4m)
M09<-lm(E4m[,1]~E4m[,3]+E4m[,5]+E4m[,6]+E4m[,7]+E4m[,5]*E4m
   [,6]+E4m[,5]*E4m[,7],E4m)
M10<-lm(E4m[c(1:40),1]~E4m[c(1:40),5],E4m[c(1:40),])
M11<-lm(E4m[c(41:80),1]~E4m[c(41:80),5],E4m[c(41:80),])
M12<-lm(E4m[c(81:120),1]~E4m[c(81:120),5],E4m[c(81:120),])
M13<-lm(E4m[,2]~E4m[,4],E4m)
M14<-lm(E4m[,2]~E4m[,4]+E4m[,5]+E4m[,6]+E4m[,7],E4m)
M15<-lm(E4m[,2]~E4m[,4]+E4m[,5]+E4m[,6]+E4m[,7]+E4m[,5]*E4m
   [,6]+E4m[,5]*E4m[,7],E4m)
M16<-lm(E4m[c(1:40),2]~E4m[c(1:40),5],E4m[c(1:40),])
M17<-lm(E4m[c(41:80),2]~E4m[c(41:80),5],E4m[c(41:80),])
M18<-lm(E4m[c(81:120),2]~E4m[c(81:120),5],E4m[c(81:120),])
```

Appendix F

Parameters and Analyses: Study 2 (Chapter 5)

```
##LIBRARIES/DIRECTORIES
require(psych)
require(nlme)
require(MplusAutomation)

##PARAMETERS
set.seed(2500)
subjects<-80
experiences<-63
maxIter<-900
goals<-2;multsize<-.05
k1BAS=1;k1FFFS=1;k1BIS=1;MaxE=1;MaxEBIS=1
k2BAS<-k2FFFS<-10*0.20
k3BAS<-k3FFFS<-k3BIS<-0.05
k4BAS<-k4FFFS<-10*0.20
MaxI<-0.5

##...CTA-RST Processing...##

##MEASUREMENT-INTERVAL
xIter<-300

##COMPONENTS
```

```
source("CTA-RST-Component-Functions.r")
FVelTArray<-VelTF()
PosArray<-Pos()
PoArray<-PosArray+2.225074e-308

##STATES
ulPlSarray<-scale(PoArray[,,1])+scale(PoArray[,,4])
ulUSarray<-scale(PoArray[,,2])+scale(PoArray[,,5])
EAFarray<-scale(FVelTArray[,,1])+scale(FVelTArray[,,4])
FFArray<-scale(FVelTArray[,,2])+scale(FVelTArray[,,5])
VelEarray<-scale(DurArray[,,1])+scale(DurArray[,,4])

##S01:ANALYSES

#Means
mStExE<-rowMeans(ExtraSarray);mStNeE<-rowMeans(NeuroSarray);
mVelEE<-rowMeans(VelEarray);mStPAE<-rowMeans(PAarray);mStNAE<-
   rowMeans(NAarray)
mStEAE<-rowMeans(EAarray);mStPlAE<-rowMeans(PlAarray)
mStUAE<-rowMeans(UAarray);mStTAE<-rowMeans(TAarray)

#WPSD
sdStExE<-describe(t(ExtraSarray))$sd;sdStNeE<-describe(t(
   NeuroSarray))$sd
sdVelEE<-describe(t(VelEarray))$sd
sdStPAE<-describe(t(PAarray))$sd;sdStNAE<-describe(t(NAarray))
   $sd
sdStEAE<-describe(t(EAarray))$sd;sdStPlAE<-describe(t(PlAarray
   ))$sd
sdStUAE<-describe(t(UAarray))$sd;sdStTAE<-describe(t(TAarray))
   $sd

#WPalpha
aStEx<-alpha(ExtraSarray);aStNe<-alpha(NeuroSarray);aVelE<-
   alpha(VelEarray)
aStPA<-alpha(PAarray);aStNA<-alpha(NAarray)
aStEA<-alpha(EAarray);aStPlA<-alpha(PlAarray);aStUA<-alpha(
   UAarray);aStTA<-alpha(TAarray)

#Typical.mean
GmStEx<-mean(mStExE);GmStNe<-mean(mStNeE);GmVelE<-mean(mVelEE)
```

```
GmStPA<-mean(mStPAE);GmStNA<-mean(mStNAE)
GmStEA<-mean(mStEAE);GmStPlA<-mean(mStPlAE);GmStUA<-mean(
   mStUAE);GmStTA<-mean(mStTAE)

#Typical.WPSD
GsdStEx<-mean(sdStExE);GsdStNe<-mean(sdStNeE);GsdVelE<-mean(
   sdVelEE)
GsdStPA<-mean(sdStPAE);GsdStNA<-mean(sdStNAE)
GsdStEA<-mean(sdStEAE);GsdStPlA<-mean(sdStPlAE);GsdStUA<-mean(
   sdStUAE)
GsdStTA<-mean(sdStTAE)

#Descriptives
Means<-c(GmStEx,GmStNe,GmVelE,GmStPA,GmStNA)
Alphas<-c(aStEx$total$raw_alpha,aStNe$total$raw_alpha,
   aVelE$total$raw_alpha,
        aStPA$total$raw_alpha,aStNA$total$raw_alpha)
WPSDs<-c(GsdStEx,GsdStNe,GsdVelE,GsdStPA,GsdStNA)
Means;Alphas;WPSDs

##Bivariate-MLM.Data
ctStEx<-scale(t(ExtraSarray),center=TRUE,scale=FALSE)
ctStNe<-scale(t(NeuroSarray),center=TRUE,scale=FALSE)
ctVel<-scale(t(VelEarray),center=TRUE,scale=FALSE)
ctStPA<-scale(t(PAarray),center=TRUE,scale=FALSE)
ctStNA<-scale(t(NAarray),center=TRUE,scale=FALSE)

pers<-rep(1:80,each=63);obs<-rep(1:63,80)
stackedEx<-c(ctStEx);stackedNe<-c(ctStNe);stackedVel<-c(ctVel)
stackedPA<-c(ctStPA);stackedNA<-c(ctStNA)

mlmd<-data.frame(pers=pers,obs=obs,STEX=stackedEx,STPA=
   stackedPA,STNE=stackedNe,
        STNA=stackedNA,VEL=stackedVel)

##Bivariate-MLMs:

#1:Ex->V
N.Ex2Vel<-lme(VEL~STEX,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Ex2Vel<-lme(VEL~STEX,random=~STEX|pers,data=mlmd,method="ML
```

```
   ",control=list(opt="optim"))
anova(N.Ex2Vel,A.Ex2Vel);summary(A.Ex2Vel);intervals(A.Ex2Vel,
   which="fixed")

#2.Ne->V
N.Ne2Vel<-lme(VEL~STNE,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Ne2Vel<-lme(VEL~STNE,random=~STNE|pers,data=mlmd,method="ML
   ",control=list(opt="optim"))
anova(N.Ne2Vel,A.Ne2Vel);summary(A.Ne2Vel);intervals(A.Ne2Vel,
   which="fixed")

#3.V->PA
N.Vel2PA<-lme(STPA~VEL,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Vel2PA<-lme(STPA~VEL,random=~VEL|pers,data=mlmd,method="ML",
    control=list(opt="optim"))
anova(N.Vel2PA, A.Vel2PA);summary(A.Vel2PA);intervals(A.Vel2PA
   ,which="fixed")

#4.V->NA
N.Vel2NA<-lme(STNA~VEL,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Vel2NA<-lme(STNA~VEL,random=~VEL|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
anova(N.Vel2NA, A.Vel2NA);summary(A.Vel2NA);intervals(A.Vel2NA
   ,which="fixed")

#5.Ex->PA
N.Ex2PA<-lme(STPA~STEX,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Ex2PA<-lme(STPA~STEX,random=~STEX|pers,data=mlmd,method="ML
   ",control=list(opt="optim"))
anova(N.Ex2PA, A.Ex2PA);summary(A.Ex2PA);intervals(A.Ex2PA,
   which="fixed")

#6.Ne->NA
N.Ne2NA<-lme(STNA~STNE,random=~1|pers,data=mlmd,method="ML",
   control=list(opt="optim"))
A.Ne2NA<-lme(STNA~STNE,random=~STNE|pers,data=mlmd,method="ML
   ",control=list(opt="optim"))
```

```
anova(N.Ne2NA, A.Ne2NA);summary(A.Ne2NA);intervals(A.Ne2NA,
   which="fixed")

##111-MLM.Data
mtStEx<-t(mStExE);mtStNe<-t(mStNeE);mtVel<-t(mVelEE);mtStPA<-t
   (mStPAE);mtStNA<-t(mStNAE)
mtStEA<-t(mStEAE);mtStPlA<-t(mStPlAE);mtStUA<-t(mStUAE);mtStNA
   <-t(mStTAE)

MEX<-rep(mtStEx,each=63);MNE<-rep(mtStNe,each=63);MVEL<-rep(
   mtVel,each=63)
MPA<-rep(mtStPA,each=63);MNA<-rep(mtStNA,each=63)

mlmdb<-data.frame(pers=pers,obs=obs,STEX=stackedEx,MEX=MEX,
   STPA=stackedPA,MPA=MPA,
        STNE=stackedNe,MNE=MNE,STNA=stackedNA,MNA=MNA,VEL=
           stackedVel,MVEL=MVEL)

prepareMplusData
        (mlmdb,"C:/Users/Ashley/Documents/@-Projects-Hot/Y4-
           PhD/PhD-Simulation/mlmdb.dat")

##Bivariate-MLMs.MPlus:

#1.Ex,V,PA
TITLE: Study 2 (1-1-1 MLMs)
DATA: FILE = "C:/Users/Ashley/Documents/@-Projects-Hot/Y4-PhD/
   PhD-Simulation/mlmdb.dat";
VARIABLE: NAMES ARE
pers obs STEX MEX STPA MPA STNE MNE STNA MNA VEL MVEL;
USEVARIABLES ARE
pers, STEX, STPA, VEL;
MISSING ARE all (-999);
CLUSTER IS pers;
ANALYSIS: TYPE IS TWOLEVEL RANDOM;
!ALGORITHM = INTEGRATION;
MODEL:
%WITHIN%
STPA ON STEX(pew);
STPA ON VEL(pvw);
VEL ON STEX(vew);
```

```
%BETWEEN%
VEL STPA STEX;
STPA ON STEX(peb);
STPA ON VEL(pbb);
VEL ON STEX(veb);
MODEL CONSTRAINT:
NEW(indep total);
indep = vew*pvw;
total = pew + vew*pvw;
OUTPUT: CINTERVAL;
!TECH1 TECH8

#2.Ne,V,NA
TITLE: Study 2NA (1-1-1 MLMs)
DATA: FILE = "mlmdb.dat";
VARIABLE: NAMES ARE
pers obs STEX MEX STPA MPA STNE MNE STNA MNA VEL MVEL;
USEVARIABLES ARE
pers, STNE, STNA, VEL;
MISSING ARE all (-999);
CLUSTER IS pers;
ANALYSIS: TYPE IS TWOLEVEL RANDOM;
MODEL:
%WITHIN%
STNA ON STNE(pew);
STNA ON VEL(pvw);
VEL ON STNE(vew);
%BETWEEN%
VEL STNA STNE;
STNA ON STNE(peb);
STNA ON VEL(pbb);
VEL ON STNE(veb);
MODEL CONSTRAINT:
NEW(indep total);
indep = vew*pvw;
total = pew + vew*pvw;
OUTPUT: TECH1 CINTERVAL;
!TECH8
```

Appendix G

Parameters and Analyses: Study 3 (Chapter 6)

```
##LIBRARIES/DIRECTORIES
require(psych)
require(nlme)
library(matrixStats)

##PARAMETERS
set.seed(1500)
subjects<-82
experiences<-63
maxIter<-900
goals<-2;multsize<-.05
k1BAS=1;k1FFFS=1;k1BIS=1;MaxE=1;MaxEBIS=1
k2BAS<-k2FFFS<-10*0.20
k3BAS<-k3FFFS<-k3BIS<-0.05
k4BAS<-k4FFFS<-10*0.20
MaxI<-0.5

##...CTA-RST Processing...##

source("CTA-RST-Component-Functions.r")

##MEASUREMENT-INTERVAL(Traits)
xIter<-900
```

```
##COMPONENTS(Traits)
PosArray<-Pos()
PoArray<-PosArray+2.225074e-308
FVelTArray<-VelTF()
##TRAITS:
ulPlSarray<-scale(PoArray[,,1])+scale(PoArray[,,4])
ulUSarray<-scale(PoArray[,,2])+scale(PoArray[,,5])
EAFarray<-scale(FVelTArray[,,1])+scale(FVelTArray[,,4])
FFArray<-scale(FVelTArray[,,2])+scale(FVelTArray[,,5])

EAFm<-rowMeans(EAFarray)
FFm<-rowMeans(FFArray)
ulPlm<-rowMeans(ulPlSarray)
ulUm<-rowMeans(ulUSarray)

TEA<-EAFm
TTA<-FFm
TPA<-ulPlm
TUA<-ulUm

##MEASUREMENT-INTERVAL(States:EA,TA)
xIter<-300
##COMPONENTS(States:EA,TA)
FVelTArray<-VelTF()
##STATES(EA,TA)
EAFarray<-scale(FVelTArray[,,1])+scale(FVelTArray[,,4])
FFArray<-scale(FVelTArray[,,2])+scale(FVelTArray[,,5])

SEA<-EAFarray
STA<-FFArray

##MEASUREMENT-INTERVAL(States:PA,UA)
xIter<-30
##COMPONENTS(States:PA,UA)
PosArray<-Pos()
PoArray<-PosArray+2.225074e-308
##STATES(PA,UA)
ulPlSarray<-scale(PoArray[,,1])+scale(PoArray[,,4])
ulUSarray<-scale(PoArray[,,2])+scale(PoArray[,,5])

SPA<-ulPlSarray
```

```
SUA<-ulUSarray

##S03: ANALYSES

##Find "mean across all reports"
AmStEAE<-mean(rowMeans(SEA));AmStUAE<-mean(rowMeans(SUA))
AmStTAE<-mean(rowMeans(STA));AmStPAE<-mean(rowMeans(SPA))

##Find "SD aggregated across all reports"
AsdStEAE<-sd(SEA);AsdStUAE<-sd(SUA)
AsdStTAE<-sd(STA);AsdStPAE<-sd(SPA)

##Find "WP SD aggregated across participants"
WpsdStEAE<-mean(rowSds(SEA));WpsdStTAE<-mean(rowSds(STA))
WpsdStUAE<-mean(rowSds(SUA));WpsdStPAE<-mean(rowSds(SPA))

##Find "BP SD aggregated across participants"
BpsdStEAE<-mean(colSds(SEA));BpsdStTAE<-mean(colSds(STA));
BpsdStUAE<-mean(colSds(SUA));BpsdStPAE<-mean(colSds(SPA))

EASim <-c(AmStEAE,AsdStEAE,BpsdStEAE,WpsdStEAE)
TASim <-c(AmStTAE,AsdStTAE,BpsdStTAE,WpsdStTAE)
PASim <-c(AmStPAE,AsdStPAE,BpsdStPAE,WpsdStPAE)
UASim <-c(AmStUAE,AsdStUAE,BpsdStUAE,WpsdStUAE)
EASim;TASim;PASim;UASim

Traits<-data.frame(TEA,TTA,TPA,TUA)
describe(Traits)

##Mean-center predictor variables (TA, UA) around each person'
   s mean
ctStTA<-scale(t(STA),center=TRUE,scale=FALSE);ctStUA<-scale(t(
   SUA),center=TRUE,scale=FALSE)

##Make a data frame for all multilevel models:
pers<-rep(1:82,each=63);obs<-rep(1:63,82)
stackedEA<-c(t(SEA));stackedTA<-c(ctStTA);stackedPA<-c(t(SPA))
   ;stackedUA<-c(ctStUA)

mlmd3<-data.frame
(pers=pers,obs=obs,STEA=stackedEA,STTA=stackedTA,STPA=
```

```
   stackedPA,STUA=stackedUA)

##Bivariate multilevel models
##1. TA predicts EA
Null3EA<-lme(STEA~STTA,random=~1|pers,data=mlmd3,control=list(
   opt="optim"))
PtaOea<-lme(STEA~STTA,random=~STTA|pers,data=mlmd3,control=
   list(opt="optim"))
BRPtaOea<-coef(PtaOea)
anova(Null3EA, PtaOea);summary(PtaOea);describe(BRPtaOea)

##2. UA predicts PA
Null3PA<-lme(STPA~STUA,random=~1|pers,data=mlmd3,control=list(
   opt="optim"))
PuaOpa<-lme(STPA~STUA,random=~STUA|pers,data=mlmd3,control=
   list(opt="optim"))
BRPuaOpa<-coef(PuaOpa)
anova(Null3PA, PuaOpa);summary(PuaOpa);describe(BRPuaOpa)

##Find correlation between EA-TA and UA-PA relationship
slopes<-data.frame(BRPtaOea[,2],BRPuaOpa[,2])
corSl<-cor(slopes)
corSl

##Do trait-interaction multilevel models
##Mean-center predictor variables:
cTrEA<-scale(TEA,center=TRUE,scale=FALSE);cTrTA<-scale(TTA,
   center=TRUE,scale=FALSE)
cTrPA<-scale(TPA,center=TRUE,scale=FALSE);cTrUA<-scale(TUA,
   center=TRUE,scale=FALSE)

##Stack data and make data frames:
stackedTrEA<-rep(cTrEA,each=63);stackedTrTA<-rep(cTrTA,each
   =63)
stackedTrPA<-rep(cTrPA,each=63);stackedTrUA<-rep(cTrUA,each
   =63)

mlmd4<-data.frame
(pers=pers,obs=obs,STEA=stackedEA,STTA=stackedTA,TREA=
   stackedTrEA,TRTA=stackedTrTA)
```

```
mlmd5<-data.frame
(pers=pers,obs=obs,STPA=stackedPA,STUA=stackedUA,TRPA=
   stackedTrPA,TRUA=stackedTrUA)

mlmd6<-data.frame
(pers=pers,obs=obs,STEA=stackedEA,STTA=stackedTA,TRPA=
   stackedTrPA,TRUA=stackedTrUA)

mlmd7<-data.frame
(pers=pers,obs=obs,STPA=stackedPA,STUA=stackedUA,TREA=
   stackedTrEA,TRTA=stackedTrTA)

##1.MLMs predicting state EA from state TA, trait EA, trait TA
   :
NulModEA1 <- lme(STEA ~ STTA + TREA + TRTA + STTA:TREA + STTA:
   TRTA + TREA:TRTA + STTA:TREA:TRTA, random = ~1|pers, data =
    mlmd4, method = "ML", control=list(opt="optim"))

AltModEA1 <- lme(STEA ~ STTA + TREA + TRTA + STTA:TREA + STTA:
   TRTA + TREA:TRTA + STTA:TREA:TRTA,random = ~STTA|pers, data
    = mlmd4,method = "ML",control=list(opt="optim"))

anova(NulModEA1, AltModEA1); summary(AltModEA1); intervals(
   AltModEA1)

##2. MLMs predicting state EA from state TA, trait PA, trait
   UA:
NulModEA2 <- lme(STEA ~ STTA + TRPA + TRUA + STTA:TRPA + STTA:
   TRUA + TRPA:TRUA + STTA:TRPA:TRUA, random = ~1|pers, data =
    mlmd6, method = "ML", control=list(opt="optim"))

AltModEA2 <- lme(STEA ~ STTA + TRPA + TRUA + STTA:TRPA + STTA:
   TRUA + TRPA:TRUA + STTA:TRPA:TRUA, random = ~STTA|pers,data
    = mlmd6,method = "ML",control=list(opt="optim"))

anova(NulModEA2, AltModEA2); summary(AltModEA2); intervals(
   AltModEA2)

##3.MLMs predicting state PA from state UA, trait PA, trait UA
   :
NulModPA1 <- lme(STPA ~ STUA + TRPA + TRUA + STUA:TRPA + STUA:
```

```
   TRUA + TRPA:TRUA + STUA:TRPA:TRUA, random = ~1|pers, data =
    mlmd5, method = "ML", control=list(opt="optim"))

AltModPA1 <- lme(STPA ~ STUA + TRPA + TRUA + STUA:TRPA + STUA:
   TRUA + TRPA:TRUA + STUA:TRPA:TRUA,random = ~STUA|pers,data
   = mlmd5,method = "ML",control=list(opt="optim"))

anova(NulModPA1, AltModPA1); summary(AltModPA1); intervals(
   AltModPA1)

##4.MLMs predicting state PA from state UA, trait EA, trait TA
   :
NulModPA2 <- lme(STPA ~ STUA + TREA + TRTA + STUA:TREA + STUA:
   TRTA + TREA:TRTA + STUA:TREA:TRTA, random = ~1|pers, data =
    mlmd7, method = "ML", control=list(opt="optim"))

AltModPA2 <- lme(STPA ~ STUA + TREA + TRTA + STUA:TREA + STUA:
   TRTA + TREA:TRTA + STUA:TREA:TRTA, random = ~STUA|pers,data
    = mlmd7,method = "ML",control=list(opt="optim"))

anova(NulModPA2, AltModPA2); summary(AltModPA2); intervals(
   AltModPA2)

##EXPLORATORY REGRESSIONS & MLMs

##Regress traits on instigating (z) and consummation (zC)
   weights
SDF<-data.frame(TPA,TUA,TEA,TTA,zBAS,zFFFS,zBIS,zCBAS,zCFFFS,
   zCBIS)

SDF03<-lm(SDF[,1]~SDF[,5]*SDF[,6]*SDF[,7]*SDF[,8]*SDF[,9]*SDF
   [,10],SDF) #TPA
SDF04<-lm(SDF[,2]~SDF[,5]*SDF[,6]*SDF[,7]*SDF[,8]*SDF[,9]*SDF
   [,10],SDF) #TUA
SDF05<-lm(SDF[,3]~SDF[,5]*SDF[,6]*SDF[,7]*SDF[,8]*SDF[,9]*SDF
   [,10],SDF) #TEA
SDF06<-lm(SDF[,4]~SDF[,5]*SDF[,6]*SDF[,7]*SDF[,8]*SDF[,9]*SDF
   [,10],SDF) #TTA
summary(SDF06); confint(SDF06,level=.95)

##Predict state affective synchrony from instigating cue
```

```
   weights

##Data frame:
pers<-rep(1:82,each=63);obs<-rep(1:63,82)
stackedEA<-c(t(SEA));stackedTA<-c(ctStTA);stackedPA<-c(t(SPA))
   ;stackedUA<-c(ctStUA)
szBA<-rep(zBAS,each=63);szBI<-rep(zBIS,each=63);szFF<-rep(
   zFFFS,each=63)
md1<-data.frame(pers=pers,obs=obs,STEA=stackedEA,STTA=
   stackedTA,zBA=szBA,zBI=szBI,zFF=szFF)
md2<-data.frame(pers=pers,obs=obs,STPA=stackedPA,STUA=
   stackedUA,zBA=szBA,zBI=szBI,zFF=szFF)

##1.MLMs predicting state EA from state TA & instigating
   sensitivities
NEA <- lme(STEA ~ STTA + zBA + zBI + zFF + zBA:zBI + zBA:zFF +
    zFF:zBI + zBA:zBI:zFF, random = ~1|pers, data = md1,
   method = "ML",control=list(opt="optim"))

AEA <- lme(STEA ~ STTA + zBA + zBI + zFF + zBA:zBI + zBA:zFF +
    zFF:zBI + zBA:zBI:zFF, random = ~STTA|pers, data = md1,
   method = "ML", control=list(opt="optim"))

anova(NEA, AEA); summary(AEA); intervals(AEA)

##2.MLMs predicting state PA from state UA & instigating
   sensitivities
NPA <- lme(STPA ~ STUA + zBA + zBI + zFF + zBA:zBI + zBA:zFF +
    zFF:zBI + zBA:zBI:zFF, random = ~1|pers, data = md2,
   method = "ML", control=list(opt="optim"))

APA <- lme(STPA ~ STUA + zBA + zBI + zFF + zBA:zBI + zBA:zFF +
    zFF:zBI + zBA:zBI:zFF, random = ~STUA|pers, data = md2,
   method = "ML", control=list(opt="optim"))

anova(NPA, APA); summary(APA); intervals(APA)
```