The Contribution of Reinforcement Sensitivity Theory to Personality Theory

William Revelle
Northwestern University

12/18/06

William Revelle
Department of Psychology
Northwestern University
Evanston, Illinois
email: revelle@northwestern.edu

Contents

The breadth of Jeffrey Grays contributions to psychology is barely touched in the various chapters of this book. This is not a criticism of the authors, but of the task that the editor set them. For to honor Jeffrey Gray in a single volume is an impossible task. As all of the contributers have said so well, Jeffrey was a remarkable man. Warm and caring to his friends and colleagues, linguistically skilled, familiar with theater and music in multiple languages, Jeffrey developed a biological model of personality that has become so well established that it has persisted even after he drastically changed it. To those who study personality psychology, the preceding chapters of this book are essential requirements that allow us to take advantage of the careful research and broad scholarship that went into Jeffrey Grays work.

Personality theory is concerned with describing and explaining the observed complexity of individual differences in the patterning of affect, behavior, cognition, and desires over time and space. At the descriptive level, there is moderately strong consensus that three to six broad domains can be used to organize how people describe themselves and others. People differ in the power they exert over others, in the affection they show others, in the quality of their work, in their tendency to react emotionally, in the openness of their interests to intellectual inquiry, and in their basic intellectual ability. In his discussions of adult personality, particularly in the work place, Hogan and his colleagues has distinguished between personality as we see ourselves (our identity) and personality as others see us (our reputation). He categorizes the tasks that we face as those of getting along and getting ahead and suggests that much of human behavior is in interaction with other humans as we are faced with agentic and communal roles (R. Hogan, J. Hogan and Roberts, 1996; R. Hogan and Kaiser, 2005; R. Hogan and Roberts, 2004).

Although there is agreement that most if not all of the dimensions of personality at the descriptive level have moderate additive heritabilities (McGue & Bouchard, 1998), this says nothing about the mechanism. For those who worry about mechanism (i.e., the readers of this book), the path from genes to behavior must go through proteins, biological networks, structures and systems which in turn are tuned by the individuals history of environmental experience. (Individual differences in the capacity to learn language or to be religious are heritable, but what language one uses or what religion one espouses is a function of environmental experience.)

The proposed biological mechanisms should be consistent with our knowledge of evolutionary theory. A fundamental assumption of evolution is preservation of function across species. Hearts, lungs, the circulatory system, and the central nervous system did not evolve independently in different species but rather were selectively modified to meet particular evolutionary challenges. It might be comforting to think that human affect, behavior, and cognition are unique, but the uniqueness is one of degree, not of kind.

Jeffrey Gray took advantage of this assumption of evolutionary preservation and studied the rat, particularly the rats response to anti-anxiety drugs, to study human personality.

Earlier versions of "the Gray model" have become so well established in the field that they can be thought of as the standard model for the biological basis of personality. Variations of the dimensional analysis of the Behavioral Inhibition System and the Behavioral Approach System has been integrated into most causal theories of personality. Whether it is the BIS and BAS of Gray (1970), of Fowles (1980, 1984), or of Depue, and Collins (1999) the two systems approach has been the foundation of biologically driven theories.

But the original Gray model was wrong. It failed the basic criterion of science in that it did not predict what it should have predicted and did not match our growing understanding of biological mechanisms. The best evidence for the models failure was, of course, in the later work of Jeffrey Gray and his colleagues (Corr, 2004; Gray and McNaughton, 2000; Pickering and Gray, 2001) . This was the real strength of Jeffreys contributions: He was able to develop a testable theory that could be rejected based upon data. Not only did he develop such a theory, he was the one who was able to put together the evidence of its failure. But not content with merely testing (and rejecting) his prior work, Jeffrey continued to develop a much more complex model that is both biologically more accurate and simultaneously harder to test. For the newer theory attempts to model the complexity of three interacting brain systems, each with multiple subsystems.

Although many of us have referred to "Grays theory of personality", as Pickering (Chapter 16) points out, Jeffrey himself thought this was pompous and preferred to use the term – suggested by Pickering – Reinforcement Sensitivity Theory. I think this label is unfortunate, for although this book discusses Reinforcement Sensitivity Theory, Jeffrey's major contribution to personality theory should probably better be known as Three Systems Theory (3ST). The biologically and behaviorally based distinctions between the Behavioral Inhibition System, the Behavioral Approach System, and the Fight/Freeze/Flight System has had a greater impact upon personality theory than the fundamental analysis of how these systems respond to appetitive and aversive cues in the process of learning. Just as the original distinction between analyzing behavior in terms of action tendencies (BAS) and inhibition of action (the BIS) made contact with non-biological models of action (e.g., Atkinson and Birch, 1970) that focused on general action and negaction tendencies, so can we hope that the Three Systems Theory will lead to broader conceptualizations of the complexity of human behavior.

RST is a projective test: Everyone brings their own theoretical biases and fits them into RST. RST is a fundamental approach to the problem of the biological basis of personality. RST recognizes the importance of analyzing ongoing behavior in the face of potentially aversive and rewarding cues in a goal directed motivational context that depends upon past histories of rewards and punishments. In contrast to many personality researchers who limit themselves to single behaviors associated with single dimensions of personality, the contributors to this volume have focused on the complex interplay of three systems and two types of cues as they lead to learning and performance.

Much of this book focuses on differential responses of the three systems to cues for punishment and reward (with a particular emphasis upon the BIS and BAS), but ignores that much of behavior is not directly motivated by environmental reinforcers. Most of animal and human learning is not directly reinforced but is rather learning how to parse the continuous flow of input from the sensory system into meaning. Infants exposed to the buzzing confusion of the stream of sounds and light that adults associate with words, sentences and objects need to learn to parse the stream into words that then lead to associations with objects and action (Hespos and Spelke, 2004). For adults, much of our learning is implicit and semantic and not directly associated with reinforcement (Reber, Gitleman, Parrish, & Mesulam, 2003). For the rat, as we have known for years, exploring mazes will lead to knowing the structure of mazes, even without direct rewards. It is only when reinforcements are made explicit that the learning results in motivated action (Blodgett, 1929; Tolman, 1948; Tolman and Honzik, 1930).

Although it is possible to learn without reinforcement, that there are different brain systems to [learn how to] approach, to avoid, and to inhibit is fundamental to current personality theory. The discussions in the preceding chapters have focused on learning to approach given appetitive cues and learning to avoid given aversive cues, and the use of an inhibitory system to resolve conflict in the case when the cues themselves are conflicting. However, a broader interpretation of these three systems (Furnham and Jackson, Chapter 15) is that we can distinguish between general tendencies for action and inaction, approach, avoidance, and inhibition (Atkinson and Birch, 1970; Elliot and Thrash, 2002; Humphreys and Revelle, 1984; Kuhl and Blankenship, 1979; Revelle, 1993; Revelle and Michaels, 1976).

The authors of the preceding chapters uniformly have recognized the complex patterns of predictions that a thoughtful analysis of RST produces. The scope of the chapters is inspiring, for they range from discussions of evolutionary reasons for three systems (Âvila & Torrubia, Chapter 7), to discussions of the molecular genetics and neurophysiology of each system, (Reuter, Chapter 10) to techniques in measurement (Pickering and Smillie, Chapter 4; Torrubia, Âvila, & Caseras, Chapter 6), and the implications for psychopathology (Knyazev, Wilson, and Slobodskaya, Chapter 14; Newman and Wallace, Chapter 13; Zinbarg and Yoon, Chapter 12), health (Keltikangas-Järvinen, Chapter 11) and performance in the work place (Furnham and Jackson, Chapter 15).

18.1   RST1 and the RST2: Does it make a difference?

The basic contribution of the original theory, RST1, was not that specific neural structures related to specific behaviors, but that it was possible to integrate aspects of the Central Nervous System (CNS) into a conceptual nervous system (cns) composed of approach and inhibition systems. RST1 gave plausible (and different) neural underpinnings to the

descriptive and causal theory of Hans Eysenck (1967). RST1 and its variants became the standard model of biologically oriented personality theories. Indeed, even as the axes of RST1 slowly rotated towards Extraversion and Neuroticism, the explanation of the Eysencks E/N dimensions became more Gray like (e.g. Corr, Chapter 1; Gomez and Cooper, Chapter 9; Matthews and Gilland, 1999).

Most tests of the human personality aspects of the theory were not at the biological level but rather at the conceptual systems level. This is a point of irritation for Pickering (Chapter 16) who thinks that the human level results were not directly applicable to the theory. Tests of reward or punishment sensitivity could involve conditioning (e.g., Zinbarg and Mohlman, 1998; Zinbarg and Revelle, 1989) and classification tasks (Rogers and Revelle, 1998) as they related to self report inventories. These studies did not attempt to relate these behaviors to parts of the septal-hippocampal loop, but were expressed at the conceptual level of rewards, punishments, and positive and negative affect.

RST2 was a major revision of RST1 at both the physiological level and at the conceptual level, with much of the emphasis of aversive cue sensitivity switching towards the FFFS, but the changes in predictions were not as immediately obvious. Indeed, most references to Gray's theory of personality are still references to RST1 and are oblivious of RST2. One of the goals of this book is to bring the newer model to the broader research community. In addition to this text, the recent articles by Corr (2004) and Smillie, Pickering and Jackson (2006) do much to enlighten personality researchers not directly connected to the RST mafia.

18.2   Basic theory

The distinctions between RST1 and RST2 are laid out in the early chapters by Corr (Chapter 1), McNaugton and Corr (Chapter 2) and are formally modeled by Pickering (Chapter 16). These reviews are very helpful in highlighting the changes that RST underwent both in the biological bases of the three systems as well as the behavioral implications of the new model.

In a masterful review, Corr (Chapter 1) directly addresses the question of why use RST2 when RST1 is so well understood. Does the human theorist really need to care about the subtleties of brain function (McNaugton and Corr, Chapter 2)? Yes and No. No in terms of distinctions between the relative importance of the medial hypothalmus versus the amygdala. But yes in terms of the revised importance of the Fight/Freeze/Flight System and the modified role of the Behavioral Inhibition System. RST2 has dropped the conditioned/unconditioned aversive stimuli distinction and now emphasizes that the FFFS is involved in reactions to all aversive stimuli, the BAS in reactions to all appetitive stimuli, and the BIS in resolving goal conflicts in general. (That is, not just approach-avoidance,

but also avoidance-avoidance and approach-approach, see also Zinbarg and Yoon, Chapter 12).

In agreement with some of the human clinical literature, RST2 now distinguishes between fear and anxiety. It adds an emphasis upon both defensive distance and defensive direction. An analysis of the defensive distance for eliciting Fight, Freezing, or Flight will be an important addition to human models. In the rat, defensive distance may be directly measured, but in the human the concept is more in the cognitive representation of threat. Evaluating biases in cognitive estimates of the strength and distance of aversive cues will be a useful framework for testing the application of the model as it relates to trait and state measures of anxiety.

As a powerful demonstration of the utility of the original model, RST1, Âvila and Torrubia's review (Chapter 7) integrates RST1's biological approach with modern studies of the effect of anxiety and extraversion on condtioning. They review Newmans interpretation of Gray (Newman and Wallace, chapter 13; see also Patterson & Newman, 1993) and point out that it is the differences in the learning of aversive associations in conditions that can lead to reward and punishments that are essential to understanding disinhibition. They also examine Corr's Separable versus Joint Sub Systems distinction (SSH and JSH) and outline the conditions when SSH is true (strong appetitive/aversive stimuli, extreme scores, no mix of aversive and appetitive stimuli, no need for rapid attentional shifts). They make the interesting point that approach tendencies are continuously reviewed and can be extinguished whereas inhibitory tendencies are prone to be maintained through successful advoidance. As is seen also in many of the other chapters, Âvila and Torrubia discuss how human research is different from animals and emphaisize the role of expectations and the motivational context.

Too many theories in personality are verbal descriptions of complex patterns of affect, cognition, and behavior. RST tends to fall into this trap and has many different interpretations. This is a particular problem in the way many of us address the implications of RST. Some of these theoretical descriptions include structural models of relationships (e.g., Furnham and Jackson, Chapter 15; Humphreys and Revelle, 1984) that are admittedly unclear in the specific pattern of predictions.

Pickering (Chapter 16) addresses this problem and begins to formalize the model in terms of a static model of the pattern of predicted relationships. He recognizes the need to convert these static models into dynamic simulations and reports some progress in this endeavor. (See also Smillie, et al., 2006). He further points out that although logically the variables of interest are the sensitivity weights associating cues for reward and punishment with activation of the BAS and FFFS, that these weights are probably not open to self awareness, but that rather the affective, behavioral, and cognitive reactions produced by BAS and FFFS are and thus to self report. This highlights a major complication in terms of testing the theoretical predictions with respect to self report personality variables, for

the fundamental variables are not open to awareness.

Although RST is a theory of three systems, the emphasis has traditionally been upon the BIS and more recently the FFFS. The biological underpinnings of anxiety from the RST perspective are well reviewed by Gray and McNaughton (2000) but those of the BAS are underplayed. This lack of emphasis on the BAS is compensated for by Pickering and Smillie (Chapter 4) who do an excellent job at reviewing the physiological structures proposed to underlie the BAS.

In addition to the biological correlates, they ask whether the BAS is perhaps oversold. In 1995 I suggested that the BAS has been associated with most of the approach traits (extraversion, impulsivity, novelty seeking, and positive affectivity) as well as states of positive affect (Revelle, 1995). Pickering and Smillie (Chapter 4) ask whether this promiscuity of associations is due to the importance of the BAS, or due to theoretical difficulties in defining the concept?

The original association of BAS with impulsivity was unfortunate, and Pickering and Smillie review the biopsychological and behavioral data suggesting that dopaminergic functioning and the reward system are more aligned with extraversion than with impulsivity. Unfortunately, although extraversion can be better measured than impulsivity (e.g., Revelle, 1997) purified measures of extraversion leave something to be desired. Even correcting for attenuation does not lead to perfectly correlated measures (Goldberg et al, 2006). This is a problem not for RST but for the broader study of personality. This problem of measurement can partly be resolved by focusing on discriminant validity, particularly focusing on what a putative measure does not measure, rather than the convergent validities of our psychological constructs.

18.2.1 Biological Correlates

Reuter (Chapter 10) provides a very thoughtful analysis of current work in molecular genetics and the current work trying to associate candidate genes with candidate brain structures. This work is hindered by the problems of measurement of RST constructs at the personality item/scale level. The proposed measurement of endophenotypes of RST identified by fMRI and PET requires better identification of those endophenotypes than is currently possible. RST is a link between genetics, the proteins for which genes code, the structures and transmitters associated with those proteins, and higher level approach and inhibitition and fighting/fleeing behaviors that we see at the level of the whole organism. The multiplicity of genes associated with the dominergic system that is presumably the underpinning of the BAS makes it quite clear that the data force us to reject the One Gene-One Disease (OGOD) model, or its generalization to the One Gene-One System Hypothesis (OGOSH).

De Pascalis (Chapter 8) thoroughly reviews the problems of correlating psychophysiological measures with putative measures derived from RST. The patterns of correlations are generally supportive of BIS/BAS distinctions, but are less consistent with respect to specific localizations of function. The early work by Sutton and Davidson (1997) suggesting a lateralization of functioning for the BIS versus the BAS has been called into question by the effects of anger and approach on frontal activation (Harmon-Jones, Sigelman, Bohlig, & Harmon-Jones, 2003). DePascalis' conclusion matches that of many of the other authors in this volume, that there is a great deal of promise but a great deal that needs to be known before we can definitively assess RST1 or RST2.

Keltikangas-Järvinen (Chapter 11) reviews the physiological and health risks associated with a reactive negative system (equated with the BIS) and a reactive positive system (equated with the BAS.) The emphasis upon BIS vs. BAS leads to thinking of affect as one bipolar dimension and similarly tends to focus on BIS-BAS as a bipolar dimension. Perhaps this is a misinterpretation, for the evidence for the multidimensionality of affect is compelling (Matthews et al., 2002; Rafaeli and Revelle, 2006). Keltikangas-Järvinen recognizes the problem of measurement with BIS -BAS but do not focus so much on the complexity of a theory where the two systems interact with each other. Furthermore, by focusing on health correlates of BIS and BAS, the possibility of curvilinear, non-monotonc, associations is ignored. That is, it is important to consider the functional significance of both systems. It is likely that there are benefits and costs associated with each system and the optimal level is at some middle point (see below for a further discussion of the challenge on optimality for theory testing.)

18.2.2   Are animal models appropriate for human personality?

The RST was derived from work primarily based upon rats. (Jeffrey used to confess that he was a rat psychologist who would surface every seven years to see what had happened in human personality. Concluding that not much had happened, he would return to the rat lab.) A recurring theme throughout the book is that Gray's theory is rat based and humans are more complicated. Is a biological driven theory, derived from rats, applicable to humans? Many of the chapters address this issue, and most suggest that the answer is yes. Much of behavior is reflexive or routine and does not involve complex cognition (Ortony, Norman & Revelle, 2005). That the behavior reflects prior experience and situational context is given, but that a rat based model is adequate for developing theories for complex human feelings, thoughts, and actions may be challenged.

Matthews (Chapter 17) rises to this challenge and suggests that a cognitive science approach of computation is the most useful level of analysis. But this fails to address the important contributions of behavior genetics and physiology. Even taking a cognitive science approach, we should distinguish between the information being processed and the

computational limitations of the processor. Individual differences in the activation of the three systems lead to different information being stored in associative memories and different bias settings in the coding of environments based upon these prior memories as well as different resource limitations. None-the-less, Matthews' request for a formal computational model is appropriate and a connectionist architecture along the lines sketched out by Pickering (Chapter 16) is a beginning.

McNaughton and Corr (Chapter 3) respond to Matthews' critique in a forthright defense of the use of animal models to study human personality. To add to this debate is not necessary, but irresistable. It should be noted that almost 60 years ago Tolman (1948) discussed animal models of cognition in a computational form that would appeal to the most cognitive of psychologists. To Matthews, the contribution of cognitive science is the requirement for computational models, but to others, the great breakthrough from strict behaviorism to cognitive psychology was the introduction of latent constructs and intervening variables. The conceptual nervous system approach of Hebb, particularly as utilized by Gray, follows this approach by creating high level summaries (BAS/BIS/FFSS) and idealizations of low level wetware.

18.3   Can biological theories address cognitive problems?

That one's biology affects one's cognitions is the primary focus of those who study energetic components of mood and emotion. Changes in arousal and affect, by changing resource availability and resource allocation, affect the relative strengths and foci of competing cognitive processes (Humphreys and Revelle, 1984; Revelle, 1993). But this is a two way street, for changing cognitions will change arousal and affect. That the internal, cognitive, representations of external contingencies of reinforcement need to be considered has been clear for decades. Weiner and Schneider (1971) showed that the interaction of test anxiety with task difficulty effects so clearly demonstrated by Spence, Farber and McFann (1956), was due to subjects interpreting their performance as success or failure (and thus rewarding or punishing). When success/failure was made explicit by the experimenter, the anxiety by feedback interaction drove the results. That is, the experimental participants' cognitive interpretations of the situation overcame any potential physiological differences in anxiety related drive.

A recurring theme in cognitive models of personality and performance is that there are multiple levels of control. Low level, reactive responses to stimuli can be controlled by routines that have become automaticized and require little to no cognitive resources. Although much of human behavior probably occurs at this routine level, higher level reflective processes can (and do) override them to facilitate the approach of long term goals and to resolve conflict (Ortony et al., 2005). Similar hierarchical models of control have been

proposed by Broadbent (1971), Mulder (1986), and Revelle (1993).

But such levels of control are not unique to cognitively oriented models. RST2 specifically incorporates levels of control into the neurophysiology of anxiety with prefrontal cortically based linguistic control systems overriding and interacting with subcortical and archicortical control of anxiety (Gray and McNaughton, 2000). (See also MacLean, 1990 for an earlier neurologically based hierarchical model of control and Corr, Chapter 1 for the place of cognitive control in RST2.)

18.4 RST and psychopathology

A fundamental hypothesis to the dimensional approach to psychopathology is that psychopathology represents extreme scores on normal personality traits. Separately, a basic statistical concept is that small differences in the means or variances of two groups on an underlying dimension can lead to large differences in the relative frequencies in which these two groups are seen at extreme scores of this dimension. The combination of the continuity hypothesis with the extremity effect allows for powerful tests of theory by using psychopathological groups, for small mean differences in the groups can be detected in the relative frequency of rare events. The phenomenon of disinhibition and psychopathy is an excellent example of this approach. Psychopathy is characterized as extreme disinhibition. Although the incidence of psychopaths is less than 1% in the population, they represent up to 25% of population of prisons. Using prisoners is thus a powerful technique for studying subtle correlates of disinhibition with personality traits associated with the RST.

Wallace and Newman (Chapter 13) consider the proposed distinction between primary and secondary psychopaths in terms of the old and the new RST and find that the pattern of relationships is much more compatible with the revised model. Traditionally, primary psychopaths are characterized as having low BIS activity while secondary psychopaths have a super active BAS. Factor analyses of the Psychopathy Check List leads to a two (correlated) factor solution with Factor 1 reflecting BIS- and factor 2 BAS+. Wallace and Newman suggest that BIS/BAS moderate the expression of the general trait.

Knyazev, Wilson, & Slobodskaya (Chapter 14) take an evolutionary perspective to suggest that both an overactive BIS or BAS can lead to maladjustment and suggest that some balance in the activation level of the two systems is optimal. Based upon an analysis of Corr's Joint Subsytem Hypothesis they suggest that low levels of both might lead to more favorable adjustment. In emphasizing the role that cognitive ability plays as a control process, they report that ability moderates the negative effects of high BAS, with low ability students showing larger BAS effects. In their treatment of the BIS/BAS system as a gargantuan two they describes what is functionally Wundt's organization of Galen's four temperaments in terms of BIS/BAS balance. Although not described in so many words,

their conclusion is that the sanguine (low BIS and low BAS) is most socially adjusted and that extreme scores and particular a lack of BIS/BAS balance leads to resiliency and stability.

Zinbarg and Yoon (Chapter 12) provide a very compelling review of how RST (both old and new) fit into the clinical literature on depression and different subtypes of anxiety. They take full advantage of the experimental psychopathology literature and relate it to the structural differences discussed in Gray and McNaughton. Zinbarg and Yoon show how a hierarchical model of anxiety can be conceptualized in terms of RST. More importantly, Zinbarg and Yoon propose, based upon the experimental and correlational clinical data, theoretical challenges to RST that need to be addressed.

18.5   RST in the broader context

A potential weakness of RST is that it has failed to be seen as relevant to the broader psychological community. Furnham and Jackson (Chapter 15) provide a brief review of the lack of contact between RST and work related concepts. As they point out, this partly reflects the general lack of connection between mainstream personality researchers and organizational theorists, but also reflects the lack of an emphasis upon measurement of the RST constructs. Although personality trait approaches are making inroads into Industrial/Organizational psychology, this is either through poorly defined (but popular) measures such as the Myers-Briggs type inventory (see Pittenger, 2005 for a strong critique of this practice) or much better defined but still non-RST related inventories such as the NEO-PI-R (Costa and McCrae, 1992) or Hogan Personality Inventory (Hogan and Hogan, 1995). What is missing is a theory based set of predictions of what leads to effective performance in the work place.

Furnham and Jackson (Chapter 15) follow earlier proposals (e.g., Humphreys and Revelle, 1984) and view RST constructs as biological/temperamental distal causes that combine with those motivational and situational demand variables that have been referred to as characterological and that lead to behavior. Their model abstracts the common thread of RST and recent work on human motivation (e.g. Elliott and Church, 1997; Elliot and Thrash, 2002) with a focus on approach and inhibition of behavior. It would be useful also to generalize the model to examine how goals such as promotion of gains versus prevention of losses (Higgins et al., 2003) relate to the more distal components of the RST. Their broad approach, which integrates the biological structures of RST with cognitive representations of goals as well as past histories of experience emphasizes the strength of RST as an integrative theory.

18.6 Issues for the future

18.6.1 How to measure the basic components

A requirement for any theory is that it should specify how to be tested. To test RST requires adequate measurement of the fundamental constructs. But, as has been suggested in multiple chapters, there has been remarkably little concern as to how to measure individual differences in BIS, BAS, or FFFS. This is partly because each of these constructs is specified at the biological/conceptual nervous system level and results in a complex patterning of Affect, Behavior, and Cognition in the context of an individuals Desires (or goals). Many of the predictions of the model are of the type that (e.g.,) people who have an active Behavioral Approach System and a weak Behavioral Inhibition Sytem and a weak Fight/ Freeze/ Flight System when faced with a cue for reward and also a cue for punishment will do X. But this does not lead to an easy self report or observational item. Torrubia, Âvila, and Caseras (Chapter 6) address the measurement problem and attempt to directly develop self report scales to assess the RST. Part of the difficulty in measurement results from what Wittmann (1988) describes as Brunswickian Symmetry. Broad constructs predict at broad levels of analysis and are not nearly as effective at narrow levels. Similarly, narrow trait measures are not particularly useful for predictions of broad criteria.

The failure to line up reinforcement sensitivity scales directly with the axes of the Eysenck PEN or the Big 5/Five Factor models does not invalidate RST, but rather suggests that the fundamental axes of RST are not well understood in their identification at the trait level. This might be a result of between measuring traits using self report inventories. As Pickering proposes, people are probably not be aware of their sensitivities to cues for positive and negative reinforcement, but rather are aware of the patterning of behaviors associated with those cues. In addition, it is likely that the observed surface traits reflect complex combinations of the underlying parameters of the RST. A similar problem occurred with the original formulation of BIS/BAS theory with respect to the surface traits of extraversion and neuroticism. Extraversion was seen as the difference between BAS and BIS, while neuroticism was seen as their sum.

In an attempt to relate biological constructs to surface traits, Gray mapped the BAS and BIS onto impulsivity and anxiety. Given the importance of the Eysenck (1967) two space of E and N, and the original model of anxiety as reflecting the BIS, this was a rational choice at the time as Gray was attempting to integrate BIS/BAS within the Eysenckian framework while simultaneously explaining anxiety. The identification of the BAS with impulsivity was an unfortunate choice partly because of the multidimensional structure of impulsivity. Just as the meaning of Extraversion has changed over the years, so has the meaning and measurement of impulsivity. Impulsivity is probably a mix of sensitivity to reward and a lack of inhibitory control.

Just as the measurement of sensitivity to aversive events and tendencies towards behavioral inhibition can be shown to have a hierarchical structure (Zinbarg and Yoon, Chapter 12), so part of the confusion on the measurement of BAS (Pickering and Smillie, Chapter 4) probably reflects a hierarchical structure of approach like traits. Hierarchical factor models are typical in the measurement of intelligence, where lower level ability factors are found to be correlated and a higher order factor (g) can be extracted (Carroll, 1993) and loadings on g and orthogonal lower order factors can be found using the Schmid Leiman transformaton (Jensen and Weng, 1994; Schmid & Leiman, 1957). Unfortunately, with the exception of Zinbarg and his colleagues (Zinbarg et al., 2005, 2006) such hierarchical models are not common in the non-cognitive domains of personality. The BAS like traits are a probable candidate for showing a general factor running through all of the measures with specific subdomains represented in each test.

At the broadest level, a general factor common to the entire approach domain can be taken as representing BASg, but each particular test is only partially saturated with BASg, and has specific components. A classic example of this are the BAS scales from Carver and Whites BIS/BAS scales (Carver and White, 1994). The BAS scale breaks down into 3 correlated subfactors, the common component of which could be seen as BASg. This point is reminiscent of the debate between Eysenck (1967) and Guilford (1975) as to whether there is a common factor to sociability and impulsivity (i..e., Extraversion), with Eysenck arguing for a common higher order factor, and Guilford criticizing Extraversion as shotgun marriage of two unrelated traits.

18.6.2 Optimal level of activation of all three systems, separately and combined

Activation of any one system is going to have costs and benefits. But all three systems are working in unison and the problem becomes one of finding the optimal value for all three systems in a changing world. Are the costs of one system the benefits of the others? Consider this analysis for the Behavioral Approach/ Activation System. Low levels of BAS, and low responsiveness to cues for reward, will lead to sluggish behavior and a low frequency of reward. Increasing the BAS will lead to an increasing frequency of reward. But activation bears an energetic cost and risk from predation or other threats. Even more activation leads to more energy expenditure (and greater risk). If the benefits saturate and the costs grow, there will be an optimal level of BAS activity. An unfettered BAS is manic, not adaptive. More moderate levels of BAS will avoid both sluggishness and mania

Similarly, low activation of the FFFS will make one a convenient target for predators and lead to inadequate avoidance (by either fighting or fleeing). But too high a setting will lead to constant avoidance and an inability to ever engage in a task. In that the BIS resolves conflicts between BIS and FFFS through inhibition of both, once again, a moderate level

leads to more systematic and sensitive control.

The challenge of optimality is that it implies that correlations with effective functioning will be curvilinear rather than monotonic. The "Goldilocks principle" of not too little, not too much likely applies to the three systems, separately as well as in combination.

An important question from an evolutionary point of view is how is diversity maintained. If the functioning of a system uniformly increases survival and offers reproductive advantage, genetic variation in the function of that system will be reduced over time. If, however, there are disadvantages for very high or very low functioning of a system, e.g., of the BIS or of the BAS then selection can lead to a balanced polymorphism. Although most researchers examine the characteristics of high BIS and high BAS, Âvila makes the interesting suggestion that those with an less active BIS are able to learn a broader repertoire of approach tasks and have a lower resistance to extinction than those with a more active BIS, and that an active BAS leads to more difficulty extinguishing a dominant response for reward than a less active BAS.

RST1 and its variants became the standard model for the biologically oriented and RST2 should as well. However, RST1 and RST2 propose different mechanisms. At the basic structural level, these are very different. At the affective, behavioral, and cognitive they are also different, but in more subtle ways. Teasing out the implications of the RST2 (3ST) will be a research agenda for years to come.

Unresolved problems: Is the BIS a universal conflict resolver? Corr (chapter 1) and Zinbarg and Yoon (chapter 12) raise the questions about approach-approach and avoidance-avoidance conflicts. Impulsive behavior is not just not inhibiting actions that lead to aversive outcomes, it is also being distracted from doing one positive action by some other cues for some other positive action. (Writing a chapter in honor of Jeffrey Gray is a positive behavior, but impulsively checking just one more reference on the web is another positive act that needs to be inhibited.)

What is rewarding and what is aversive? There are some prepared stimuli that are threatening (for primates, snakes and spiders seem to be prepared, guns and electric outlets are not) and rewarding (infant faces, be they humans or dogs elicit positive affect), while there are many stimuli that come to be seen as threatening or rewarding after many trials. A sailboard is not automatically a cue for either reward or a threat of danger, but is a cue for reward for the experienced windsurfer, a cue for danger from a novice who has fallen into cold water while trying to stand up on a board, and a flat plank of fiberglass to someone who has never seen the water. But where do these expectations come from? That some people find windsurfing or sailing small boats in high winds to be rewarding

15

while others find it aversive presumably reflects years of experience on the water, with or without traumatic events.

What is the role of cognition?   There is a clear role of cognition as a high level control process in the manifestations of BIS/BAS/FFFS interactions. Cognitive control processes presumably are used to inhibit behavior without involving BIS activation, but these control processes can be hindered by fatigue or distraction. People learn to inhibit exhibiting stereotypic thinking but this takes cognitive and energetic resources. When evening people are asked about minority groups in the morning, they are more likely to stereotype than they are later in the day (Bodenhausen, 1990). (That is not just a direct time of day effect is shown when morning people stereotype more later in the day.)

18.6.3   Testability: How could we know we were wrong?

Matthews (Chapter 17) raises the important question that a theory needs to be open to the possibility of disproof. That RST1 has morphed into RST2 is a sign of the testability (and rejection) of the original model. Unfortunately, the formal models of Pickering and his colleagues (Chapter 16 and Smillie, et al., 2006) suggest that RST2 makes more complicated predictions than RST1 and that these predictions might be harder to reject. For the very complexity of the model allows it to make many, seemingly contradictory predictions. However, we can hope that as the formal modelers are forced to specify particular ranges of parameter values, that testable prediction will result.

The application of 3ST to complex behaviors in complex organizations (Furnham & Jackson, Chapter 15) suggests the possibility of integration with other motivational models (e.g. achievement motivation theory as developed by Atkinson and his colleagues has seen a resurgence of interest with models of Elliot and his colleagues and is very compatible with 3ST).

RST1 and RST2 are conceptual frameworks thought to organize human and rat brain structure data. But RST1 and RST2 have been applied by some at a level that is untouched by biological theorizing. The BAS/FFFS/BIS theory at this higher level uses the terms of the biological systems but does not stand or fall on changes in our understanding of biology. This is both a strength and a weakness of the theory. Gray and his closest collaborators thought of RST2 as a biological model that could be tested at the biological level. As formulated at a generalized level, in terms of making predictions about job performance and psychopathology, the real contribution was a focus on separating a BAS from a FFFS and BIS.

18.7   Conclusion

This book is a small tribute to the contribution of Jeffrey Gray. Wrapped up as he was in his broad ranging studies of how the brain works he perhaps did not realize the enormous contribution he had made to the study of personality. His careful and analytical approach, combined with an openness to the ideas of others were an inspiration to far more people than just those who collaborated in this book. The impact of his work upon the field of personality in particular and psychology in general will continue to influence the research agenda in personality theory for the coming decades.

18.8   References

Atkinson, J. W. and D. Birch (1970). The dynamics of action. New York, Wiley.

Blodgett, H.C. 1929. The effect of the introduction of reward upon the maze performance of rats. Univ. Cal. Pub. Psychol. 4: 113-134.

Bodenhausen, Galen V. (1990) Stereotypes as judgmental heuristics: evidence of circadian variations in discrimination. Psychological Science, 1 319-322.

Broadbent, D. E. (1971). Decision and stress. London, Academic Press.

Carroll, J.B. (1993). Human cognitive abilities: A survey of factor-analytical studies. New York : Cambridge University Press.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. Journal of Personality and Social Psychology, 67, 319-333.

Corr, P.J. (2002). J.A. Gray's reinforcement sensitivity theory: tests of the joint subsystem hypothesis of anxiety and impulsivity. Personality and Individual Differences, 33, 511-532.

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. Neuroscience and Biobehavioral Reviews, 28, 317-332.

Costa P.T., Jr., and McCrae R.R.(1992) Revised NEO Personality Inventory (NEO PI-R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Depue, R. A. and P. F. Collins (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. Behavioral & Brain Sciences 22(3): 491-569.

Elliot, A. J. and M. A. Church (1997) A hierarchical model of approach and avoidance achievement motivation. Journal of Personality and Social Psychology, 72, 218-232.

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

Elliott, E.S., & Dweck, C.S. (1988) Goals: An approach to motivation and achievement. Journal of Personality and Social Psychology, 54, 5-12.

Eysenck, H. J. (1967). The biological basis of personality. Springfield, Thomas.

Fowles, D. C. (1980). The three arousal model: Implications of Gray's two factor learning theory for heart rate, electrodermal activity and psychopathy. Psychophysiology 17: 87-104.

Fowles, D. C. (1984). Arousal: implications of behavioral theories of motivation. Psychophysiological perspectives: Festschrift for Beatrice and John Lacey. M. G. H. Coles, J. R. Jennings and J. P. Stern. New York, Van Nostrand: 143-156.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. Journal of Research in Personality, 40, 84-96.

Gray, J. A. (1970). The psychophysiological basis of introversionextraversion. Behaiour Research and Therapy, 8, 249266.

Gray, J. A. (1982). The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System. Oxford: Oxford University Press.

Gray, J.A., & McNaughton, N. (2000). The Neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system. Oxford: Oxford University Press.

Guilford, J. P. Factors and factors of personality. Psychological Bulletin, 1975, 82, 802-814.

Harmon-Jones, E., Sigelman, J. D., Bohlig, A., & Harmon-Jones, C. (2003). Anger, coping, and frontal cortical activity: The effect of coping potential on anger-induced left frontal activity. Cognition and Emotion, 17, 1-24.

Hespos, S. J. & Spelke, E. S. (2004). Conceptual precursors to spatial language. Nature, 430, 453 - 456.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., & Molden, D. C. (2003). Transfer of value from fit. Journal of Personality and Social Psychology, 84, 1140-1153.

Hogan, R., & Hogan, J. (1995). The Hogan Personality Inventory manual (2nd ed.). Tulsa, OK: Hogan Assessment Systems.

Hogan, R., Hogan J. & B.W. Roberts ((1996) Personality measurement and employment decisions: Questions and answers. American Psychologist, 51, 469-477.

Hogan R, Kaiser RB. (2005). What we know about leadership. Review of General Psychology , 9(2), 169180.

Hogan, R. & B. W. Roberts (2004) A Socioanalytic Model of Maturity. Journal of Career Assessment, Vol. 12. 207-217 (2004)

Humphreys, M. S. and W. Revelle (1984). "Personality, motivation, and performance: A theory of the relationship between individual differences and information processing."

Psychological Review 91(2): 153-184.

Jensen, A. R., & Weng, L.-J. (1994). What is good g? Intelligence, 18, 231-258.

Kuhl, J. and V. Blankenship (1979). The dynamic theory of achievement motivation: From episodic to dynamic thinking. Psychological Review 85: 239-248.

MacLean, P. D. (1973). A triune concept of brain and behavior. Toronto: University of Toronto Press.

MacLean, P.D., 1990: The Triune Brain in Evolution: Role in Paleocerebral Functions, Plenum, New York.

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

Matthews, G., Campbell, S.E.. ,Falconer, S ; Joyner, L A. ; Huggins, J., Gilliland, K. , Grier, R., Warm, J. S. (2002), Fundamental dimensions of subjective state in performance settings : Task engagement, distress, and worry. Emotion, 4, 315-340

McGue, M. & Bouchard, T. J. Jr. (1998). Genetic and environmental influences on human behavioral differences. Annual Review of Neuroscience, 21, 1-24.

Mulder, G. (1986). The concept and measurement of mental effort. In G. R. J. Hockey,A. W. K. Gaillard, & M. G. H. Coles (Eds.), Energetics and human information processing (pp. 175-198). Dordrecht: Matinus Nijhoff

Ortony, A., Norman, D.A. & Revelle, W. (2005): Effective Functioning: A Three Level Model of Affect, Motivation, Cognition, and Behavior. in J. M. Fellous & M. A. Arbib (Eds.), Who Needs Emotions? The Brain Meets the Machine. New York: Oxford University Press (pp: 173-202)

Patterson, C.M., & Newman, J.P. (1993). Reflectivity and learning from aversive events: toward a psychological mechanism for Syndromes of Disinhibition. Psychological Review, 100, 716-736.

Pickering, A.D., & Gray, J.A. (2001). Dopamine, appetitive reinforcement, and the neuropsychology of human learning: An individual differences approach. In A. Eliasz & A. Angleitner (Eds.), Advances in individual differences research (pp. 113-149). Lengerich, Germany: PABST Science Publishers.

Pittenger, D. J. (2005). Cautionary Comments Regarding the Myers-Briggs Type Indicator. Consulting Psychology Journal: Practice and Research, 57, 210-221

Rafaeli, E. & Revelle, W. (2006). A premature consensus: Are happiness and sadness truly opposite affects? Motivation and Emotion, 30, 1-12.

Reber, P.J, Gitelman, D.R., Parrish, T.B. & Mesulam, M.-M. (2003). Dissociating explicit and implicit category knowledge with fMRI. Journal of Cognitive Neuroscience, 15, 574-685.

Revelle, W. (1993). Individual differences in personality and motivation: 'Non-cognitive' determinants of cognitive performance. Attention: selection, awareness and control: a tribute to Donald Broadbent. A. Baddeley and L. Weiskrantz. Oxford, Oxford University Press.: 346-373.

Revelle, W. (1995). Personality Processes. Annual Review of Psychology, 46, 295-328.

Revelle, W. (1997) Extraversion and Impulsivity: The lost dimension. In Nyborg, H. (Ed.) The scientific study of human nature: Tribute to Hans J. Eysenck at eighty. Elsevier Science Press.(189-212)

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

Rogers, G. M. and W. Revelle (1998). Personality, mood, and the evaluation of affective and neutral word pairs. Journal of Personality & Social Psychology 74: 1592-1605.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. Psychometrika, 22, 53-61.

Smillie, L. D. & Jackson, C. J (2006). Functional Impulsivity and Reinforcement Sensitivity Theory. Journal of Personality, 74, 47-83.

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, 320-335

Spence, K. W., Farber, I. E., & McFann, H. H. (1956). The relation of anxiety (drive) level to performance in competitional and non-competitional paired-associates learning. Journal of Experimental Psychology, 52, 296-305.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. Psychological Science, 8, 204-210.

Tolman, E. C. (1948) Cognitive Maps in Rats and Men. Psychological Review, 55. 189-208.

Tolman, E.C. and Honzik, C.H. 1930. Introduction and removal of reward and maze performance in rats. Uni. Cal. Pub. Psychol. 4: 257275.

Weiner, B. and K. Schneider (1971). Drive versus cognitive theory: a reply to Boor and Harmon. Journal of Personality and Social Psychology, 18: 258-262.

Wittmann, W. W. (1988). Multivariate reliability theory. Principles of symmetry and successful validation strategies. In R. B. Cattell & J. R. Nesselroade (Eds.), Handbook of multivariate experimental psychology (pp. 505-560). New York: Plenum.

Zinbarg RE, & J. Mohlman (1998). Individual differences in the acquisition of affectively valenced associations. Journal of Personality and Social Psychology, 74, 1024-1040

Zinbarg, R. and W. Revelle (1989). Personality and conditioning: A test of four models. Journal of Personality & Social Psychology 57: 301-314.

Zinbarg, R.E., Revelle, W., Yovel, I., & Li. W. (2005). Cronbach's Alpha, Revelle's Beta, McDonald's Omega: Their relations with each and two alternative conceptualizations of reliability. Psychometrika. 70, 123-133.

Zinbarg, R., Yovel, I., Revelle, W. & McDonald, R. (2006). Estimating generalizability to a universe of indicators that all have one attribute in common: A comparison of estimators for omega. Applied Psychological Measurement, 30, 121-144.