Temperament, reward and punishment sensitivity, and clinical disorders:
Implications for behavioral case formulation and therapy

Richard F. Farmer
University of Canterbury

Abstract
Recent research in psychology, psychiatry and neuroscience has demonstrated reliable associations between temperament and individual differences in sensitivity and responsiveness to environmental cues and behavioral consequences. Temperament–influenced behavior patterns evident in infancy have also been found to predict behavioral tendencies in adulthood. Such observations suggest that neurophysiological structures and physiological events associated with temperament concepts exert a mediating or moderating influence between current environmental events and behavior. This paper summarizes relevant research on individual differences in sensitivity and responsiveness to environmental cues and behavioral consequences with reference to Jeffrey Gray’s neuropsychological theory of temperament. Implications of temperament research for behavioral case formulation and therapy are described.

Key words. Temperament, reward sensitivity, punishment sensitivity, behavioral case formulation, behavior therapy

The concept of temperament is emerging as a potentially important consideration in behavioral case formulation and therapy. There are at least three reasons why temperament may be important: (a) in a variety of studies temperament constructs have been linked to individual differences in reward and punishment sensitivity (e.g., Gray & McNaughton, 2000), (b) individual differences in temperament have been postulated to affect environments in which they are expressed, and that such environments, in turn, reciprocally affect future displays of temperamentally–influenced behavior (e.g., Rothbart, Ahadi, & Evans, 2000), and (c) temperament has emerged as an influential factor in the form or expression of several varieties of psychopathology (e.g., Claridge & Davis, 2003).

A central thesis of this article is that behavioral assessors and therapists can potentially benefit from a consideration of clients’ temperament features, as underlying processes associated with temperament (e.g., the capacity to be sensitive and reactive to reinforcers and punishers) likely mediate or moderate current environment–behavior relations. Similarly, knowledge about how temperament relates to certain behavioral disorders can suggest hypotheses about maintaining factors as well as intervention strategies likely to exert the greatest therapeutic effect.

Form and Function of Behavior in Behavior Theory and Therapy

The functional analytic approach is a distinguishing feature of behavior therapy (Ferster, 1973; Kanfer & Saslow, 1969). Influenced by Darwinian principles of adaptability, functional analysis involves the identification of functional relationships between clinically relevant behaviors and the environmental variables that select, influence, and maintain them (Dougher & Hayes, 2000; Farmer & Nelson-Gray, 2005). Behavior therapy, therefore, has traditionally been concerned about the whole person and functional aspects of an individual’s behavior in particular environmental contexts (Follette, Naugle, & Linnerooth, 2000), and perhaps less concerned about the form or topography of behavior among populations or subgroups of persons. That is, whereas topographical description and classification methods such as the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV–TR; American Psychiatric Association, 2000) are primarily concerned with how people behave, behavioral functional analyses are primarily concerned about a behavior’s purpose with reference to the consequences it produces, or why people behave as they do (Farmer & Nelson-Gray, 2005; Nelson & Hayes, 1986; Nelson-Gray & Farmer, 1999).
Traditional topographical and functional forms of assessment and client conceptualization are not necessarily in conflict with each other, however (Farmer & Nelson-Gray, 2005; Nelson-Gray & Farmer, 1999). Behavior theorists have long acknowledged that some forms of behavior are, in part, influenced by heritable biological endowments that have been inter-generationally transmitted via phylogenetic selection processes. Skinner (1971, 1981), for example, suggested that an organism’s susceptibility to response consequences is inherited and a result of natural selection processes. Several forms of behavior, or behavior topographies, that demonstrate reliable patterns of covariation across persons likely reflect, in part, the influence of phylogenetic selection contingencies associated with fitness or survival over successive generations. These heritable individual differences, in turn, might also manifest themselves as variability in sensitivity to the effects of rewarding or punishing consequences, and subsumed under the broader concept of temperament (Gray, 1987; Gray & McNaughton, 2000).

Temperament

Derryberry and Rothbart (1984, p. 132) have defined temperament as “constitutional differences in reactivity and self-regulation.” They go on to clarify and expand their definition, “with 'constitutional' referring to the relatively enduring biological makeup of the individual, influenced over time by heredity, maturation, and experience...By reactivity we mean the functional state of the somatic, endocrine, autonomic, and central nervous systems as reflected in the response parameters of threshold, latency, intensity, rising time, and recovery time. By self-regulation we mean higher level processes functioning to modulate (enhance or inhibit) the reactive state of these systems...Self-regulatory processes are best approached in terms of emotions or affective-motivational processes” (italics added). Apparent in this definition is the notion that temperament is heritable and a temporally stable feature of persons, but potentially modifiable through basic maturational processes and by environmental learning and experience. This definition also stresses two other central features of temperament: differences among persons in the reactivity of various biological systems and in the emotional/motivational processes involved in the regulation of these activated states. The present article places emphasis on these last two features of temperament, and how consideration of such factors can have utility for behavioral assessment and therapy.

Claridge and Davis (2003, pp. 38-40; see also Cloninger, Svrakic, & Przybeck, 1993) have provided a useful distinction between temperament and personality (or character). Similar to Derryberry and Rothbart (1984), they regard temperament as having a strong heritable component and directly reflective of individual differences in central nervous system (CNS) functioning which, in turn, is primarily manifested as basic emotional reactions that are first apparent as normal variations in infancy. Variations in temperament demonstrate some degree of cross-species similarity, as evident in Pavlov’s observations of temperamental differences in his dogs that, in turn, were associated with individual differences in the acquisition and extinction of conditioned reflexes (see also Gosling, 2001, and Farmer & Nelson-Gray, 2005).

From a behavioral perspective, personality can be regarded as a complex constellation of learned behaviors that has, at best, an indirect association to CNS functioning (Claridge & Davis, 2003). Learned behavior patterns and repertoires that become cross-situationally consistent and temporally stable (i.e., personality) first begin to emerge from childhood, and are subject to ongoing shaping from family and social environments, cultural experiences, and other instrumental outcomes that behavior occasions. Whereas temporally stable and cross-situational manifestations of behavior consistent with the labels “anxiety” or “impulsivity”, for example, might be regarded as examples of the influence of temperament, behavioral prototypes that are collectively consistent with concepts such as “honest” or “mature” might be thought of as examples of personality. Personality, although influenced by temperament, is largely shaped and maintained by ontogenetic learning processes, whereas brain processes, functions, and structures
associated with the concept of temperament are in large part inherited, and ultimately the product of phylogenetic selection processes (Farmer & Nelson-Gray, 2005).

Theories that incorporate temperament/emotional constructs tend to emphasize responses to incentives (e.g., rewards) and disincentives (punishments) in immediate or ongoing contexts (Nigg, 2001). Individual differences in sensitivity and/or responsiveness to environmental cues are thought to reflect the functioning of predominantly subcortical systems. Among persons with extreme or deviant behavioral patterns or certain clinical conditions, abnormalities might exist in responsiveness to environmental cues (e.g., signals associated with unconditioned or conditioned reward or punishment) or events that follow behavior (e.g., hypo-responsiveness to rewards or punishers). These abnormalities, in turn, may contribute to a wide variety of outcomes, such as learning deficits, excessive behavioral inhibition, impulsivity, socially inappropriate actions, and poor self-regulation.

Although proposals on the exact number and nature of basic temperament dimensions have varied across theorists and researchers (e.g., Cloninger, 1987, Gray, 1987, Rothbart et al., 2000), there is general consensus that among the basic temperament types are: (a) self-controlled, well–adjusted, or “easy”, (b) impulsive or under-controlled, and (c) anxious, inhibited, or shy. In longitudinal studies, temperament patterns displayed by infants have demonstrated strong associations with adult personality patterns (Caspi, 2000). As will be discussed later in this article, anxiety/behavioral inhibition/shyness and impulsivity/under-controlled/novelty seeking temperament dimensions have both descriptive and theoretical relevance to a host of clinical disorders. Prior to turning attention to that topic, the issue of reward and punishment sensitivity in relation to temperament is first discussed.

Jeffrey Gray’s Theoretical Model of Temperament and Reward and Punishment Sensitivity

Jeffrey Gray’s comprehensive and integrative theory of temperament, neurobiology, neurochemistry, reward and punishment sensitivity, and psychopathology has been highly influential within the field of psychology. While other influential motivational-emotional-temperament theories have been proposed and researched (Nigg, 2000), such as Robert Cloninger’s psychobiological model (1987, 2003; Cloninger et al., 1993), probably no other has received as much empirical scrutiny and multiple levels of support as Gray’s theoretical model (Corr, Pickering, & Gray, 1997; Gray & McNaughton, 2000; Pickering et al., 1997; Pickering, Díaz, & Gray, 1995; Pickering & Gray, 1999; Rothbart et al., 2000). Furthermore, many of the alternative theories of temperament often share several features in common with Gray’s theory.

Gray (1970, 1973, 1981, 1987, Gray & McNaughton, 2000) advanced a modification of Eysenck’s (1957, 1967) two-dimensional theoretical model that emphasized the dimensions of anxiety and impulsivity rather than introversion-extraversion as indicators of differences in sensitivity and responsiveness to various environmental cues. In early drug research with animals, Gray (1970) observed that some drugs had the effect of reducing the sensitivity or responsiveness of an endogenous punishment mechanism, later termed the Behavioral Inhibition System (BIS; Gray, 1981). Gray and colleagues have since hypothesized that the activities of the BIS function to (a) inhibit behavior in situations where cues associated with punishment are present, (b) increase arousal to energize subsequent behavior, and (c) increase attentional resources to initially threatening novel stimuli (Corr et al., 1997; Gray & McNaughton, 2000). High state and trait anxiety are associated with the activation of the BIS (a conceptual nervous system concept that largely reflects the activities of the septo-hippocampal system), with those most sensitive or reactive to the effects of punishment or frustrating non-reward (i.e., the non-occurrence of an expected reward) being highest on the anxiety dimension. Conversely, those least sensitive or reactive to punishment or non-reward would be those who are lowest on the state and trait anxiety dimensions.
Gray (1973) proposed a second temperament dimension, impulsivity, which he believed to be uncorrelated with the dimension of anxiety. Gray (1973) proposed that as one's level of impulsivity increases, one becomes increasingly more susceptible to the effects of reward and relieving non-punishment (or the non-occurrence of an anticipated punishment, where the contingent termination or omission of a stimulus results in an increase in the future probability of a response). Those lowest on the impulsivity dimension would be the least sensitive or reactive to the effects of reward and non-punishment. The function of this system, later termed the Behavioral Approach (or Activation) System (BAS; Fowles, 1980; Gray, 1987), is to activate behavior or facilitate approach in situations where cues associated with reward or non-punishment are present. The underlying neurobiology associated with the BAS is found in the mesolimbic and mesocortical dopaminergic pathways.

Gray (1987; Gray & McNaughton, 2000) has also proposed the fight-flight-freezing system (FFFS) that mediates responses to unconditioned aversive stimuli (e.g., pain), innate fear stimuli (e.g., snakes, spiders, the dark), unconditioned frustrative nonreward, and conditioned aversive stimuli. In Gray’s model, emotional states that most commonly arise from this system are rage and panic, the latter of which is distinguished from anxiety. Behavioral tendencies most strongly associated with the activation of this system are active avoidance (escape) and defensive aggression in response to stimuli associated with threat or danger. Associated neurobiological structures and functions associated with the FFFS are the periaqueductal grey, medial hypothalamus, and amygdala, and serotonergic inputs into these structures.

In the most recent elaboration and revision of his model, Gray (Gray & McNaughton, 2000, pp. 83-87) has suggested that the BIS is most strongly activated when both the BAS and the FFFS are concurrently activated. The simultaneous activation of the BAS and FFFS result in an approach–avoidance conflict, whereby approach and avoidance tendencies are both inhibited by the BIS. The activation of the BIS, in turn, results in increased arousal, behavioral inhibition, environmental scanning to evaluate risk, threat, or danger, as well as an internal scan of memory that is geared toward the detection of information linked with negative affect and that is contextually related to threat and danger information tied to the current conflict. As such, the activities of the BIS often function to resolve the approach–avoidance tension by tipping the balance in the direction of avoidance.

In Gray’s model, various expressions of psychopathology are thought to arise from the activities of one or more of these systems. Most obvious are anxiety disorders associated with FFFS and/or BIS, and impulse-control disorders associated with BAS. Additive combinations of these systems may also be related to other expressions of psychopathology, such as hypomania (weak-FFFS/strong-BAS/weak-BIS), borderline personality disorder (strong-FFFS/strong-BAS/strong-BIS), and conditions characterized by high rates of anger and aggression, such as psychopathy and paranoid personality disorder (strong-FFFS/strong-BAS/weak-BIS) (Farmer, Nash & Dance, 2004; Farmer & Nelson-Gray, 1995; Fowles, 2001).

Temperament, Reward and Punishment Sensitivity, and Clinical Disorders

With regard to clinical disorders, Gray’s temperament model perhaps has the greatest relevance for at least three broad functional response classes: (a) behavioral excesses, (b) behavioral deficits, and (c) inhibited and avoidant behavior patterns.

Behavioral excesses may take on a variety of forms and, in the course of clinical assessment, it would be important to distinguish the various forms of behavioral excess according to the functions that they serve. Often behaviors are displayed excessively because of their associated reinforcement function. For example, many forms of behavioral excess among children have attention as a maintaining function (Scotti, Morris, McNeil, & Hawkins, 1996). Other forms of behavioral excesses that likely have a strong associated reinforcement function include pathological forms of gambling and promiscuous sex. As will
be discussed in greater detail below, conditions characterized by behavioral excess might be more likely to occur among individuals who are hypersensitive and responsive to the effects of reward or, in Gray’s model, who are BAS-dominant.

Behavioral deficits may be evident when individuals do not display an adequate range or repertoire of behavior for functioning effectively in a variety of contexts. There are many possible explanations for why such deficits might arise. For example, impoverished environments might not provide opportunities for the effective modeling and shaping of behavioral repertoires. Similarly, an environment might be overly punitive, and might punish the emergence of new behavioral repertoires, thus accounting for their low rate occurrence. Alternatively, the environment may not reinforce and positively support the emergence of new behaviors. All of these possibilities might account for absent or under-developed behavioral repertoires in particular areas. In the context of Gray’s model, one might also suggest that individuals who are constitutionally overly sensitive and reactive to the effects of punishment and extinction contingencies, who actively avoid novel contexts, who are overly reactive to innate fear stimuli, and/or who are somewhat less sensitive and reactive to the effects of innate and conditioned reinforcers might be among those most likely to display marked and generalized behavioral deficits. In relation to Gray’s tri-dimensional systems model, such individuals would likely be strong-FFFS/strong-BIS/weak-BAS.

Inhibited and avoidant behavior patterns may take on many different forms. Behavioral inhibition can be evident through “freezing” responses, through active avoidance of fear stimuli (as exemplified by active escape behavior), and through passive avoidance (e.g., as indicated by an unwillingness to enter certain environmental contexts). Of these, freezing and active escape responses are perhaps most specific to immediate environmental contexts (i.e., a response following detection of a perceived immanent threat or danger), and likely reflect FFFS activation or hyperactivity. More passive forms of avoidance, however, are likely mediated by BIS involvement, and are common concomitants to conditions such as social anxiety disorder or generalized anxiety disorder (GAD). Some theorists and therapists conceptualize worry, a central feature of GAD, as a negatively reinforced form of cognitive avoidance (e.g., Borkovec & Sharpless, 2004). Similarly, individuals with social anxiety disorder, by definition, frequently decline to enter into feared social or performance situations.

Hayes, Wilson, Gifford, Follette, & Strosahl (1996) have proposed another functional response class, termed experiential avoidance, that is evident when “a person is unwilling to remain in contact with particular private experiences (e.g., bodily sensations, emotions, thoughts, memories, behavioral predispositions) and takes steps to alter the form or frequency of these events or contexts that occasion them” (p. 1154). The individual who avoids often regards avoided private experiences as bad, threatening, or dangerous. Experiential avoidance processes are thought to have a maintaining function in many forms or expressions of psychopathology (Hayes et al., 1996).

Motivational/emotional processes involved in specific forms of problematic behavior have also been researched. For purposes of illustration, selective reviews of research on reward and punishment sensitivity in relation to attention deficit hyperactivity disorder, eating disorders, substance abuse disorders, and anxiety disorders are provided below.

**Attention Deficit Hyperactivity Disorders (ADHD)**

Among some developmental theorists, a distinction is made among mechanisms of behavioral inhibition. Derryberry and Rothbart (1997) and Nigg (2001), for example, distinguish between inhibition associated with executive functions (i.e., the activities of the frontal lobes) and inhibition associated with emotional/motivational processes (e.g., fear, anxiety, punishment sensitivity). The following discussion of ADHD focuses primarily on research from the latter perspective.
Emotional, temperament, and motivational theories applied to the study of ADHD are generally of two types: those that emphasize impairments in inhibition processes (e.g., BIS functioning) and those that emphasize impulsivity (e.g., BAS functioning). Early theories of ADHD placed emphasis on sensitivity and responsiveness to immediate contextual cues associated with punishment or threat as indicators of behavioral inhibition capacities (Nigg, 2001). The impaired inhibition theory of ADHD is best exemplified in the work of Quay (1988a, 1997). Basing his theory solidly within Gray’s framework, Quay proposed that weak or deficit BIS activation is central to ADHD. Low BIS reactivity or activation would impair the ability to interrupt ongoing activity and to detect and effectively respond to stimuli that signal the potential for punishment.

Although there is some evidence in support of the weak-BIS hypothesis (Beauchaine, Katkin, Strassberg, & Snarr, 2001; Quay, 1997), including psychophysiological data that suggests reduced skin conductance among youth with ADHD (Lazzaro et al., 1999) and patterns of abnormal skin conductance following the removal of reward (Iaboni, Douglas, & Ditto, 1997), there are some indications that the weak-BIS hypothesis has limited value as a theory of ADHD. For example, anxiety disorders are observed to frequently co-occur with ADHD (Biederman, Newcorn, & Sprich, 1991), with the presence of an anxiety disorder an almost certain indicator of BIS and/or FFFS over-activity. Similarly, laboratory-based research has produced findings that are inconsistent the weak-BIS model (Hartung, Milich, Lynam, & Martin 2002; Iaboni, Douglas, & Baker, 1995).

Even though empirical support for the weak-BIS hypothesis of ADHD is sparse, there is suggestion that inhibitory control of motor activity demonstrates the greatest impairment among youth with ADHD without comorbid anxiety disorders compared to youth with ADHD and co-occurring anxiety disorders (Nigg, 2001). Such findings suggest that the presence of anxiety might be associated with better self-control and behavioral regulation, perhaps as a result of inhibitory processes associated with BIS activation.

An alternative to the weak-BIS model is the strong-BAS model (or increased reward sensitivity) as espoused by Douglas (1989; Douglas & Parry, 1994) and others. Support for the strong-BAS/enhanced reward sensitivity model is indicated in laboratory based research (e.g., Iaboni et al., 1995; Tripp & Alsop. 1999, 2001), neurobehavioral research (e.g., Sonuga-Barke, 2002), and from research on the effects of stimulant medication which, among other functions, appears to attenuate responses to reward cues or reward sensitivity (Arnett, Fischer, & Newby, 1996; Taylor & Jentsch, 2001).

Part of the difficulty, however, in evaluating the utility of motivational/emotional models of ADHD is that ADHD itself is a heterogeneous condition that commonly co-occurs with other conditions. Research on the cognitive, emotional, and behavioral manifestations of ADHD have, for example, suggested two independent dimensions of disorder symptomatology, inattention and hyperactivity–impulsivity. The same persons sometimes display symptoms associated with these two dimensions, with these individuals referred to as “combined types” (Lahey et al., 1994). As such, individuals diagnosed with ADHD might actually constitute members of different groups that, although described with a common diagnostic label, may differ in important respects. Relatedly, conduct disorder (CD) is quite common among those with ADHD, with some studies suggesting that 30 to 50% of children with ADHD also have co-occurring CD (Biederman et al., 1991) and that upwards of 70% of children with CD have ADHD (Klein et al., 1991). Some theories of ADHD and CD distinguish these two disorder concepts based on reduced BIS activation and enhanced BAS activation, respectively (Quay, 1988a, 1988b). Anxiety disorders, as previously acknowledged, are also common among a subset of individuals with ADHD (Biederman et al., 1991).
**Eating Disorders**

A variety of theoretical models and observational findings emphasize the functional nature of disordered eating behavior. Escape and avoidance models have become increasingly influential in the past several years. These models have placed different emphases on what is avoided, as some have emphasized temporary relief from aversive self-awareness (Heatherton & Baumeister, 1991; McManus & Waller, 1995), negative emotions (Johnson, Schlundt, Barclay, Carr-Nangle, & Engler, 1995), anxiety associated with eating (Rosen & Leitenberg, 1985), and actual or anticipated negative social feedback (Craighead, Allen, Craighead, & DeRosa, 1996).

Other theoretical models and observational findings have alternatively placed emphasis on the positive reinforcing functions that disordered eating may serve. Reinforcers associated with disordered eating include the satiation of hunger with a binge episode following a period of dietary restraint (McManus & Waller, 1995), the taste of desired foods (Loro & Orleans, 1981), the presence of others in the social context of eating (Johnson et al., 1995), social approval for a slim appearance (Garner, Vitousek, & Pike, 1997), the anticipation or experience of interpersonal success for realizing socially-defined ideals for female identity and attractiveness (Striegel-Moore, 1993), and the attainment of self-mastery or self-control (Garner et al., 1997).

Gray’s (1987) theory has proposed that sensitivity and responsiveness to reward as well as relief from aversion are associated with BAS activation, as indicated by behavioral impulsivity. From the existent research, there is an indication that some forms of disordered eating behavior frequently occur within a larger context of impulsivity. Binge eating, for example, has been associated with substance abuse disorders (Claridge & Davis, 2003) and engagement in a wide variety of impulsive behaviors (Vitousek & Manke, 1994). Similarly, Wolfe, Jimerson, and Levine (1994) reported that self-ratings of impulsivity among women with bulimia were significantly greater when compared to age-matched controls.

A large subset of persons with bulimia nervosa also has comorbid Cluster B (i.e., borderline, histrionic, narcissistic, and antisocial) personality disorders (e.g., Ames-Frankel et al., 1992; Gartner, Marcus, Halmi, & Loranger, 1989; Lilenfeld et al., 1997; Rossiter, Agras, Telch, & Schneider, 1993). These personality disorders, strongly linked to impulsivity (Farmer & Nelson-Gray, 1995; Farmer et al., 2004), have also been observed among samples of persons with anorexia nervosa and binge eating disorder, but usually at a rate less than that observed for individuals with bulimia (e.g., Gartner et al., 1989; Wonderlich, 1995; Yanovski, Nelson, Dubbert, & Spitzer, 1993). Similarly, a subset of individuals with bulimia and anorexia nervosa often display a variety of impulsive behaviors, such as stealing, drug and/or alcohol use, suicidal behavior, and self-mutilation (e.g., DaCosta & Halmi, 1992; Lilenfeld et al., 1997).

In addition to an association with impulsivity, there is also an indication that eating disorders are strongly linked to anxiety disorders (e.g., Braun, Sunday, & Halmi, 1994; Claridge & Davis, 2003; Schwalberg, Barlow, Alger, & Howard, 1992) or elevations on self-report measures of BIS functioning (Loxton & Dawe, 2001). Relative levels of anxiety/punishment sensitivity and impulsivity/reward sensitivity appear to vary as a function of the type of eating disorder. That is, individuals with anorexia tend to have comparatively low sensitivity to reward relative to the general population (weak-BAS), whereas those with bulimia tend to be somewhat higher on reward sensitivity (strong-BAS) (Davis & Woodside, 2002) and more likely to engage in active escape via purging, which can interpreted as consistent with the concept of relieving non-punishment and BAS activation (Farmer, Nash & Field, 2001).
There is empirical evidence for increased reward sensitivity among people who demonstrate dysfunctional eating patterns, particularly among those who have bulimia or engage in binge eating. Kane, Loxton, Staiger, & Dawe (2004) found that women with bulimia, who were more impulsive on self-report measures of impulsivity than controls, displayed behavioral enhancement with financial reward incentives during a card sorting task. In this study, performance on the card-sorting task was also associated with scores on a self-report measure of state anxiety, thus suggesting an influence of anxiety/BIS activation on task performance as well.

In Farmer et al. (2001), among people diagnosed with bulimia, reward sensitivity as indicated by performance on a verbal learning task was marginally significantly correlated with purge frequency \( r = .56, p < .05 \), and moderately but not significantly associated with binge frequency \( r = .38, \text{ns} \), with the latter finding perhaps failing to reach significance due to a relatively small sample size \( n = 13 \). Similarly, among those diagnosed with eating disorder not otherwise specified (EDNOS, \( n = 21 \)), reward sensitivity displayed a moderate and significant correlation with purge frequency \( r = .46, p < .05 \) but not binge frequency \( r = .02, \text{ns} \).

Craighead et al. (1996) found that college women with a current diagnosis of bulimia nervosa and past history of depression learned a computerized mental maze task significantly faster when they received negative social feedback (i.e., punishment) for errors than when they received positive social feedback (i.e., reward) for correct responses. Participants with bulimia also learned the maze faster under the negative social feedback condition than did two control groups without a history of an eating disorder who received positive social feedback following correct responses. Craighead et al. (1996) interpreted these findings as indicative of the degree to which women with bulimia experienced negative social feedback as aversive that, in turn, stimulated the use of avoidance responses that resulted in exceptional performance. They go on to speculate that the active avoidance of anticipated negative social feedback may mediate the use of compensatory (e.g., vomiting) behaviors and contribute to an over-concern with shape and weight.

Craighead et al.'s (1996) interpretation of their findings is consistent with Schmauk's (1970) conceptualization of the mental maze task as an avoidance learning procedure in that performance during the task requires the adoption of an alternative problem solving approach (i.e., active avoidance response) as a means of avoiding an aversive interpersonal event (i.e., negative social feedback). Interpreted in this way, performance on this task may also be indexed by the degree of sensitivity to Gray's concept of relieving non-punishment, which is associated with the BAS hyperactivity, impulsivity and reward sensitivity. Alternatively, women with bulimia under the condition of negative social feedback may have learned the maze quicker due to a heightened sensitivity to punishment cues (and a corresponding passive avoidance tendency) that, in Gray's model, is linked to the anxiety dimension and BIS activation. The anxiety dimension alone, however, may be less relevant in accounting for behavioral differences in this research as trait anxiety appeared to have no relationship to task performance (Craighead et al., 1996, Footnote 1, p. 555).

Taken together, individuals who habitually engage in purge and perhaps binge behaviors may do so for their reinforcing functions. Specifically, such individuals may be more inclined to respond to aversive internal and external events by actively engaging in disordered eating behaviors that provide an escape from these aversive events. Similarly, some individuals might engage in disordered eating so as to experience the attainment of immediate reinforcers, or to increase the likelihood of attaining anticipated distal reinforcers (e.g., social approval for slim appearance).
Substance Abuse

As is the case with eating disorders, theories on the functional nature of addictive behavior have emphasized two processes, each of which constitute an “abuse of reinforcers” (Fowles, 2001). As negatively reinforced behavior, substance abuse is instrumental in producing relief (escape) from aversive states or environments (e.g., negative emotion, social anxiety). As positively reinforced behavior, substance abuse is instrumental in producing pleasant or desirable consequences (e.g., euphoria). This is exemplified in taxonomies of problem drinkers, which suggest two basic subtypes: those with anxious-dependent traits and those with a more chronic abuse pattern coupled with antisocial traits (Wulfert, Greenway, & Dougher, 1996). Recent research, however, has tended to more strongly implicate impulsive and sensation seeking traits rather than anxiety with excessive alcohol consumption (Grau & Ortet, 1999; Loxton & Dawe, 2001). Relatedly, there is evidence for individual differences in the sensitivity to consequences that follow alcohol consumption, which is experienced as more rewarding among some individuals than others (Tabakoff & Hoffman, 1988). Those who experience alcohol consumption as more rewarding are, in all likelihood, at greatest risk for alcohol misuse.

The reinforcing effects of drugs have long been recognized and, as Leshner (1997) has reviewed, most drugs of abuse either directly or indirectly affect the activities of the mesolimbic dopaminergic reward system pathways. Similarly, Fowles (2001, p. 94) has suggested the following:

“A dominance of the BAS over the BIS would produce both an impulsive temperament and a bias toward the positively reinforcing effects of drugs of abuse over the delayed negative consequences. Therefore, to the extent that this approach–avoidance conflict model applies, appetitive and aversive motivational constructs are particularly useful to understanding substance abuse.”

A series of laboratory tasks also point to the dominance of BAS over BIS in relation to features associated with alcohol consumption or drug abuse. Franken (2002), in a study consisting of alcohol drinkers, found that self-reported BAS but not BIS predicted self-reported urge to drink following exposure to alcohol-related cues. Similarly, Kambouropoulos and Staiger (2001) found that students who consume relatively high amounts of alcohol demonstrated enhanced performance following exposure to alcohol cues on a card-sorting task with financial reward incentives. The authors interpreted this finding as consistent with strong-BAS activated approach behavior following exposure to idiosyncratically-based rewarding stimuli, namely alcohol cues. Madden, Petry, Badger, & Bickel (1997) and Kirby, Petry, and Bickel (1999) found that heroin users relative to controls were more likely to choose immediate and smaller rewards over larger and delayed rewards (i.e., demonstrate higher discount rates for delayed rewards), and that delay-discounts rates were also positively associated with self-reports of impulsivity.

Other findings (e.g., Kambouropoulos & Staiger, 2004) suggest that depending on the nature of the cues present and the availability of alcohol for consumption, both appetitive and aversive motivational processes exert influence on the valence of affect and the urge to drink. In Kambouropoulos and Staiger (2004), findings suggested that in frustrating and nonrewarding drinking situations, BIS activation might be associated with an increased urge to drink as a means of terminating or modifying resultant negative affective states.

Anxiety Disorders

In Gray’s model (Gray & McNaughton, 2000), BIS hyperactivity results in a cognitive set that results in the exaggeration of a negative bias, or bias that amplifies the experience of threat, danger, or potential punishment. The affective consequence of this activation pattern is the experience of excessive
anxiety. BIS hyperactivity and associated processes are thought, most prototypically, to underlie the experience of generalized anxiety disorder.

Panic episodes and “alarm reactions” would be associated with the FFFS system, as would active avoidance (escape) of innate and conditioned fear stimuli. However, the anticipatory anxiety and environmental scanning for safety signals might better reflect the activities of the BIS. Other anxiety conditions, such as agoraphobia and social anxiety disorder, can represent the joint effects of strong activation and reciprocal feedback among BIS and FFFS systems (Gray & McNaughton, 2000). As noted by Fowles (2001), Gray’s BIS theory also provides a theoretical account for the associations between the experience of anxiety and characteristics like shyness, worry, anxious apprehension, social withdrawal, discouragement, and pessimism.

A number of lines of research converge to support Gray’s basic theory of anxiety (BIS) and pure fear (FFFS). In addition to a number of studies that support BIS-activation with responsiveness to punishment and BAS-activation with responsiveness to rewards (e.g., Boddy, Carver, & Rowley, 1986; Corr, 2002; Gupta, 1976; McCord & Wakefield, 1981; Zinbarg & Mohlman, 1998), drug studies with clinical populations suggest that anxiolytics (specifically benzodiazepines) reduce the experience of anxiety (BIS) but have little or no effect on panic or phobias (Gray & McNaughton, 2000), thus supporting the BIS/FFFS distinction. Considerable evidence also has accumulated that associates the experience of anxiety with heightened attention to stimuli associated with threat, punishment, or danger (Williams, Watts, MacLeod, & Mathews, 1997).

Implications for Behavioral Case Formulation and Therapy

General Treatment Implications

In behavior therapy, a useful pre-treatment assessment goal is to evaluate whether problematic response classes are primarily maintained by positive reinforcement, negative reinforcement processes associated with active escape, or by passive avoidant behavior. These different possibilities have implications for the nature of the therapy that might be most beneficial for the client concerned (Wulfert et al., 1996). Generally, behavior maintained through the process of positive reinforcement is influenced by the reward or pleasure–enhancing function it serves, whereas active escape behavior maintained by negative reinforcement processes associated with the avoidance of aversive contexts, while passive avoidance functions to maintain anxiety in the various contexts where it is generally experienced (Wulfert et al., 1996).

Response classes maintained by positive reinforcement can often be addressed by interventions that eliminate the reinforcers that support problematic behavior. Similarly, differential reinforcement procedures are often employed. These involve the replacement of an inappropriate response with another response, with this accomplished through the reinforcement of one set of responses while another set of responses (i.e., the inappropriate ones) is placed under an extinction schedule. Similarly, self–control strategies such as Premacking, public statement related to behavioral goals, cost–benefit analyses, and counter–regulation methods are among those that can be employed to bring problematic impulsive behavior under the influence of alternative reinforcement contingencies incompatible with impulsive action.

Other therapeutic interventions for reducing impulsive behavior include those that perhaps enhance the development of an approach–avoidance conflict in the presence of rewarding stimuli. Such interventions would increase the salience of stimulus cues that might signal aversive outcomes associated with such behavior and, as a result, facilitate the inhibition of behavior. For example, Pliner
and Iuppa (1978) found that the presence of a mirror when eating reduced food consumption among obese participants.

For problematic behaviors that are maintained by reinforcement and enacted where an approach–avoidance conflict is already evident (i.e., where FFFS, BAS and BIS are simultaneously activated), therapeutic strategies that specifically target a client’s verbal behavioral repertoire that reveals ambivalence or a desire to change problematic behaviors should be considered (Wulfert et al., 1996). These would include strategies such as motivational interviewing (Miller & Rollnick, 1991) or aspects of Acceptance and Commitment Therapy (ACT; Hayes, Wilson, & Strosahl, 1999).

For problematic response classes that are manifestations of active or passive avoidance strategies, exposure–based interventions might be implemented to block behavioral avoidance while promoting habituation or extinction to stimuli that are associated with negative emotional states. Similarly, behavioral activation strategies (Martell, Addis, & Jacobson, 2001) might be employed to facilitate exposure and to assist the individual in coming into contact with naturally occurring reinforcers associated with behavioral action. Both of these types of interventions are further described below in the section on anxiety and anxiety disorders.

**ADHD**

Given the heterogeneity with which ADHD is expressed across persons (as reflected in the ADHD subtypes), the different patterns of disinhibition noted among males and females with ADHD, and the high degree of comorbidity that ADHD has with other conditions, careful detailed idiographic assessments of youth with ADHD are indicated.

From the findings presented above, family and classroom-based psychosocial interventions that involve self-monitoring, attention to and discrimination of response-consequence relations, problem solving skills training, a token economy reward system to strengthen desired behavior, and a response-cost intervention to reduce misbehavior may be among the most efficient intervention strategies given the nature of motivational anomalies observed among members of this group. Other forms of therapy (e.g., parent training, social skills training, medication monitoring) not directly related to motivational anomalies may also be indicated.

**Substance Abuse**

Given research that suggests BAS over BIS dominance among substance abusers, treatment programs such as those that offer tangible rewards (e.g., gift vouchers, movie passes) for drug non-use, as indexed by clean urine samples, may prove to be effective (e.g., Silverman, Higgins, Brooner, & Montoya, 1996), as well as other programs like Alcoholics Anonymous and Narcotics Anonymous that provide social reinforcement for abstinence. Similarly, given that there is some evidence of BIS activation among at least a subset of those who misuse substances, therapeutic strategies that attempt to work with the resultant approach–avoidance conflict or ambivalence might be especially helpful for increasing the salience of aversive outcomes associated with continued substance use (e.g., Miller & Rollnick, 1991). In Gray’s terminology, such techniques would facilitate some degree of BIS control, or perhaps even BIS dominance over BAS, in situational contexts where alcohol-related cues are present.

**Eating Disorders**

Given that eating disorders are often comorbid with substance abuse (e.g., Kane et al., 2004), interventions similar to those used for substance abuse may be beneficial in the treatment of eating disorders. These would include motivational interviewing and the contingent awarding of tangible and social reinforcers for the non-engagement in disordered eating behaviors.
Other intervention strategies have also proved to be useful for treating eating disorders. In the case of bulimia, periods of self-starvation often precede binge episodes. Often, for people who restrict food intake, self-starvation has associated social reinforcement qualities (e.g., social approval for a slim appearance). However, intense hunger, a consequence of food deprivation, is itself a strong establishing operation that increases the reinforcing properties of food consumption, thus making binge episodes more probable. For these reasons, several effective approaches to the treatment of bulimia place strong emphasis on the elimination of self-starvation practices and the subsequent normalization of eating habits and behavior (e.g., Agras, Schneider, Arnow, Raeburn, & Telch, 1989).

Following binge episodes, persons with bulimia will often experience a panic-like phobic response to feelings of fullness. They often equate such feelings with the physical manifestation of becoming fat. Because the idea of becoming fat is equated with the loss of social reinforcers (e.g., social approval, important social relationships), purge episodes often follow binge episodes, whereby purges function to terminate the aversive state of fullness and associated anxiety. For this reason, some approaches to treatment for bulimia seek to simultaneously expose clients to feelings of fullness while blocking the purge response (e.g., Leitenberg, Rosen, Gross, Nudelman, & Vara, 1988). These approaches are premised on the notion that purging following eating is partly responsible for the maintenance of binge eating, as purging is conceptualized as a form of escape response maintained by the relief from discomfort that it provides. The non-reinforced exposure to feelings of fullness, it is thought, promotes the habituation or extinction of anxiety to these somatic cues.

Another therapeutic approach might be to inform the client that the escape response, purging, is not a particularly effective form of escape. Studies on the effectiveness of vomiting and laxative use indicate these acts only eliminate a modest amount of calories consumed during a typical binge episode. In the case of laxatives, only about 12% of consumed calories are eliminated through this process (Bo-Linn, Santa Ana, Morawski, & Fordtran, 1983), and less than 50% of calories consumed during a binge, on average, are expelled as a result of self-induced vomiting (Kaye, Weltzin, Hsu, McConaha, & Bolton, 1993).

Although the binge/purge cycle of bulimia can be regarded as maintained by short-term active avoidance strategies, there are a number of longer-term negative health-related consequences that make bulimia-related behaviors problematic (e.g., electrolyte abnormalities, renal damage or failure, chronic diarrhea and/or constipation, stretching, weakening and/or tearing of the esophageal wall). As such, therapeutic interventions might seek to capitalize on the conflict or ambivalence that many people with bulimia have about their relationship with food by highlighting the negative distal consequences associated with this condition. Motivational interviewing techniques (Miller & Rollnick, 1991) might be especially well–suited for this purpose, as they might tip the balance in favor of BIS dominance and, consequently, reduce the impact of reward contingencies that maintain dysfunctional eating behaviors.

Anxiety Disorders

Even though Gray (Gray & McNaughton, 2000) has discounted a classical conditioning role in the development of most phobias, he does ascribe to the view that phobias will continue to be maintained to the extent that the phobic stimuli are avoided. Perhaps for this reason, there is considerable evidence that associates greater therapeutic gains with the amount of time the individual is exposed to the phobic situation or object (Gray & McNaughton, 2000), which perhaps allows for greater opportunities for phobic reactions to feared stimuli to habituate or extinguish.

Cognitive and cognitive behavioral therapies have likewise been demonstrated to be effective treatments for anxiety and panic disorders. The cognitive elements of these therapies are likely effective as they facilitate alterations in thinking about events and experiences initially regarded as threatening. For
example, those who experience episodic panic attacks demonstrate a tendency to exaggerate the significance of certain bodily sensations, whereby such sensations and experiences are interpreted as reflective of a catastrophic physical event (e.g., heart attack, impending death) or mental event (e.g., loss of control, going crazy) rather than as concomitants to the experience of anxiety (Clark, 1997; Craske & Barlow, 2001).

Similarly, in the case of social anxiety disorder, individuals with this condition might hold a number of assumptions about themselves in relation to others or social contexts (e.g., “Unless someone shows that they like me, they dislike me;” “Unless I am liked by everyone, I am worthless”) (Clark, 1997). In both cognitive and cognitive behavioral therapies for panic and social anxiety disorder, some form of exposure to feared stimuli is often employed (e.g., to internal somatic sensations, to feared objects, to anxiety-evoking social contexts). Similarly, to test the validity of assumptions that underlie the client’s beliefs or rules that influence behavior, behavioral experiments are often employed (Beck, Rush, Shaw & Emery, 1979). Response prevention, avoidance blocking, and the elimination of safety cues and behaviors are also frequently used to facilitate habitation and extinction of fear or anxiety to cues and contexts erroneously perceived as excessively threatening or dangerous. Relatedly, the therapeutic focus on maladaptive or erroneous assumptions and beliefs can further reduce the disproportionate weight that is placed on possible but improbable threats, punishers, or negative outcomes. Taken together, behavioral, cognitive, and cognitive behavioral therapies might function to reduce erroneous adverse associations, perhaps principally by reducing the perceptual/cognitive bias toward excessive threat and, correspondingly, the experiences of both fear and anxiety (Gray & McNaughton, 2000).

Behavioral activation (BA) is a therapeutic approach that attempts to increase behavior that is at a low rate or has become less frequent for a variety of reasons, including previous punishment of behavior, the habitual engagement in escape and avoidance behaviors, and the frequent enactment of behaviors that result in inactivity (e.g., rumination, seclusion). As negative mood states are thought to reflect BIS activation (Gray & McNaughton, 2000), BA can be viewed as an attempt to override the effects of BIS as the client is encouraged to become behaviorally active even when mood states are negative. One feature of this therapy involves the identification of activities to try, and the problem-solving of any difficulties that make it difficult to carry out these activities. BA also actively discourages the use of avoidant coping strategies, and instead encourages re-entry into situations that are currently avoided but that were previously reinforced. BA also aims to teach clients to be astute observers of their own behaviors, and to make connections between behavior and variations in mood. In relation to this aspect of therapy, a client would be encouraged to notice the types of situations that either set the occasion for behavior to occur or inhibit a behavior’s occurrence. Similarly, clients are instructed in how to identify the consequences that follow their behavior, and how such consequences are related to variability in mood. All of this is done in the service of helping the client achieve his or her life goals (distal positive reinforcers), while at the same time challenging inhibited patterns of behavior based on anticipated punishers or disappointing outcomes.

Summary

As this review suggests, a consideration of temperament features in relation to problematic client behavior can usefully inform a behavioral assessment and therapy. Temperament, largely regarded as biologically-based and initially evident early in life as variations in reactivity and self-regulation, reciprocally influences the types of environments one enters as well as the effects that environments have on one’s behavior. In relation to clinically relevant behavior, temperament may mediate or moderate environment–behavior relations, and thus have relevance for the complete account of behavior. Of particular importance is the association between an individual’s temperament and sensitivity and responsiveness to environmental cues and behavioral consequences, as individual differences in these
areas can have implications for functional understandings of behavior as well as the selection of therapeutic interventions that are potentially most effective.

References


Please address editorial correspondence to:

Richard Farmer, Ph.D.
Department of Psychology
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

Tel: +64-3-364-2987, ext. 7196
Fax: +64-3-364-2181
E-mail: richard.farmer@canterbury.ac.nz

Author's note. This research was partially supported by Research Grant U6520 from the University of Canterbury. Correspondence regarding this article should be addressed to Richard Farmer, Department of Psychology, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, or by e-mail at richard.farmer@canterbury.ac.nz.

---

**Advertising in The International Journal of Behavioral and Consultation Therapy**

Advertising is available in IJBCT. All advertising must be paid for in advance. Make your check payable to Joseph Cautilli. The copy should be in our hands 3 weeks prior to publication. Copy should be in MS Word or Word Perfect format and advertiser should include graphics or logos with ad copy.

The prices for advertising in one issue are as follows:

- 1/4 page: $50.00
- 1/2 page: $100.00 horizontal or vertical
- Full Page: $200.00

If you wish to run the same ad in all four issues for the year, you are eligible for the following discount:

- 1/4 Pg.: $40 - per issue
- 1/2 Pg.: $75 - per issue - vertical or horizontal
- Full Page: $150.00 - per issue

For more information, or place an ad, contact: Joe Cautilli via e-mail at jcautilli@cctckids.com or by phone at (215) 462-6737.