Conceptual Models of the Nervous System:
Implications for Antisocial Behavior

JOSEPH P. NEWMAN

An organism responds in a particular way to a particular situation because of the state in which it currently finds itself. That state can only be contained as a physical specification of the cells which constitute the organism, above all, as a specification of the neuro-endocrine system. Thus, both genetic and environmental factors and interactions between them must act by altering the physiological basis of behavior (Gray, 1972, pp. 372-373, emphasis in original).

To the extent that this statement is correct, knowledge about the entire array of factors governing antisocial behavior is likely to be advanced by a good working model of the neuroendocrine system. Moreover, by relating the diverse influences of genetic, developmental, and situational variables to a unitary context (i.e., the neuroendocrine system), a working model of this type encourages investigators to conceptualize these variables in a more integrated fashion. Such working models have been referred to by Gray (1972) and others (e.g., Hebb, 1955) as conceptual nervous system (CNS) models.

In essence, the CNS is a theory about the real nervous system (CNS). The significance of the CNS derives from our incomplete understanding of the real nervous system. Without such understanding, we can only speculate about how its functions are achieved and how it influences behavior. Thus, CNS models are hypothetical constructs that may be found useful or wanting and will probably require revision as knowledge about the real nervous system continues to evolve. In light of this uncertainty, it is essential to recognize that a CNS model may limit as well as facilitate research by determining which psychological processes and dimensions of behavior are examined and how.

Recognizing the shortcomings of an earlier CNS model, Hebb (1955) noted that the CNS "of 1930 was evidently like the gin that was being drunk about the same time; it was homemade and none too good." Continuing in the same vein, he proposed, "If we must drink we can now get better liquor; likewise, the conceptual nervous system of 1930 is out of date and—if we must neurologize—let us use the best brand of neurology we can find" (p. 243). The new brand of neurology espoused by Hebb related to then recent advances relating the brain stem structures to learning and motivation via the concept of arousal. Based on more recent discoveries relating the septo-hippocampal system (SHS) to learning and motivation via a more elaborate understanding of the brain's arousal systems, Gray (1972, 1987) appears to have uncorked a vintage CNS model (see also, Fowles, 1980).

Although CNS models are based on and constrained by knowledge of the actual nervous system, the CNS models employed by psychologists are heavily influenced by behavioral data. Indeed, Gray (1987) proposed that "the structure of the neuro-endocrine system inferred by the psychologist from purely behavioral data may be at certain points in advance of physiological knowledge" (p. 241, emphasis in original). The CNS, then, is a hypothetical construct formed by physiology on the one hand and elaborated by psychology on the other. Its purpose is to guide theorizing about causal influences on behavior and to enhance understanding of both psychological and physiological processes by integrating the two.

Owing to the complexity of CNS functioning, CNS modeling may be attempted at various levels. At the physiological level, CNS models might involve any of a wide variety of cortical, subcortical, or neurotransmitter systems. At the psychological level, CNS models might target learning, emotion, attention, or a range of other factors that influence behavior regulation, propensity to antisocial behavior, or both. Whereas one CNS model may focus on the implications of the ascending reticular activating system (ARAS) for arousal regulation (Eysenck, 1967), another may be concerned with the implications of the septo-hippocampal system (SHS) for behavioral inhibition (Gray, 1987), and a third model may highlight the implications of serotonergic pathways for impulsive aggression (see Berman, Kavoussi, & Coccaro, Chapter 28, this volume). Thus, the number of potential CNS models may be quite large.

The hypothetical nature of CNS models implies that no one CNS model can lay claim to being correct. How, then, should we select or evaluate alternative CNS models? CNS models are essentially heuristic devices; therefore, the metric for evaluating them rests, to a large extent, on their utility. The utility of a CNS model, in turn, is likely to depend on a number of factors, including, but not limited to, (a) the accuracy and clarity with which the CNS model reflects the physiological system modeled; (b) the applicability of its biopsychological constructs to behavioral syndromes of interest.
THEORETICAL CONTEXT AND CONTROVERSIAL ISSUES

To illustrate the use of cNS models to elucidate antisocial behavior, I focus on the application of Gray’s cNS model to psychopathy. Gray (e.g., 1972) is an articulate proponent of the cNS approach, and his model provides a particularly useful context for theorizing about psychopathy (e.g., Fowles, 1980; Fowles & Mischel, 1994; Lykken, 1995; Newman & Wallace, 1993a, 1993b; Patterson & Newman, 1993).

Eysenck’s Model

Before describing the specifics of Gray’s cNS model, I briefly discuss Eysenck’s (1967, 1981) theoretical framework because it provides an important context for Gray’s model and the cNS modeling of antisocial behavior more generally. Eysenck’s cNS model provides an excellent example of this approach to antisocial behavior (Eysenck, 1981; Eysenck & Eysenck, 1978; Eysenck & Gardnerson, 1989). According to the model, genetic factors contribute to individual differences in physiological processes (e.g., ascending reticular activating system [ARAS]) that, in turn, drive fundamental psychological processes (e.g., motivation, learning) that give rise to stable differences in personality (e.g., introversion-extraversion [I-E]) and, ultimately, influence a person’s social behavior and risk for psychopathology (e.g., antisocial personality disorder).

A common confusion that arises when considering biological influences on social behavior relates to whether particular behaviors, such as driving under the influence of alcohol or committing an armed robbery, can be rooted in our biology. Are there really genetic or physiological processes responsible for such specific behaviors? In most cases, scientists investigating antisocial behavior are searching for characteristics that predispose a person to such behavior rather than direct causal relationships.

This predisposing role of physiological factors is intrinsic to Eysenck’s (1981) model. Specifically, personality is employed as an intervening variable. Rather than relating in specific differences in physiological functioning to antisocial behavior directly, Eysenck links them to general personality styles, which, in turn, make antisocial behavior more or less likely. In contrast to the diversity that characterizes antisocial behavior, personality dimensions are relatively homogeneous. Thus, linking personality to specific biopsychological processes seems more plausible. In addition, situating personality constructs between a physiological predisposition and antisocial behavior clarifies why particular predispositions do not always become manifested as antisocial behavior. Lykken (1982), for example, has proposed that a biologically based trait of fearlessness may predispose a person to heroic or antisocial deeds, depending on their socialization.

Within Eysenck’s (1981) model, the relation between personality and antisocial behavior is mediated by psychological processes such as perception, motivation, emotion, and cognition, which, in turn, influence learning, development, and social adjustment. To the extent that personality does, in fact, influence a person’s risk for antisocial behavior by systematically influencing that person’s reaction to environmental events, the contribution of personality to antisocial behavior must be understood in light of myriad person by situation interactions that shape an individual’s response to the environment. Accordingly, the significance of personality-based explanations of antisocial behavior depends on the extent to which the personality constructs invoked implicate specific psychological processes. Eysenck’s use of the ARAS to achieve this specificity provides a classic example of a cNS model.

Though recognizing the crucial role of experience in shaping behavior, Eysenck’s (1981) model focuses on the role of biopsychological processes in mediating one’s transactions with the environment. His general framework provides a model for other investigators interested in the development and application of cNS models. Recognizing its virtues, Gray (1981) has adopted Eysenck’s theoretical framework to develop his own cNS model. Whereas Eysenck’s cNS model is predicated on the ARAS, Gray’s (1987) model is based primarily on the septo-hippocampal system.

Gray’s Model

As shown in Figure 30.1, Gray’s model has three interacting systems: the behavioral activation system (BAS), the behavioral inhibition system (BIS), and the nonspecific arousal system (NAS). Each system plays a crucial role in the regulation of behavior. The BAS is sensitive to cues for reward and active avoidance and functions to (a) increase NAS activity, (b) inhibit activity in the BIS, and (c) initiate motor behavior in the service of approach or active avoidance. The BIS is sensitive to cues for punishment and nonre-
ward and serves to (a) increase NAS activity, (b) interrupt ongoing or anticipated motor behavior, and (c) direct attention to significant stimuli. The NAS receives inputs from both the BAS and the BIS and acts to increase the intensity (i.e., speed and force) of behavior. As indicated by the two switches in the decision mechanism that "turn on" approach behavior or the stop/inspect response, the BAS and BIS compete to influence the focus of behavior. If BAS activation is stronger, people will maintain their focus and initiate goal-directed behavior, whereas they will pause and redirect attention to environmental cues (i.e., initiate passive avoidance) if BIS activity predominates. NAS activity, on the other hand, influences qualitative rather than directional aspects of behavior. In addition to increasing the speed and force of whatever behavior eventually occurs, as proposed by Gray, increases in NAS activity will tend to increase the focus of attention and speed with which behavior is initiated, thus limiting the amount of time and concomitant processing accorded to the simultaneous evaluation of BAS and BIS inputs (see Wallace, Bachorowski, & Newman, 1991; Wallace & Newman, in press).

Although schematic representations of Gray’s model are often limited to the three systems described earlier (e.g., Fowles, 1980; Newman & Wallace, 1993), an additional system—the fight/flight system (FFS)—has been receiving increased attention (Gray, 1987), especially with regard to reactive aggression (e.g., Gray, 1991; Quay, 1993) and panic disorder (Fowles, 1993a). In contrast to the BAS and BIS, which respond to conditioned stimuli, the FFS is activated by unconditioned aversive stimuli such as pain and the termination or omission of reward. Once activated, the FFS is associated with extreme autonomic nervous system activity and intense motor activity of the type associated with frantic struggle, fleeing an aggressor, or lashing out in counterattack.

**Evaluation of the Eysenck/Gray Model**

With regard to the first criterion for evaluating CNS models, the accuracy and clarity with which the CNS model reflects the physiological system modeled, Gray’s model is based primarily on the implications of the septo-hippocampal system (SHS), although it is also informed by the effects of antianxiety drugs and, to a lesser extent, other CNS pathways (see Gray, 1994). The psychological counterpart of the SHS is the BIS. Gray has been developing the implications of SHS functioning for behavioral inhibition and anxiety for more than 25 years (see Gray, 1970). A detailed critique of his proposals regarding the physiology of anxiety and behavioral inhibition is beyond the scope of this chapter. Although there is no consensus regarding the exact functions of the SHS or the physiological substrate of anxiety, Gray’s proposals are consistent with a great deal of experimental evidence concerning the consequences of lesioning the SHS as well as with the evidence concerning the effects of antianxiety drugs (see Gray, 1982, 1987). Although Gray acknowledges that his formulations will require additional modification, his characterization of SHS functioning is plausible, comprehensive, and up-to-date (see also, Gray, 1991, 1994).

With regard to (b), the applicability of the model to relevant behavioral syndromes, Fowles (1988, 1993b) has provided numerous examples of the applicability of Gray’s model to psychopathological syndromes, including psychopathy. Supporting the applicability of the model to antisocial behavior, numerous investigators have used Gray’s model as a framework for describing their perspective on antisocial behavior (Fowles, 1980; Kilian & Cloninger, 1993; Lykken, 1995; McClellan, 1991; Milich, Hartung, Martin, & Haigler, 1994; Newman & Wallace, 1993a, 1993b; Quay, 1993; Walker et al., 1991).

The third criterion for CNS models noted earlier relates to (c), the ease with which the model’s biopsychological constructs connect with other relevant constructs so that it is useful for organizing related theory and evidence. Investigators of antisocial behavior have offered a variety of proposals concerning the behavioral, personality, psychophysiological, and biochemical correlates of antisocial behavior. For example, numerous investigators have identified deficient passive avoidance learning as a key correlate of psychopathy as well as criminality and weak socialization in general (Lykken, 1957; Patterson & Newman, 1993; Trase, 1978). Passive avoidance involves the inhibition of specific behaviors that have been associated with punishment. Passive avoidance is integral to Gray’s model and has been used by Gray to label the output of the BIS (e.g., 1987). Because any of the factors characterized in Gray’s model could influence the output of the BIS, the model provides a rich source of hypotheses concerning passive avoidance deficits (see Newman & Wallace, 1993a).

With regard to personality, antisocial behavior has been linked to a wide variety of personality traits and other traitlike individual difference variables, including, but not limited to, aggressiveness, impulsivity, sensation seeking, thought disorder, low anxiety/harm avoidance, low arousal, low intelligence, lack of empathy, poor...
conditionability, and impaired executive functioning. Though these characteristics are quite diverse, it is possible to conceptualize them as (a) driving motor activity (e.g., aggressiveness, impulsivity, sensation seeking), (b) weakening inhibition (e.g., poor fear conditioning, lack of empathy, low anxiety/harm avoidance), or (c) interfering with judgment/information processing (e.g., intelligence, thought disorder, impaired executive functioning). Moreover, these traitlike correlates of antisocial behavior have been linked to the three major factors that emerge consistently in factor analytic studies of personality: These factors are typically labeled extraversion, neuroticism, and psychoticism (e.g., Eysenck & Eysenck, 1978); impulsivity, anxiety, and socialization (Barratt & Patton, 1983); novelty seeking, harm avoidance, and reward dependence (Cloninger, 1987); or positive emotionality, negative emotionality, and constraint (Tellegen, 1985).

As was the case with passive avoidance, Gray's model has been linked to the major dimensions of personality and, thus, may be suitable for conceptualizing the biopsychological correlates of antisocial behavior at this level as well. The association between Gray's BAS (i.e., behavioral approach) and impulsivity is quite straightforward as is the relation of Gray's BIS (i.e., behavioral inhibition) to anxiety/negative emotionality (see Gray, 1991; Lauren & Ketelaar, 1989).

The significance of Gray's model for the third dimension of personality, sometimes described as psychoticism, psychopathy, socialization, constraint, conformity, and reward dependence, is less clear. Both Gray (1991) and Fowles (1993a) associate psychopathy with a weak BIS, thus positing that anxiety and psychopathy anchor opposite ends of the same continuum (see Gray, 1991). In identifying psychopathy with low anxiety and a weak BIS, however, this strategy does little to distinguish the third dimension of personality. Toward this end, Gray (1991) relates psychosomatic to the FFS and, though acknowledging potential problems, he also associates Tellegen's (1982) constraint factor with the FFS. Fowles and Missel (1994), however, propose a different strategy for distinguishing the anxiety and constraint factors. They relate the constraint factor, which includes a fear/harm avoidance component, to the BIS, and they map the more reactive aspects of anxiety onto the FFS. Concerning this proposal, it is worth noting that Fowles and Missel relate psychopathy to low constraint (i.e., a weak BIS) rather than low anxiety (see also, Lykken, 1995).

The uncertain applicability of Gray's model to the socialization/constraint dimension is ironic given the model's immediate bearing on passive avoidance learning—a process regarded as fundamental to the socialization process. Within Gray's model, the most obvious explanation for poor passive avoidance learning involves the direct effects of the BIS because passive avoidance is the behavioral output of this system. Though psychopathy is associated with poor passive avoidance, it remains to be established whether passive avoidance learning is more closely linked with the anxiety/negative emotionality (e.g., Gray, 1991) or socialization/constraint dimensions (e.g., Fowles & Missel, 1994). Alternatively, as discussed in the section on current findings, individuals with low levels of anxiety and constraint may display weak passive avoidance for different reasons—a possibility with important implications for the relation of Gray's model to socialization/constraint.

Gray's biopsychological constructs also have been associated with numerous psychophysiological and biochemical correlates of antisocial behavior. A relatively large number of investigators have reported that psychopaths display less electrodermal activity in anticipation of aversive events, whereas their heart rate (HR) is equal to or greater than that of nonpsychopathic controls (Hare, 1978). In a thoughtful and thorough analysis of the relation of Gray's model to psychophysiological measures of arousal, Fowles (1980) proposed that electrodermal activity may be associated with Gray's BIS, whereas HR is often associated with BAS activity.

Concerning the biobehavioral evidence on antisocial individuals, the relation between Gray's CNS model and antisocial behavior has been the subject of several reviews (e.g., Lewis, 1991; McBurtney, 1991; Quay, 1993; Rogeness, Javors, & Ploszka, 1992). Although recognizing the inherent complexity of neurotransmitter systems and the pitfalls involved in mapping neurochemistry to neuroanatomy and behavior, the authors have offered tentative proposals concerning the relation of dopamine (DA), norepinephrine (NE), and serotonin (5-HT) to processes specified by Gray's model. In brief, there appears to be good agreement that DA is involved in BAS functioning because of its association with the localization of rewarding and punishment stimuli and mobilization of active behavior to approach or avoid them (Cloninger, 1987; Depue, Luciana, Arbisi, Collins, & Leon, 1994). According to Rogeness et al. (1992, p. 771), "Regulation of the sephrohippocampal system [i.e., BIS] appears to be noradrenergic with additional regulation from serotonergic projections from the median raphe." More specifically, NE appears to be instrumental in regulating DA-dependent behaviors (i.e., BAS activity), whereas "serotoninergic mechanisms appear stronger in inhibiting irritable, aggressive behavior" associated with the fight/flight system (FFS; Rogeness et al., 1992, p. 772; cf. Spoornt, 1992).

The last criterion enumerated earlier concerns the extent to which a CNS model generates novel and valid predictions, thus promoting hypothesis testing and enhanced understanding. At a general level, Gray's model appears to accommodate many of the major factors that have guided research on the psychobiological underpinnings of antisocial behavior. Indeed, numerous chapters and journal articles have been written on the good "fit" or applicability of Gray's model to a wide variety of personality types and psychopathological conditions (e.g., Fowles, 1988, 1993b). Though such presentations promote the potential utility of Gray's model, they do not necessarily document its ability to advance the field.

In order to realize the contribution of Gray's CNS model for antisocial behavior, it is important for investigators to be clearer about the interrelated but distinct issues of applicability, involving the breadth and coverage of the model, and utility, stemming from the model's ability to generate specific, testable, and novel hypotheses. Although the applicability of a CNS model contributes to its potential utility, investigators must go beyond post hoc mapping of personality and psychopathological constructs onto Gray's model to realize this potential. Breadth and flexibility are important, but if the model is so broad and flexible that it can explain just about anything, then the validity of the "explanations" must be suspect.

Consider the predisposition to impulsive behavior. By one account, impulsivity may reflect a strong BAS, conferring
enhanced sensitivity to reward cues and a bias toward behavioral approach as opposed to avoidance (Gray, 1981; Gray, Owen, Davis, & Tsaltis, 1983). Alternatively, a weak BIS may diminish inhibitory control over response inclinations from the BAS, resulting in a tendency to act quickly without adequate regard for negative consequences (Fowles, 1980). A third explanation highlights the role of NAS activity in promoting the rapid initiation and intense maintenance of goal-directed behavior (Wallace et al., 1991). Finally, a fourth explanation focuses on the reciprocal inhibition of the BAS and BIS (Newman & Wallace, 1993a). Given the range of potential "explanations," Gray's model can account for impulsive behavior in diverse ways, but the etiology of specific instances of the behavior remains obscure.

This example raises an important issue concerning the distinction between applicability and utility. Though the broad applicability and provision of multiple pathways to dysregulation are strengths of Gray's model, investigators must be careful not to accept ostensibly compelling explanations without critically evaluating alternative ones. Relatedly, they must resist attributing phenotypically similar behaviors (e.g., impulsive behavior) to identical mechanisms without explicit examination of the potential pathways. It is possible to generate and test specific explanations for particular behaviors using Gray's model, but this generally requires the manipulation or measurement of multiple components of the model (see the next section on methodological issues). Overall, Gray's model appears to be broad enough to accommodate a wide range of behaviors, elaborate enough to suggest multiple explanations, and detailed enough to allow specific hypothesis testing. Parallelizing this discussion of impulsivity, a fundamental controversy emerging from the literature on cNS models of antisocial behavior concerns the extent to which psychopathy and other antisocial syndromes may be conceptualized as exaggerated approach behavior (e.g., strong BAS), low anxiety (e.g., weak BIS), or some other biopsychological process (e.g., low constraint, poor response modulation).

METHODOLOGICAL ISSUES

Sample Characteristics

To the extent that investigators hypothesize specific biopsychological predispositions to antisocial behavior, obtaining meaningful results depends on the homogeneity of the sample identified. In light of the heterogeneity that characterizes antisocial individuals, it is crucial to define subject groups carefully. Many investigators (e.g., Blackburn, 1983; Raine & Venables, 1981; Trasler, 1978) regard antisocial behavior as a failure of socialization influenced by a combination of biological and environmental factors. However, some cases of antisocial behavior appear to reflect environmental factors primarily, whereas biological factors seem to be paramount in others. If one's goal is to identify a biological predisposition, it is important to study subjects whose antisocial behavior is likely to have important biological determinants.

Though strategies for achieving homogeneous groups have not, as yet, been well elaborated, they may involve (a) eliminating subjects with strong environmental inducements to antisocial behavior, (b) focusing on subjects with especially chronic or extreme forms of antisocial behavior, (c) focusing on subtypes with a demonstrated family history, or (d) identifying and using biological/cognitive markers or personality traits to identify more homogeneous groups.

Laboratory Measures

As noted earlier, investigators have related the components of Gray's model to various biochemical, psychophysiological, behavioral, and self-report measures. Examination of this work suggests that each of these measures contributes a relatively unique and complementary perspective on Gray's cNS model, although the relation of these measures to specific components of Gray's model is inferential. Thus, investigators must exercise caution in using any one of these measures to draw conclusions about which aspect of Gray's model has been assessed.

For example, Fowles (1980) offers a compelling argument concerning the association of Gray's BIS with EDA, but the link between BIS functioning and EDA is, itself, only a hypothesis. In outlining his proposal, Fowles (1980) stated, "It is not necessarily being argued that EDA responds only to stimuli which activate the BIS. . . . EDA is seen in response to a wide range of stimuli" (p. 93). Thus, there is reason to be cautious in attributing group differences in EDA to BIS functioning in particular. Even when subjects' reactions to aversive events are assessed under well-controlled laboratory conditions, individual differences in EDA may relate to a variety of factors, including some that are independent of BIS functioning (e.g., active avoidance; Miller, 1979; Spitzer & Epstein, 1976). Thus, low EDA per se is not sufficient to identify psychopathy with a weak BIS.

Interpreting the biochemical evidence on antisocial individuals using Gray's model has been the subject of several reviews (e.g., Gray, 1991; Lewis, 1991; McBurnett, 1991; Quay, 1993; Rogeness et al., 1992). Although there appears to be some consensus regarding the relation of Gray's subsystems to the major neurotransmitter systems, the proposals are quite tentative and exceedingly complex. Although recent efforts to bridge the biochemical and behavioral domain using Gray's model as a framework represent an exciting development, the use of biochemical manipulations or measures to draw conclusions about group differences in BAS or BIS functioning is also premature.

A third means of assessing the components of Gray's model involves behavioral performance under well-controlled laboratory conditions. For example, evaluating the probability and intensity of responding for reward provides a reasonable index of BAS activity; the probability and amount of response suppression in response to punishment cues affords a measure of BIS activity; and a bias toward approach or avoidance under mixed-incentive conditions provides information about the balance between the systems (see Arnett, Smith, & Newman, in press; Newman & Wallace, 1993a; Quay, 1993). Moreover, there is some evidence that response speed and, to a smaller degree, attentional focus are informative with regard to NAS activity (Wallace et al., 1991; Wallace & Newman, in press).

Here too, it would be hazardous to draw conclusions without assessing other elements of the model. As already noted, the com-
ponents of Gray’s model are interactive and, thus, yield multiple pathways to a given outcome. Impulsive behavior may reflect intense BAS activation, high NAS reactivity, or a weak BIS. Although laboratory research on passive avoidance learning in psychopaths is often attributed to individual differences in the strength of the BIS, it too may reflect multiple mechanisms. The behavioral outputs of Gray’s model (i.e., approach, passive avoidance) reflect the resolution of a go/no-go conflict with multiple factors influencing the outcome (Newman & Wallace, 1993a). Multiple pathways notwithstanding, an advantage of behavioral measures concerns their face validity owing to Gray’s behavioral description of each system’s function.

Investigators have also proposed using self-report measures of personality and affect to characterize the psychological processes contributing to antisocial behavior using Gray’s model. Based on Tellegen’s mapping of positive emotionality to the BAS and negative emotionality to the BIS, for instance, it may be informative to assess positive and negative affect in antisocial offenders to draw inferences about BAS- and BIS-related processes, respectively (see Fowles, 1987). More recently, Fowles and Missel (1994) proposed that measures such as Schalling’s (1978) psychic anxiety and somatic anxiety scales may be useful for assessing BIS and FFS activity, respectively. Here too, however, investigators must be cautious because trait differences may reflect diverse psychological processes, including the cumulative effects of diverse experiences that are influenced by personality-related response inclinations (see Fowles, 1987; Frick, in press).

Overall, it appears that biochemical, psychophysiological, behavioral, and self-report measures may be used as rough estimates of BAS and BIS functioning suitable for characterizing antisocial individuals as BAS or BIS dominant. Regardless of the global nature of this distinction, it represents an important step in recognizing the diverse processes contributing to antisocial behavior and provides a context for integrating the correlates of antisocial behavior across multiple measurement domains.

On the other hand, a major thesis of this chapter is that CNS models offer the opportunity for greater precision in specifying the psychological processes underlying antisocial behavior; and, in this regard, the distinction between BAS and BIS dominant styles seems short of the mark. Indeed, the position advocated in this chapter is that realizing the unique potential of Gray’s model involves moving beyond the classification of antisocial syndromes according to the major components of his model. In my opinion, the unique contribution of CNS models relates to their ability to generate testable hypotheses with a degree of novelty and specificity that would not be possible without the detailed perspective that they provide (see Gorenstein & Newman, 1980).

CURRENT FINDINGS

This section reviews a select literature on the psychopath as a means of exploring the application of Gray’s model to antisocial behavior. In particular, I attempt to evaluate alternative mechanisms for psychopathic behavior derived from Gray’s model. This strategy is intended to highlight methodological issues pertaining to diverse interpretations, particularly the importance of using experimental manipulations and multiple measures. Among other reasons, psychopathy was selected because it is generally regarded as the prototypical antisocial syndrome (Gorenstein & Newman, 1980), it appears to have important psychobiological underpinnings (Hare, 1991; Hart & Hare, Chapter 3, this volume), and there is considerable evidence that it may be assessed reliably and validly using Hare’s Psychopathy Checklist (PCL; Hare, 1991). Moreover, evidence suggests that the PCL identifies a relatively homogeneous and extremely antisocial group with regard to number and range of crimes, frequency of violent and sex crimes, degree of substance abuse, and risk for recidivism (see Hare, 1991, for a review). Given the reliability of the PCL, its ability to identify a relatively homogeneous group of subjects, and its prolific antisocial behavior of the subjects identified, PCL-defined psychopaths represent a good sample for evaluating the application of Gray’s model to antisocial behavior.

The Weak BIS Hypothesis

In a particularly influential review of the implications of Gray’s model for psychopathy, Fowles (1980) proposed that many symptoms of psychopathy are “interpretable as a direct manifestation of a weak or deficient BIS” (p. 96). According to Fowles (1980), a weak BIS could explain psychopaths’ absence of anxiety and difficulty inhibiting behavior in the presence of punishment cues, problems learning from past punishments (passive avoidance) and inattention (extinction), strong reward-seeking behavior, tendency toward unsocialized active-avoidance (which, like reward, is mediated by the BAS), and their low tolerance for alcohol.

With regard to experimental evidence, Fowles (1980, 1993a) relies on pre-existing psychophysiological, behavioral, and personality data to link psychopathy and the BIS. Numerous studies document psychopaths’ weaker electrodermal (EDA) in response to stimuli that have been paired with aversive events and in anticipation of aversive events (Hare, 1978). If Fowles’s assumption linking EDA and BIS activity is correct, then such evidence is consistent with a weak BIS. With regard to behavioral evidence, Fowles (1988) cites laboratory studies demonstrating deficient passive avoidance learning (e.g., Lykken, 1957; Newman & Kosson, 1986; Newman, Wulom, & Nathan, 1985) and extinction (Newman, Patterson, & Kosson, 1987; Siegel, 1978) as evidence of deficient BIS functioning in psychopaths. The weak BIS hypothesis is also consistent with clinical characteristics of psychopaths emphasizing “absence of nervousness or psychomotoric manifestations” (Checkley, 1976, p. 224).

Fowles’s (1980) paper has been important for demonstrating the applicability of Gray’s model to psychopathy and antisocial behavior more generally (see also Fowles, 1988). To advance understanding, however, investigators must use the weak BIS hypothesis as a springboard to generate and test more specific alternative hypotheses rather than evaluating the global hypothesis.

The BIS carries out numerous functions related to checking stimuli, interrupting behavior, shifting the focus of attention, and initiating passive avoidance, all of which are subserved by different physiological systems (Gray, 1987, 1991). If the multicomponent nature of the BIS is granted, then it is important for proponents of the weak BIS hypothesis to be clear about which aspect of BIS
functioning is presumed dysfunctional in psychopaths or, alternatively, whether psychopaths are deficient in all aspects of BIS functioning. In the latter case, it would be useful to specify whether (a) some integrative dysfunction is presumed to render all of the separate processes comprising the BIS ineffective, (b) a problem at the input stage (e.g., insensitivity to punishment cues) precludes enactment of other BIS functions (i.e., the interruption of approach and redirection of attention), or (c) some other process (e.g., exaggerated approach) is presumed responsible for psychopaths’ failures to stop, reflect, and initiate passive avoidance.

Such elaboration might also clarify the extent to which the weak BIS hypothesis represents an advance over the low-fear theory. In discussing what he calls the Fowles-Gray-Lykken theory, Lykken (1995) stated that it “assimilates the Lykken low-fear theory, now the ‘weak-BIS’ theory” (p. 163). Using Gray’s model to recast long-standing theory and research involving psychopaths’ insensitivity to fear stimuli exercises the broad applicability of the model but may do little to advance the field. The utility of Gray’s model involves its ability to generate alternative hypotheses about the psychological processes contributing to inhibitory deficits—processes that may then be tested and further elaborated to increase understanding.

An example of the need for greater specificity concerns the common assumption that psychopaths’ poor passive avoidance learning supports a weak BIS interpretation. Psychopaths do not display a global deficit in passive avoidance learning (Newman & Kosson, 1986; Schmaek, 1970). Available evidence suggests that psychopaths fail to master passive avoidance contingencies when they are latent (Lykken, 1957) or otherwise require subjects to suspend goal-directed behavior to process secondary contingencies (Newman & Wallace, 1993b). When passive avoidance contingencies are salient from the outset of a task, so that subjects do not need to alter the focus of their goal-directed behavior to process them, then psychopaths perform as well as controls (Newman & Wallace, 1993b). Relatedly, psychopaths perform passive avoidance tasks as well as controls when trials are well spaced, allowing ample time to suspend goal-directed behavior and process punishment feedback (Arnett, Howland, Smith, & Newman, 1993; Newman et al., 1987).

Similar evidence exists regarding the situation specificity of psychopaths’ EDA (Arnett et al., in press). Low-anxious psychopaths displayed significantly less EDA than low-anxious nonpsychopathic in response to punishment cues while they were responding for rewards even though their EDA to the same cues was nonsignificantly greater than that of controls in the absence of reward cues.

The situation-specific nature of psychopaths’ deficient passive avoidance and EDA indicates that they are not globally insensitive to punishment cues. Deficient passive avoidance is seen most clearly on tasks in which subjects must alter a pre-established or otherwise dominant response set to process punishment feedback. Such evidence is not consistent with a global, weak BIS hypothesis or with proposals targeting the BIS input (i.e., sensitivity to punishment cues). More specifically, the data suggest a problem with the BIS interrupt function or some other process responsible for updating a person’s response strategy (see Damasio, 1994; Newman et al., in press; Patterson & Newman, 1993).

The BAS Dominance Hypothesis

The same year that Fowles (1980) related Gray’s SHS model (i.e., the BIS) to psychopathy, Gorenstein and Newman (1980) posed a “septal lesion” model for psychopathy and other syndromes of disinhibition. Although this model has been misrepresented as positing a “brain defect” (Lykken, 1995, p. 180), Gorenstein and Newman (1980) stated unambiguously that “suggesting such an analogy, we do not imply that human disinhibition can be traced to septal dysfunction. Rather, we suggest that behavioral analysis of the septal syndrome can elucidate basic psychological components of human disinhibition” (p. 302). On this basis, Gorenstein and Newman (1980) proposed that “one may hypothesize that hypersensitivity to rewards . . . characterizes human disinhibition and in turn plays a significant role in sustaining the known avoidance deficits” (p. 312). A more elaborate version of the reward dominance hypothesis has been set out by Quay (1988, 1993) with regard to the behavior problems of children with conduct and attention deficit disorder.

To explore this hypothesis, Newman and Kosson (1986) examined passive avoidance learning under conditions involving monetary punishments only or a combination of reward and punishment incentives. Although the tasks used in the two conditions were identical, psychopaths’ deficient passive avoidance was specific to the condition requiring them to inhibit reward seeking to avoid punishment (see also, Moses, Ratliff, & Ratliff, 1979; Newman et al., 1985; Schmaek et al., 1990; Thormiss & Zuckerman, 1995). Replicating and extending earlier work by Siegel (1978), Newman et al. (1987) found that psychopaths also perseverated responding for reward in a card-playing task despite the fact that their excessive reward seeking resulted in their earning significantly less money than controls (see also, S. K. Shapiro, Quay, Hogan, & Schwartz, 1988).

Arnett et al. (in press) employed a combination of behavioral and psychophysiological measures to evaluate BAS-related activity in psychopaths and nonpsychopathic offenders during a serial-reaction-time task. Subjects earned 5 cents each time that their response rate for five consecutive responses exceeded a criterion. As predicted by the strong BAS hypothesis, low-anxious psychopaths responded significantly faster than controls, but this effect was qualified by a higher order interaction indicating that support was due primarily to response rate during later trials. Because the reward-only trials were interspersed with reward-punishment trials, response speed during the later reward-only trials may have been influenced by factors other than reward.

Whereas the findings for passive avoidance and response perseveration are consistent with proposals that reward seeking has overshadowed avoidance, they provide little or no evidence that hypersensitivity to reward per se is responsible for this response style. Because weak inhibition may reflect multiple pathways in Gray’s model, establishing such influences requires more focused experimental manipulations. To examine sensitivity to reward cues more directly, Newman, Patterson, Howland, and Nichols (1990) employed a pattern-matching task involving reward-only and punishment-only incentives. Unlike impulsive college students, low-anxious psychopaths did not appear to be differentially activated.
by reward cues. Newman, Kosson, and Patterson (1992) examined delay of gratification under reward-only and reward-punishment conditions. Here too, low-anxious psychopaths appeared no more impulsive than controls in the reward-only condition. Similarly, low-anxious psychopathic delinquents performed as well as controls on a reward-only version of the passive avoidance task described earlier (e.g., Newman et al., 1985).

Gray (1987; see also Lykken, 1995) has proposed that the disinhibited behavior of primary or low-anxious psychopaths may reflect a weak BIS, whereas the behavior problems of high-anxious or neurotic psychopaths reflect exaggerated approach (i.e., a strong BAS). This review has focused on the evidence for hypersensitivity to reward cues in primary (low-anxious) psychopaths and, therefore, does not contradict proposals relating this mechanism to other disinhibited groups (see also, Newman & Wallace, 1993a).

The Response Modulation Hypothesis

Whereas the previous hypotheses focus on whole systems (e.g., the BIS) or their relative dominance, Gray's model also provides a useful framework for investigating a variety of specific processes that may hamper self-control (see Newman & Wallace, 1993a, 1993b; Nichols & Newman, 1986). The response modulation hypothesis provides a good example. The term response modulation was used by McCleary (1966) to describe the tendency of animals with septal and hippocampal lesions to persist in "dominant" responses despite punishment, extinction, or contingency reversal (i.e., response perseveration). Thus, response perseveration was viewed as a failure to revise response strategies based on new or unexpected information other than an imbalance between reward and punishment. Motivation. Similarly, the response modulation hypothesis attributes the self-regulatory problems of psychopaths to a reduced capacity to shift attention from the effortful organization and implementation of goal-directed behavior to its evaluation/modification (Newman et al., 1987, 1990; Patterson & Newman, 1993; see also, Damasio, 1994). Patterson and Newman (1993) proposed a four-stage model to elucidate response modulation and distinguish it from the insensitivity to punishment emphasis entailed by the weak BIS hypothesis (see also, Newman, 1987). At Stage 1 of the model, a person allocates attentional and motor resources to achieving a particular goal. If the goal-directed behavior is successful, then the sequence ends. However, if the person perceives proprioceptive or environmental feedback indicating that the goal-directed behavior is unsuccessful, then such information produces an increase in nonspecific arousal and a "call for processing" in Stage 2. Stage 3 involves the person's reaction to the call for processing. In one scenario, the person suspends goal-directed behavior and reorients attention (i.e., displays reflectivity) in which case the increase in arousal subserves cognitive processing of the unexpected/unscheduled circumstances. In another, the person fails to alter his or her ongoing (dominant) response set with the result that the arousal fuels active responding (i.e., disinhibition) related to the original goal-directed behavior. Stage 4 involves the consequences of the response style displayed at Stage 3. Whereas pausing to reflect on the arousal-eliciting event enables a person to adjust his or her response strategy and learn from experience, the disinhibited reaction facilitates active coping but reduces encoding of the stimulus conditions and behavior that gave rise to the problem. Although it is difficult to rule out problems associated with Stages 1 and 2, Patterson and Newman (1993) have proposed that the poor passive avoidance learning and other poorly regulated behavior of psychopaths relate to processes connected with Stage 3 of the model, which, in turn, interferes with consolidation of information at Stage 4 (see also, Newman & Wallace, 1993b).

Most, if not all, of the behavioral data cited in support of the weak BIS and BAS dominance hypotheses are also consistent with the poor response modulation hypothesis (i.e., psychopaths fail to suppress approach behavior despite punishment). Distinguishing among the alternative explanations requires careful analysis of the situations in which psychopaths display inhibitory failures as well as detailed assessment of the model's multiple components.

Although such studies are scarce, Newman and Wallace (1993b) suggest that available evidence is most consistent with the response modulation hypothesis. First, there is little evidence that psychopaths display regulatory deficits in reward- or punishment-only situations (Newman et al., 1990). Even though the EDA of psychopaths is distinctive in punishment-only situations, their disinhibited behavior is relatively specific to situations requiring subjects to suspend approach behavior in reaction to cues for punishment (Newman & Wallace, 1993b). Second, even in situations involving approach-avoidance conflict, psychopaths tend to perform as well as controls when the demands for response modulation are minimized by making the avoidance contingency salient from the outset (Newman et al., 1990; Newman, Wallace, Schmitt, & Amett, in press) or providing subjects with ample time to process negative feedback (Amett, Howland, Smith, & Newman, 1993). Third, psychopaths often display a paradoxical increase in response speed in reaction to punishment cues, suggesting that they are normally responsive to the arousing, though not the inhibiting, properties of such cues (Amett et al., in press). Finally, psychopaths seem to be less influenced by contextual cues even when they are unrelated to punishment (Newman, in press; Newman et al., in press).

In addition, the response modulation hypothesis suggests an explanation for poor socialization and lack of constraint that is largely independent of reward and punishment sensitivity and, thus, the impulsivity and anxiety/negative emotionality dimensions (see also, Depue & Spoont, 1986; Lynam, 1996). In its general form, the response modulation hypothesis holds that once engaged in effortful goal-directed behavior, psychopaths are relatively unaffected by contextual cues that automatically prime associations and enhance perspective in controls. Such perspective taking may be essential for good judgment, affective depth, impulse control, interpersonal commitment, and self-regulation more generally (Newman, in press; Newman & Wallace, 1993b; D. Shapiro, 1965).

Although the response modulation hypothesis was not derived from Gray's model, the framework helped form specific a priori hypotheses about the nature of the problem and its consequences for self-regulation (Patterson & Newman, 1993). According to Gray (1987), the SHS is instrumental in comparing stimulus inputs from the environment with predictions based on knowledge
of (a) past environmental regularities (acquired via Pavlovian conditioning), (b) the next intended behavior, and (c) the anticipated consequences of such behavior on the environment (inferring from instrumental conditioning). For the most part, this complex integrative process occurs in a relatively automatic fashion. However, "if there is discordance between actual and expected stimuli or if the prototypical stimulus is aversive—conditions that are jointly termed 'mismatch’—the septo-hippocampal system takes direct control over behaviour and now functions in 'control mode'" (Gray, 1987, p. 294). One interpretation of psychopath's laboratory performance is that the (automatic) shift to control mode occurs less readily.

With regard to the neurochemistry underlying these processes, Gray (1987) proposes the following:

Under appropriate conditions, the ascending noradrenergic input to the SHS prompts this into applying its computational powers to the analysis of anxiogenic stimuli while the serotinergic input prompts it to inhibit motor behaviour. At the same time, the locus corensus system increases the activity of a number of other regions widely distributed in the brain (so exercising its general arousal function). (p. 313)

Finally, Gray (1987) notes that ascending serotonergic projections are especially important in mediating behavioral inhibition in response to "high-intensity anxiety" (p. 313); i.e., cues for punishment as opposed to stimuli for nonreward (see also, Spoon, 1992). Given this formulation, one may speculate that the laboratory performance of low-anxious psychopaths is related to the ascending serotonergic projections. Accordingly, cues for punishment would tend to elicit as opposed to passive solutions without necessarily altering the arousal component of their response to punishment cues. Such individuals would also be more likely to emit dominant responses without sufficient processing of the relatively automatic associations primed by "mismatches." Moreover, laboratory evidence of disinhibited responding in psychopaths has been found principally in reaction to punishment cues rather than cues for nonreward1 and while subjects are engaged in task relevant approach behavior (e.g., Newman et al., 1985, 1990, 1992). Although space precludes a more detailed analysis of these findings (e.g., constraint) with general implications for antisocial behavior (see also, Lynam, 1996).

SUMMARY, CONCLUSIONS, AND FUTURE DIRECTIONS

It has been proposed that CNS models are valuable for organizing and integrating findings from diverse literatures in relation to specific processes of relevance to antisocial behavior. This chapter focuses on Gray's model, which has been used increasingly for this purpose. The model provides a broad framework and encompasses a variety of physiological, motivational, learning, and cognitive factors of relevance to antisocial behavior. However, the literatures subsumed by these factors are themselves extensive and complex. It stands to reason, therefore, that a schematic version of Gray's model such as the one commonly applied in personality and psychopathology research (e.g., Figure 30.1) is necessarily superficial. Proper use of the model would seem to require investigators to become conversant with the research literature being represented by the model (e.g., see Gray, 1987) in order to provide a meaningful analysis of the processes being modeled. Gray's model is especially valuable for clarifying the psychological implications of these literatures and highlighting potential interactions among the implied processes, but it cannot, without terrible loss of meaningful detail, substitute for the content of these literatures. To the extent that investigators rely on the hypothetical constructs in Gray's model using broad notions such as approach-avoidance or reward sensitivity—punishment sensitivity, we lose the opportunity to advance our understanding of these complex, multidimensional processes.

These comments notwithstanding, there is evidence that Gray's model has proven useful at the level of schematic-based formulations. Such formulations are useful for demonstrating the applicability of Gray's model (e.g., a dysfunctional BIS could explain many symptoms of psychopathy); for conceptualizing competing hypotheses (e.g., the disinhibited behavior of particular antisocial syndromes may involve exaggerated BAS processes, weak BIS processes, or some combination of the two); for suggesting interactions that may moderate disinhibited behavior (e.g., high anxiety and, presumably, a strong BIS may moderate the expression of conduct disorder in children; Frick, in press; Walker et al., 1991); as well as for identifying psychological processes in need of more detailed analysis (e.g., Does the disinhibited behavior of psychopaths reflect BIS-mediated increases in NAS arousal?). The point emphasized throughout this chapter is that such formulations are valuable to the extent that they clarify the processes contributing to antisocial behavior. Given a literature linking Electrodermal Hyporeactivity to psychopathy, it is pointless to propose that EDA reflects un regulated BIS or that EDA affected by BAS activation.

In conclusion, a thorough understanding of the predisposition to antisocial behavior appears to require integrating findings from the physiological, neuroendocrine, psychophysiological, behavioral, cognitive, emotional, and personality domains. As knowledge increases in each of these domains, so will the need for heuristic
devices to conceptualize these diverse influences and their interactions. CNS models can serve this important function. Gray’s model was used as an example to illustrate this function. The model demonstrates terrific potential for relating diverse influences on antisocial behavior as well as for identifying relatively unique processes contributing to the development of such behavior problems. However, to realize the potential of Gray’s and other CNS models, it is essential that we do not rely hypothetically constructed constructs on lose sight of the complex physiological processes underlying them. The purpose of a CNS model is to increase the depth and specificity of our formulations. Used properly, they should stimulate specific hypotheses about the biopsychological processes underlying antisocial behavior, clarify the distinct biopsychological processes contributing to the antisocial behavior of diverse subgroups, and provide a means of conceptualizing the role of environmental factors in moderating antisocial response inclinations.

REFERENCES


334 Conceptual Models of Antisocial Behavior


