PERGAMON

Personality and Individual Differences 30 (2001) 333–352

PERSONALITY AND
INDIVIDUAL DIFFERENCES

www.elsevier.com/locate/paid

Commentary

# Testing problems in J. A. Gray's personality theory: a commentary on Matthews and Gilliland (1999)

Philip J. Corr

*Department of Psychology, Goldsmiths College, University of London, New Cross, London SE14 6NW, UK*

Received 15 April 1999; received in revised form 4 October 1999

---

## Abstract

In their comparative review of H. J. Eysenck's arousal and J. A. Gray's reinforcement theories of personality, Matthews and Gilliland [(1999) *Personality and Individual Differences, 26*, 583–626] concluded that "Cognitive constructs may be more appropriate than biological ones for explaining the majority of behaviours, so that explanations of the kind offered by the Eysenck and Gray theories are relevant to a restricted range of phenomena only" (p. 620). In this commentary, I address a number of problems in Gray's account of the Behavioural Inhibition System (BIS) and the Behavioural Approach System (BAS) that need to be tackled before we can conclude that cognitive accounts of personality should supersede, rather than complement, biologically-based ones. I focus on two broad sets of issues. (1) Practical (operational) problems, that call for further experimental work, relating to: (a) BIS/BAS and conditioning/learning; (b) reinforcement parameters (i.e., defining reward/punishment, reinforcement expectancies, schedules of reward, sources of reinforcement, cues versus feedback); and (c) psychometric measures of BIS/BAS functioning. (2) The theoretical nature of BIS/BAS effects, that call for an elaboration of Gray's theory. I propose a two-process model that postulates that the BIS and BAS exert two effects: *facilitatory*, the BIS mediates responses to aversive stimuli, the BAS to appetitive stimuli; and *antagonistic*, the BIS and BAS impair responses mediated by the alternate reinforcement system. Specific directions for future research are given. I conclude that, in several crucial respects, Gray's reinforcement theory has yet to be adequately tested; and that closer attention to operational definitions, and the mutual interplay of the BIS and BAS, may enhance its experimental precision. © 2000 Elsevier Science Ltd. All rights reserved.

*Keywords:* Behavioural inhibition system (BIS); Behavioural Approach System (BAS); Reinforcement; Expectancies; EPQ; Impulsivity; Anxiety; Extraversion; Neuroticism

---

*E-mail address:* p.corr@gold.ac.uk (P.J. Corr).

PII: S0191-8869(00)00028-3

## 1. Introduction

Matthews and Gilliland's (1999) comparative review of the empirical status of H. J. Eysenck's Arousal Theory (AT) and J. A. Gray's Reinforcement Sensitivity Theory (RST) raised a number of important issues relating to the biological approach to personality. In this commentary I address a number of issues relating to RST, *viz* (1) problems with operational definitions, and (2) theoretical elements that are in need of clarification and perhaps reformulation. I argue that these issues need to be addressed before we are compelled to conclude that "Cognitive constructs may be more appropriate than biological ones for explaining the majority of behaviours, so that explanations of the kind offered ... are relevant to a restricted range of phenomena only" (p. 620). I contend that there are a number of hidden complexities in Gray's general model of RST that, at present, hinder the derivation of precise experimental hypotheses necessary for the test of the theory's specific claims in human beings; and that the range of phenomena to which RST is applicable is much broader than assumed by Matthews and Gilliland (1999). I will present several theoretical elaborations to RST that may help to close the gap between experimental prediction and data.

Gray's RST consists of three systems of emotion that underlie motivated behaviour. The first is the *behavioural inhibition system* (BIS; Gray, 1976, 1982), which is sensitive to *conditioned* aversive stimuli (i.e., signals of both punishment and the omission/termination of reward) extreme novelty, high intensity stimuli, and innate fear stimuli (e.g., snakes, blood). The BIS is proposed as the causal basis of anxiety (Anx), which, in terms of Eysenck's personality space, ranges from E−/N+ (Anx+) to E+/N− (Anx−) (Anx+ is rotated by 30° from N; Gray, 1970; Pickering, Corr & Gray, 1999). The second negative emotionality system is the *fight/flight system* (FFS; Gray, 1987), which is sensitive to *unconditioned* aversive stimuli (i.e., innately painful stimuli), which mediates the emotions of rage and panic. Gray aligns the FFS with Eysenck's psychoticism (P; Eysenck & Eysenck, 1976) factor. The third system, which is responsible for positive emotionality, the *behavioural approach system* (BAS; Gray, 1987), is sensitive to *conditioned* appetitive stimuli (i.e., signals of both reward and the termination/omission of punishment). The BAS mediates impulsivity (Imp), which ranges from E+/N+ (Imp+) to E−/N− (Imp−) (Imp+ is rotated 30 degrees from E; Gray, 1970; Pickering et al., 1999). There is not a comparative system to the FFS that mediates *unconditioned* appetitive stimuli.

In broad terms, RST predicts that Imp+ (strong BAS) individuals should be most sensitive to signals of reward, relative to Imp− (weak BAS) individuals; and Anx+ (strong BIS) individuals should be most sensitive to signals of punishment, relative to Anx− (weak BIS) individuals. The orthogonality of the BIS and BAS suggests that (1) responses to reward should be the same at all levels of BIS/Anx, and (2) responses to punishment should be the same at all levels of BAS/Imp. (The most recent elaboration of this theory is found in Pickering et al., 1997.)

In the discussion that follows, Sections 2 to 4 focus on operational problems, which hinder the construction of effective experimental tests of RST's core postulates, and thus impair our ability to evaluate its true scientific value. These issues are raised in the hope that they may encourage refinement to experimental methodologies (however, at this time, it is possible only to provide tentative suggestions as to the direction this research should take). In Section 5, I propose a definite elaboration of RST which, while it is consistent with the theoretical foundations of Gray's general model, points to new lines of investigation.

## 2. BIS/BAS and conditioning/learning

In addressing the range of issues raised by Matthews and Gilliland, I will first discuss the relevance of RST to Pavlovian conditioning, and then to more general forms of learning that depend upon (BIS/BAS) motivational influences.

### 2.1. *Pavlovian conditioning*

Matthews and Gilliland (1999, p. 603) raised the question of the applicability of RST to Pavlovian conditioning, a form of learning that has long held a privileged position in the experimental test of biological models of personality.

Gray's RST is explicit about the nature of stimuli to which the BIS and BAS are sensitive: they are sensitive to *conditioned*, not *unconditioned*, stimuli. Therefore, as discussed by Matthews and Gilliland, the BIS and BAS should not be involved in the Pavlovian conditioning of conditioned stimuli (CSs). Given its role in mediating reactions to unconditioned aversive stimuli, the FFS (P) might be expected to influence Pavlovian aversive conditioning, at least insofar as strong (FFS) reactions to unconditioned aversive stimuli facilitate stimulus salience and intensity related associability. Thus, Pavlovian conditioning studies of BIS/BAS effects might appear to be inappropriate for testing RST; and, indeed, significant results could be seen to represent a decisive refutation of the theoretical bases of the BIS and BAS. In contrast to Pavlovian conditioning, reactions to aversive/appetitive CSs on well-learned instrumental behaviour would appear to represent the ideal test bed for RST's predictions (see Pickering et al., 1997).

Now if RST relates only to the invigoration or suppression of already well-learned instrumental behaviour, and not to conditioning or learning in general, then it rather obviously has limited application; it would be difficult to argue that it could replace Eysenck's arousal theory that has a broad range of application (including Pavlovian conditioning), supported by impressive empirical data. If the BIS/BAS are not related to Pavlovian conditioning, then RST would have the difficult task of explaining why Pavlovian conditioning and personality (i.e., E and Imp) *are* empirically related (e.g., Eysenck & Levey, 1972): if the BIS and BAS are proposed as alternative causal explanations of E and N (Gray, 1970), then why should E (and sometimes N and P) relate to Pavlovian conditioning?

We are thus presented with what appears to be a glaring contradiction: in agreeing with RST's claim that the BIS/BAS do not relate to Pavlovian conditioning, and therefore Pavlovian conditioning studies are not germane to testing RST's postulates, we are faced by a number of such studies that have been used both to test and support Gray's RST, including data from Gray's own laboratory (e.g., Corr, Pickering & Gray, 1995a). However, the issue of BIS/BAS effects in Pavlovian conditioning is more complex than it may appear at first sight.

#### 2.1.1. *UCS/CS distinction*

Close inspection of what passes for Pavlovian conditioning in human studies highlights a number of important factors. For example, as discussed by Matthews and Gilliland (1999), Corr et al. (1995a) used a two-process learning paradigm in which *S*s first acquired, by Pavlovian (associative) conditioning, CSs for reward and punishment, and then, in the second phase of the experiment, these CSs signalled reinforcement contingent upon instrumental behaviour. The results for the associative phase revealed effects of Cloninger's (1986) harm avoidance (HA, in the case of aversive UCS), and reward dependence (RD, in the case of appetitive UCS); in the instrumental phase, Imp and Anx moderated reactions to contingent reward and punishment. But in this study, Pavlovian conditioning entailed *second-order* learning in which the "UCS" was itself a conditioned stimulus (i.e., money; the CS were coloured lines). (In second-order conditioning, the term UCS is used to denote one stimulus that already has the power to elicit a response, unlike the CS which has to acquire this power by its association with the UCS: the UCS is not *innately* aversive or appetitive.)

In the associative phase of the Corr et al. (1995a) study, a trial comprised (1) *S*s being shown a coloured line (CS), then (2) they were required to press one of three buttons to indicate whether a (UCS) counter on the screen would move left (aversive, lose money), right (appetitive, win money), or not move at all (neutral, no monetary effects). The rationale for predicting a BIS/BAS effect in this associative learning phase was that: after the *S* had made a response, to indicate which CS-colour was predictive of which UCS, and once the UCS had been presented, the associative strength of this CS-UCS link (still in working memory) would be influenced by the UCS (i.e., winning/losing of money), by virtue of the UCS serving as adequate input into the BIS or the BAS and thus altering arousal and emotion states. For example, *S*s sensitive to aversive stimuli were expected to show a stronger output of the BIS (arousal and negative affect) that should serve to strengthen any stimulus–stimulus links currently undergoing processing; a comparable case was made for appetitive stimuli and the BAS. Thus, the Corr et al. (1995a) study is not relevant to the issue of BIS/BAS effects in first-order Pavlovian conditioning. However, this study is germane to the issue of the relationship between personality and associative learning of pre-existing conditioned stimuli (we may assume that such learning is pervasive in human beings).

However, a number of other studies have shown that personality is related to first-order Pavlovian conditioning. Do these studies refute RST's claims? In a classic series of Pavlovian conditioning studies in Eysenck's laboratory (e.g., Eysenck & Levey, 1972), a puff of air to the cornea served as the UCS, and light or tone served as the CS. Consistent effects of E and Imp were observed. Should we here, where the UCS is truly innate (i.e., it elicits a physiologically-determined reflex, the eyeblink), expect BIS/BAS effects? In an ideal experimental situation, RST does not predict BIS/BAS effects; but in a typical eyeblink conditioning experiment there are an array of potential ambient CSs, that vary along valence and arousal dimensions (e.g., experiments are often run in potentially aversive environments, e.g., hospital settings). These stimuli alone may be sufficient to activate the BIS/BAS, the effects of which might be superimposed upon first-order conditioning effects. But even assuming that human *S*s are not reacting to these CSs, the fact of being told the nature of the experiment (i.e., puffs of air to eye) and being wired-up to laboratory equipment may lead to rapid (perhaps one-trial) Pavlovian conditioning to environmental stimuli, which then serve as CSs adequate to activate the BIS/BAS; activation of the BIS/BAS might be expected to exert a generalised emotional influence on learning and thereby influence the Pavlovian conditioned response of interest in the experiment.

Therefore, we cannot state with any confidence that what we are observing in a human Pavlovian conditioning experiment is truly first-order Pavlovian conditioning; and we have good reasons for assuming that much of so-called Pavlovian conditioning variance is BIS/BAS mediated. Thus the claim that the BIS/BAS are not involved in Pavlovian conditioning must be treated with suspicion, as must the claim that personality effects on first-order Pavlovian conditioning necessarily refute the cardinal postulates of RST.

In order to address this issue with greater rigour, we would need to employ an experimental procedure involving a physiological response system that was preferentially sensitive to first but not second-order Pavlovian conditioning (i.e., a conditioned response that was not affected by generalised emotional influences); and then we could test RST's claim that the BIS/BAS are not involved in first-order Pavlovian conditioning. This experiment could utilise a two-process learning procedure (Mowrer, 1960), that first established appetitive and aversive CSs by truly first-order Pavlovian procedures and then presented these CSs in a second instrumental phase of the experiment, during which (putatively BIS/BAS-mediated) passive avoidance and approach behaviour could be measured (see Pickering et al., 1997).

### *2.2. BIS/BAS and learning*

As with Pavlovian conditioning, according to RST, learning is not mediated by the BIS/BAS; but learning is influenced by virtue of the fact that many types of psychological processes are affected by general motivational factors, such as arousal and emotion. To illustrate this point, consider the case of procedural learning that colleagues and I have employed to test both Eysenck's arousal theory and Gray's RST.

Procedural learning is a phylogenetically old form of knowledge acquisition involving the learning of stimulus–stimulus regularities (i.e., frequency, or covariation, information). The procedural learning task used in our series of studies consisted of a long series of target movements (between four locations on a computer screen); some of these movements were random, while others were predictable. Procedural learning was demonstrated by the progressive decline in RT to predictable trials as compared with RTs to random trials. Most *S*s showed robust procedural learning without awareness of the procedural rules (Corr, 1994; Corr, Pickering & Gray, 1997b). This task is of theoretical interest because it may well represent a ubiquitous form of learning implicated, to some degree, in many diverse forms of human performance. If biological theories of personality fail to relate to this type of learning then we would indeed be justified in concluding that these theories relate to a rather restricted range of phenomena, as claimed by Matthews and Gilliland (1999).

But should we expect RST, or Eysenck's arousal theory, to relate to procedural learning? Indeed, can we make any sensible predictions at all, especially as the rules comprising procedural learning are largely nonconscious?

Despite claims that procedural learning should show few individual differences, by virtue of its phylogenetic origins, we have found that it is sensitive to arousal (caffeine; Corr, Pickering & Gray, 1995b), de-arousal (haloperidol; Corr & Kumari, 1997), both of which were moderated by E (Soc), in accordance with Eysenck's arousal hypothesis; and punishment×Anx (Corr et al., 1997b), in accordance with the specific predictions of RST, and to a reward×Imp interaction (Corr, 2000), in accordance with the general postulates of RST. (Such data suggest that Gray's RST and Eysenck's arousal theory are not mutually exclusive, but complementary accounts, each relating to different causal systems.)

These effects on procedural learning are not caused by behavioural activation or inhibition (i.e., the specific behavioural outputs of the BIS/BAS), because effects on overall RT are not found. Rather, they seem caused by the induction of arousal and emotion. We could speculate that few, if any, cognitive processes escape fully the influence of arousal and emotion; therefore, we may wish to conclude that the theoretical constructs central to both RST, and Eysenck's arousal theory, have profound effects across the whole cognitive-behavioural landscape. Thus, we may be premature to conclude that biological theories are "... relevant to a restricted range of phenomena only" (Matthews & Gilliland, 1999, p. 620).

RST has perhaps focused too much on prototypical animal behaviours (*viz*, passive avoidance and approach behaviour), leading to the impression that, when these behaviours are not possible (or measured in the experiment), BIS/BAS do not exert general motivational effects that impact upon many diverse forms of learning and performance. Attention to this issue in future revisions of RST might bring the model into closer agreement with standard cognitive accounts of the relationship of personality to performance.

## 3. Reinforcement parameters

A major problem with the evaluation of RST is the imprecise operational definition of reinforcement. Much of the disconfirmatory (and confirmatory) evidence reviewed by Matthews and Gilliland (1999) suffered from this problem. This lack of precision may underlie the diversity of findings that, while tending to support the general postulates of RST, fail to support the specific details of the model. A number of relevant issues, relating to the manipulation of reinforcement, are discussed below. These issues highlight the difficulty of evaluating Gray's RST, as well as the challenges facing future research. Where feasible, recommendations are given to guide research.

### *3.1. Defining reward and punishment*

What constitutes reward and punishment? That is, how do we define, in operational terms amenable to experimental manipulation and verification, the types of stimuli that activate the BIS and BAS; and how do we know the degree of activation induced by these stimuli? At present, we do not have a proper metric for the calculation of reinforcement parameters.

In the nature of a developing research programme, the operationalisation and manipulation of reward and punishment in previous RST studies has been rather hit-and-miss. Unlike rat experiments, where environments can be tightly controlled, we cannot, with great confidence, determine which stimuli are uniquely rewarding or punishing to human beings. Accordingly, it should perhaps come as no surprise that the effects of reinforcement are so varied.

To illustrate this point, I (Corr, 2000) found that, using monetary incentives, reward-sensitive Imp + *S*s (who were also Anx + ) responded to reward (versus non-reward control) with a marked reduction in procedural learning, which was accompanied by a reduction in hedonic tone and an increase in state anxiety. These effects, in apparent opposition to RST, appeared to reflect frustrative nonreward (FN; i.e., an aversive reaction when the actual level of reward is less than the expected level of reward; see Gray, 1987). However, upon close theoretical analysis, these apparently paradoxical effects are consistent with the general postulates of RST. We return to this issue below in the context of reinforcement expectancies.

A second recent example, relating to the induction of emotional states by reinforcing visual stimuli, comes from affective modulation of the eyeblink startle reflex by filmclips. Despite confirmation of the association of trait anxiety and aversive stimuli when using slide material (e.g., Corr et al., 1995c, 1997a), attempts to replicate these effects using filmclips proved negative (Kumari et al., 1996). Following recent neuroimaging studies, showing an anatomical dissociation of reactions to fear and disgust-inducing stimuli (Phillips et al., 1998, 1997), we reanalysed these filmclip data and found that *S*s scoring high on the Eysencks' neuroticism scale showed potentiated startle reflexes to fear-inducing filmclips only; to disgust-inducing filmclips they showed reduced startle reflexes (Wilson, Kumari, Gray & Corr, 2000). Such data point to the importance of clearly differentiating stimuli capable of inducing different negative emotional states.

In future RST studies, describing stimuli merely in terms of appetitive (reward) and aversive (punishment) valence may not be adequate. Reinforcing stimuli need to be carefully delineated in terms of their emotional components; this may have the benefit of revealing which emotional states are the product of activation of the BIS or BAS alone, and which are the product of the influence of both systems. Improved operational procedures may greatly enhance the relevance of Gray's RST to the true complexity of the emotions.

Given this serious weakness in our experimental procedures, to quote such data as disconfirmation of RST is premature — although, as Matthews and Gilliland rightly point out, these data do properly imply that RST has been weakly specified with respect to human beings and is in need of more rigorous delineation. Specifically, such data pose a challenge to RST researchers to characterise, in advance of data collection, reinforcement parameters that permit precise experimental predictions. Matthews and Gilliland (1999) were indeed correct to point to the embarrassing diversity of RST data, little of which can be cited as strong support of the simple version of RST.

### *3.2. Reward/punishment expectancies: comparator values*

Related to the definition of reward and punishment is the issue of reward/punishment expectancies brought to the experimental setting by *S*s. People volunteer to participate in psychological studies for a variety of reasons: they may be unemployed and hoping to earn money, in which case their expectation of the "going rate" for laboratory work may be calibrated against several reference sources (e.g., minimum hourly rate, incurred costs of travel, inconvenience, and compensation for effort); or they may be interested in psychology and what psychologists get up to in their laboratories; or they may participate to earn course credits. In any event, during the testing session, *S*s performance may be influenced by a variety of subjective factors (e.g., ego-enhancement/protection, experimenter approval, and perceived control). The experimenter can only hope that these factors do not become significant sources of systematic error. The association of such motivating factors and BIS/BAS-related personality factors is unknown; but their potential for compounding data is great.

The influence of reinforcement expectancies was first delineated by Gray and Smith (1969). Their mathematical model postulated that actual reinforcement value is compared against expected reinforcement value, and it is the difference between these two values that determine (1) whether an experimenter-defined reward/punishment is perceived as such by the *S*, and (2) the direction and strength of the reaction to reinforcement. For example, reward sensitive Imp + *S*s might expect more reward than the experiment actually delivers and thus perceive the experiment as somewhat aversive; whereas punishment-sensitive Anx + *S*s may be pleasantly relieved by the experiment, perceiving it as relatively appetitive (for a discussion of comparator values that set reward/punishment expectancies, see Corr, submitted). It is somewhat surprising that this important aspect of RST has received so little attention in human experimental studies.

The above discussion may throw light on Matthews and Gilliland's (1999, p. 607) conclusion that "It is disturbing that the well-designed and thorough discrimination learning studies of Zinbarg and Revelle (1989) and of Corr et al. (1995a) give entirely different results" to findings "... within more 'traditional' conditioning paradigms, notably the eyeblink and verbal conditioning paradigms". We should only expect consistent findings if reinforcement parameters are invariant across studies.

In order to remedy this state of affairs, in future experimental studies two separate components of aversive/appetitive motivation need to be delineated: (1) general (trait) BIS/BAS sensitivities (e.g., Anx and Imp); and (2) specific (state) expectancies of reward and punishment. With respect to (2), it would be appropriate to take separate measures of (subject-defined) expectancies and (experimenter-defined) reinforcement values in order to estimate the subject-perceived value of manipulated reinforcement.

### *3.3. Schedules of reward*

The manipulation of reward in RST studies is problematic for other reasons. Imp seems more related to the frequency (Sonugabarke, Taylor, Sembi & Smith, 1992), or lack of delay (Schweitzer & Sulzerazaroff, 1995), rather than level of reward per se; also, Imp + *S*s steeply discount the future (i.e., temporal discounting; Ostaszewski, 1996), which further weakens the positive effects of delayed reward. This intolerance of reward delay is associated with negative emotional states — the antithesis and possible antagonising of positive motivational states — which can lead to aggression and violence (Evenden & Ryan, 1996). Thus, BAS/Imp positive relations may be evident only under frequent reinforcers without significant delays. In future studies of RST, it would be desirable to manipulate these parameters of reward in order to examine more fully the effects of Imp.

### *3.4. Source of reinforcement*

The source of reinforcement is yet another relevant variable. For example, Gupta (1976) found that E + *S*s showed superior verbal conditioning with positive reinforcement only with female-delivered reinforcement. In support of the conclusion concerning the salience of experimenter-delivered reinforcement, when a buzzer was substituted for verbal reinforcement E + /N– (i.e., Anx–) not E + /N + (i.e., Imp + ) was associated with superior conditioning (Nagpal & Gupta, 1979); similar Anx– findings, with weak appetitive stimuli, are commonly reported (e.g., Corr et al., 1995a, c; Larsen & Katelaar, 1991).

The positive findings from Gupta's group in India, showing that, in addition to E + superiority with positive reinforcement (Gupta & Nagpal, 1978), reactions to positive reinforcement are enhanced in N + *S*s (Nagpal & Gupta, 1979), is in good agreement with Gray's RST. But few other laboratories have been so successful with this verbal operant conditioning paradigm (e.g., Hernaiz, 1991). One reason for this lack of replication may relate to the source of reinforcement. As discussed by Corr (1994), university students in India, who comprised Gupta's samples, are highly deferential and respectful of their professors, so positive reinforcement in the form of the verbal feedback "Good" should be expected to be highly salient in this population; but we should not necessarily expect the same motivational reactions in student populations in other countries. Clearly, this is an issue that demands closer attention in future reinforcement studies.

### 3.5. Cue of reinforcement versus feedback

Matthews and Gilliland (1999, p. 611) stated that "... Gray's theory does not provide a rationale for the moderating role of feedback, which is an actual outcome rather than a signal of reinforcement". For example, Gupta's work showed that E+/N+*S*s *are* sensitive to positive reinforcement in the form of verbal feedback. However, the distinction between cues of reinforcement and actual reinforcement may be much less important than first imagined. Feedback will itself lead to associative learning: it could be seen as the UCS, with salient environmental stimuli, including the experimenter, serving as CSs adequate to activate the BIS/BAS from very early stages in the experiment. Although these cues would not be discriminative, they would nevertheless activate the BIS/BAS and thus influence any form of learning or performance that is sensitive to the induction of emotion/arousal.

Although we could develop (psychophysiological) measures that were differentially sensitive to cues of reinforcement, as distinct from reinforcement per se, in order to address this issue with greater experimental rigour, we may prefer to accept this ambiguity in liberalising our view of the types of stimuli that, in human beings, activate the BIS and BAS. The more important issue is the consequences of the activation of the BIS and BAS: this is RST's pulsating heart, where the dagger of experimental refutation must be aimed.

## 4. Personality measures of BIS/BAS functions

Another issue that calls for further experimental clarification concerns the psychometric measurement of BIS/BAS functions. Many studies have used the Eysencks' E & N dimensions, some preferring the Eysenck Personality Inventory (EPI; Eysenck & Eysenck, 1964), with its two component measures of E (Soc and Imp), while others have preferred the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975), in which Imp items have been largely removed from E. There have also been many attempts to use more direct measures of the BIS/BAS; for example, Spielberger's trait anxiety, Barratt's impulsivity scale, or the impulsivenesness scale from Eysenck Personality Scales (EPS; Eysenck & Eysenck, 1991). In addition, a range of BIS/BAS-inspired measures have been developed in recent years: for example, Torrubia and Tobena's (1984) sensitivity to punishment scale; Wilson, Barrett and Gray's (1989) Gray–Wilson Personality Scales (which measures six prototypical learning paradigms); Ball and Zuckerman's (1990) Generalized Reward and Punishment Expectancy Scales (GRAPES); MacAndrew and Steele's (1991) BIS scale; and Carver and White's (1994) Behavioral Inhibition and Activation Systems (BIS/BAS) Scales. In addition, BIS/BAS scales from alternative biological models of personality have been used (e.g., Cloninger, 1986; Tridimensional Personality Questionnaire, that measures harm avoidance, HA; reward dependence, RD; novelty seeking, NS; and persistence, P).

No one set of measures has turned out to be superior in terms of predictive value. For example, in two studies (in which the same measures were taken), colleagues and I have found that Cloninger's HA predicted modulated startle reflex by unpleasant slides (Corr et al., 1995c, 1997a); and in a different study (Inoue, French, Pring & Corr, 2000) we found that Spielberger's trait anxiety measure (Spielberger, Gorsuch, Lushene, Vagg & Jacobs, 1983) predicted the same aversive modulation. In the case of appetitive motivation, Zinbarg and Mohlman (1998) found that Carver and White's (1994) BAS Reward Responsivity scale, and not (EPI) impulsivity, was related to the speed of acquisition of reward expectancies (for a survey of the effects of different BIS/BAS measures, see Pickering et al., 1997).

Personality scales have a number of potential problems that may explain the inconsistency of results observed across studies. First, they are prone to a range of extraneous influences (e.g., response sets, and social desirability); secondly, they may reflect the interplay of a number of lower-order, relatively independent sources of variance (thus, they may conflate rather than clarify fundamental causal processes); and thirdly, and more fundamentally, human beings may be incompetent in their linguistic expression of positive and negative motivation. However, the fact that existing personality scales do relate to basic neuropsychological processes (e.g., E and cortical arousal), albeit often in unpredictable ways, suggests that pencil-and-paper personality measures do have a role to play in biological accounts of personality. However, we must bear in mind Cloninger's (1986) assertion that the phenotypic expression of personality, as revealed by exploratory factor analysis, may not closely correspond to the genotype structure of personality, as represented by the BIS/BAS. Therefore, it may be necessary to construct scales on a rational basis to capture better underlying causal influence.

In any event, there is little consensus over the optimal set of BIS/BAS psychometric measures, and there is no rational basis upon which to choose between alternative scales. As few studies have used all measures, it is difficult to arrive at a consensus concerning their respective merits. It is therefore difficult to evaluate these studies in terms of the predictions of RST.

### *4.1. E/N-derived BIS/BAS factors*

It is possible to extract Anx and Imp factors directly from the Eysencks' PEN scales. As Anx is rotated 30° from N, IMP 30° from E, and Imp inclines into P, Anx away from P, EPQ-derived Anx and Imp may be derived thus: $\text{Anx} = ((21-\text{E}) + (\text{N}\times 2) - \text{P})$, $\text{Imp} = ((\text{E}\times 2) + \text{N} + \text{P})$. Corr (submitted) compared the predictive power of (EPS) Impulsiveness and EPQ-derived Imp in terms of the induction of frustrative nonreward (FN) on procedural learning: only EPQ-derived Imp(+) and Anx(+) moderated FN; specific measures of (EPS) Impulsiveness and (Spielberger's) trait Anx did not. Thus, it may not be appropriate to consider specific measures of Imp and Anx as comparable to EPQ-rotated Imp and Anx factors. According to Gray (1970), E and N should be rotated to form Imp and Anx, but these EPQ-derived Imp and Anx factors do not seem to moderate behaviour in an identical manner to specific measures of Imp and Anx. Such differences may help to explain the diversity of results reported, some more conducive to RST's postulates than others. Thus, published studies to date may provide only a partial picture of the true scientific value of RST.

To complicate matters still further, there may be important sex differences in Imp and Anx measures. Diaz and Pickering (1993) reported that questionnaire measures of Imp were predicted by differently weighted combinations of E, N and P for males and females. It is unclear whether such sex differences represent fundamental differences in underlying biological systems, or some form of psychometric artefact (e.g., response bias). In any case, such data suggest that putative sex effects should be closely examined in future RST studies.

Clearly more work is needed to provide adequate psychometric measures of BIS/BAS functioning. It would be desirable for RST studies to use a wide range of measures to allow comparison across studies. However, it is also possible that existing BIS/BAS scales poorly measure BIS/BAS functioning. But where else might we look for psychometric measures of BIS/BAS functions?

### *4.2. BIS/BAS scales: an empirical approach*

Personality research is still concerned with the venerable question of whether a larger number of lower-order factors is preferable to a smaller number of higher-order dimensions. This concern may be especially pertinent for the BIS and BAS. Depue and Collins (1999) made a strong case for associating E with positive incentive motivation, and in their theoretical exposition they identified a number of behavioural processes implicated in incentive motivation that may have distinct neurological substrates. Therefore, seeking a single psychometric measure of a motivational tendency may be inappropriate. There could exist, for example, traits that moderate common elements of both positive and negative motivation; for example, the initial processing of stimulus salience/significance, irrespective of emotional valence, which may share a common neurology in nucleus accumbens dopamine functions (Gray, Kumari, Lawrence & Young, 1999). This putative trait may be quite distinct from traits relating to response organisation and execution (e.g., approach in the case of positive incentive motivation, passive avoidance in negative incentive motivation).

This alternative conceptualisation of motivational traits suggests that research should: (1) delineate the processes involved in positive and negative incentive motivation (e.g., stimulus identification, emotional evaluation, response organisation, etc.); (2) determine the extent of shared processing (e.g., stimulus salience); (3) determine the unique processes (e.g., behavioural facilitation and inhibition); and only then (4) identify emotional, motivational and behavioural traits that correspond to these shared and unique processes. If adequate behavioural indices of these processes could be developed, then item analysis techniques could be used to construct scales that correspond directly to these underlying conceptual processes. It might be valuable to pursue this strategy alongside the search for optimal relations with existing scales of reinforcement sensitivity.

## 5. BIS/BAS effects: theoretical clarification

In this section, I turn attention from operational issues to focus on a substantive theoretical matter. I argue that many predictive failures of RST may be traced to the hypothesized functional independence of BIS/BAS effects (this line of argument follows closely that of Pickering, 1997).

Now nearly all RST-inspired studies have adopted the view that Imp+ (strong BAS) *S*s should be most sensitive to signals of reward, relative to Imp− (weak BAS) *S*s; and Anx+ (strong BIS) *S*s should be most sensitive to signals of punishment, relative to Anx− (weak BIS) *S*s. In addition it is supposed that (1) responses to reward should be the same at all levels of BIS/Anx, and (2) responses to punishment should be the same at all levels of BAS/Imp. That is, BIS (Anx) and BAS (Imp) effects should be independent. The literature fails to support these specific claims. Although these data may be cited in refutation of RST, they may simply reflect an incomplete characterisation of BIS/BAS effects.

In order to address this complex issue, I first describe two positions with respect to BIS/BAS effects: the first that assumes BIS/BAS functional independence (the *separable subsystems* hypothesis); the second that assumes inhibitory functional effects (the *joint subsystems* hypothesis). Next, I present the theoretical background to the *joint subsystems* hypothesis before evaluating each position with respect to data from human experimental studies. Finally, I describe a two-process model of BIS/BAS, which attempts to account for the mutual effects of BIS/BAS influences and which points to new lines of investigation.

### 5.1. Separable subsystems

Consistent with Gray's own theoretical claims, most RST studies have assumed separate effects of anxiety (BIS) and impulsivity (BAS) (Pickering et al., 1997). This hypothesis states that "... individual differences in the functional capacity of one system are independent of the individual differences in the functional capacity of the other system" (Pickering, 1997, p. 145). This *separable subsystems* hypothesis is most tenable under reinforcement conditions that contain neither mixed reward and punishment cues, nor demand rapid attentional and behavioural shifts between these two sets of cues. This view of BIS/BAS effects assumes that a decision mechanism determines only one dominant behavioural/affective *state* (Gray & Smith, 1969). This assumption, in part, may be traced to the behaviour observed in approach–avoidance conflict situations, in which one system tends to come to dominate over the other system, thus leading to the assumption that, at any one moment, either the BIS or the BAS is in exclusive control of the choice and execution of behaviour. According to Gray (1987, p. 180), "Conflict between the reward and punishment mechanism is resolved in the decision mechanism according to whichever input to this mechanism is stronger; the reciprocally inhibitory links between the reward and punishment mechanisms ensure a stable outcome to such conflicts."

It is highly doubtful that, in typical human experimental situations, contingencies of reinforcement are powerful enough totally to disengage functioning of either the BIS or the BAS. Thus, we must, at the very least, be in doubt concerning the functional independence of the BIS and the BAS.

### 5.2. Joint subsystems

Also consistent with the general postulates of RST is the possibility that, given a background (non-zero) level of BIS/BAS activation, Anx (BIS) and Imp (BAS) may exert functionally interdependent effects. According to this *joint subsystems* hypothesis: (1) *state* measures of appetitive responses and positive emotion should be highest in Imp+ (BAS+) and Anx− (BIS−) individuals;

and (2) *state* measures of aversive responses and negative emotion should be highest in Anx+ (BIS+) and Imp− (BAS−) individuals. That is, aversive and appetitive motivation is influenced by *both* the BIS and BAS, and that an algebraic subtraction of BIS/BAS activation values is required to derive state measures of emotion. In this respect, *state* measures of current emotion might be best conceptualised along two dimensions: pleasantness versus unpleasantness and arousal (reflecting the moment-by-moment functional interdependence of the BIS/BAS); in contrast, *trait* measures of BIS/BAS functioning may be best conceptualised by the alternative classification of positive affect and negative affect (reflecting the separable effects of BIS/BAS viewed over the longer time frame).

### *5.2.1. Theoretical rationale for joint subsystems hypothesis*

The notion of an inhibitory influence of the BIS on the BAS was first suggested by experiments which revealed that antagonising the BIS can lead to enhanced BAS functioning. For example, barbiturate drugs that impair BIS-mediated responses (Barry, Wagner & Miller, 1962; Gray, 1967, 1969; Ison & Pennes, 1969; Miller, 1959, 1964; Wagner, 1963) have sometimes been found to enhance BAS-mediated approach responses (e.g., Kamano, Martin & Powell, 1966), suggesting that during normal operation, the BIS inhibits the BAS. These releasing effects of anxiolytics on BAS responses are consistent with subjective reports from patients that benzodiazepines, in addition to reducing anxiety, also have a stimulant effect sometimes resulting in elevated positive affect (Thayer, 1989).

In addition to drug effects, additional evidence for reciprocal effects came from lesion of the septo-hippocampal system, which impairs BIS functioning in a manner similar to barbiturates (Gray, 1982), and seems capable of enhancing BAS functioning, as seen in two-way active avoidance (Green, Beatty & Schwartzbaum, 1967; Isaacson, Douglas & Moore, 1961; Olton & Isaacson, 1968; Rabe & Haddad, 1969). Similarly, lesion to the frontal cortex, the highest level of the septo-hippocampal system, enhances BAS-mediated active avoidance behaviour (Albert & Bignami, 1968). In addition, intracranial self-stimulation studies (Olds & Fobes, 1981) reveal both the neuroanatomical overlap in appetitive and aversive systems and the existence of reciprocal inhibition.

Although there is evidence that presentation of appetitive stimuli can reduce the influence of aversive stimuli (e.g., avoidance behaviour; Grossen, Kostansek & Bolles, 1969), it is clear that the antagonising effects of appetitive motivation on aversive motivation are relatively weak. In contrast to appetitive motivation, aversive motivation is easy to establish and difficult to extinguish. In addition, the link from the BIS to the BAS is stronger than that from the BAS to the BIS. This conclusion is based upon a wealth of experimental data which shows that it is easier to transform an appetitive-CS (e.g., CS paired with food) into an aversive-CS (e.g., CS paired with shock) than vice versa (Konorski, 1967; Konorski & Szwejkowska, 1952; see Mackintosh, 1983, for a review of the behavioural evidence relating to appetitive and aversive systems inhibition).

Antagonising, or mutually opposing, effects are seen in part-approach behaviour in the rat in the classic approach–avoidance conflict situation (Miller, 1959; Gray, 1987), where the rat is appetitively motivated to approach the goal box for a reward but fears approaching the box because of a prior negative experience (e.g., foot shock). The degree of actual (BAS) approach (state) behaviour is as much a function of the strength of (BIS) avoidance motivation as it is of (BAS) approach motivation: accordingly, actual approach behaviour may be strongest in BAS+/BIS− (couched in human terms, in Imp+/Anx−) individuals.

### 5.3. Experimental evidence

It is relatively common to find that Anx + is related to individual differences in aversive reactions, a finding that is consistent with the view that it is comparatively easy to induce aversive emotion. For example, the finding that potentiation of the startle reflex by unpleasant stimuli is predicted by trait anxiety is robust (e.g., Cook, Hawk, Davis & Stevenson, 1991; Corr et al., 1995c, 1997a). But it is less common to find that Imp + is related to individual differences in appetitive reactions (Matthews & Gilliland, 1999; Pickering et al., 1997; Zinbarg & Mohlman, 1998); in these cases, often Anx− moderates appetitive reactions, whether assessed by modulation of the eyeblink startle response (e.g., Corr et al., 1995c), induced positive emotion (Larsen & Katelaar, 1991), instrumental approach behaviour (Corr et al., 1995a), or appetitive classical conditioning (Mangan, 1978; Paisey & Mangan, 1988). Matthews and Gilliland (1999, Table 1, p. 600) showed that hedonic tone is consistently related to Anx− (i.e., E+/N−), not Imp+ (i.e., E+/N+).

In addition, sometimes highly complex Anx×Imp interactions are found (e.g., Zinbarg & Mohlman, 1998; Zinbarg & Revelle, 1989). For example, Barratt (1971), using EEG, found that Imp+/Anx− *S*s were less aroused at the moment of stimulus presentation, and emitted fewer classically conditioned responses, while Imp−/Anx+ *S*s emitted the highest number of conditioned responses. Interpretation of such data necessarily falls back on post hoc explanation: the theoretical elegance of RST is thus marred by inelegant justification. It would be difficult to disagree with Zinbarg and Mohlman's (1998) statement that "... the interactive effect of impulsivity by trait anxiety ... is not well understood at present" (p. 1038). It may thus seem tempting to conclude that RST provides an inadequate account of personality, and is in need either of modification or replacement by higher-level cognitive models (e.g., Matthews & Gilliland, 1999).

### 5.4. A two-process model of BIS/BAS functioning

It would be fair to say that experimental data do not conform strongly to a *separable subsystems* view of BIS/BAS functions, especially in relation to reward and Imp. A reasonable interpretation of this literature is that, sometimes, results conform to the *separable subsystems* hypothesis (especially for aversive stimuli), at other times, to the *joint subsystems* hypothesis. These two hypotheses are not mutually exclusive, but highly complementary, pointing to a two-process model of BIS/BAS effects. Pickering (1997) provides a highly valuable discussion of BIS/BAS interactions and provides tests of several neural network models (the arguments that follow, though not directly based upon these computational models, do share a number of common themes and are built upon a collaborative research programme; see Pickering et al., 1997).

#### 5.4.1. Facilitatory and antagonistic effects

Now assuming that Anx and BIS behaviours, and Imp and BAS behaviours, are related, as predicted by RST, then the experimental data summarised above is consistent with the view that the BAS and BIS have two effects: the first *facilitatory*, the second *antagonistic*. In the case of BIS behaviours, Anx+ facilitates, Imp+ antagonises; in the case of BAS behaviours, Imp+ facilitates, Anx+ antagonises. The precise pattern of personality effects is hypothesised to depend upon the relative strengths of aversive and appetitive stimuli (i.e., the degree of BIS and BAS activation). Speculatively, with weak aversive and appetitive stimuli, only antagonistic factors are at work, with Anx impairing BAS behaviour, and Imp impairing BIS behaviour; but with strong stimuli, facilitatory factors may play a more important part, with Anx facilitating BIS behaviour, and Imp facilitating BAS behaviour: in the case of very strong BIS/BAS-relevant stimuli, antagonistic effects may be of little importance.

Now it is much easier to manipulate aversive stimuli than appetitive stimuli in standard laboratory settings. Therefore, in hindsight, it is perhaps not surprising that *facilitatory* effects of the BIS are more commonly found (e.g., Corr et al., 1995c, 1997a) than *facilitatory* effects of the BAS; whereas, it is common to find *antagonistic* effects of the BIS (Corr et al., 1995a, c), *antagonistic* effects of the BAS are less common.

The *joint subsystems* hypothesis may explain the hitherto puzzling "complementary-trait" effects discussed by Pickering et al. (1997). That is, Anx− often predicts appetitive motivation, not Imp+ (e.g., Corr et al., 1995a, c). As stated by Matthews and Gilliland (1999, p. 607), such data "... suggest that, in more anxious individuals, individual differences in BIS activity elicited by the experimental setting tend to mask effects of positive reinforcement, because of the inhibitory effects of the BIS on the BAS"; but they then state that "This hypothesis does not explain the absence of effects of anxiety on passive avoidance in the Corr et al. (1995a) study" (in which we found that passive avoidance was highest in Imp− *S*s; i.e., Imp+ impaired appropriate passive avoidance behaviour in response to contingent punishment; this finding is consisted with data from other laboratories; e.g., Newman, 1987). The assumption of an antagonistic effect of Imp+ (BAS+) on reactions to punishment resolves this problem. Barratt (1971) reported a similar effect that could be accommodated within this model of antagonistic BAS effects: whereas Anx+ facilitated eyeblink conditioning, Imp+ impaired it, as shown by superiority in Imp−*S*s.

The nature of BIS/BAS interactions complicates the derivation of experimental hypotheses from the general principles of RST. However, both the *separable subsystems* and *joint subsystems* views predict that Anx and Imp, not E and N, should be most reliably related to behavioural reactions to conditioned stimuli for reward and punishment. On this basis, RST may be clearly distinguished from Eysenck's arousal theory.

The full implications of BIS/BAS interactions have yet to be realised: whether they truly clarify RST must await empirical investigation. Precise predictions required for such empirical investigation will be dependent on our improved ability accurately to characterise reinforcement parameters. Independent measures of these manipulations, including psychophysiological, behavioural and self-report, are required to evaluate the valence and strength of reinforcing stimuli. This validity evidence would permit comparison across studies and may therefore help to explain the diversity of findings that is a characteristic of RST research.

## 6. The future of biological personality models

There can be little doubt that cognitive constructs have an important role to play in personality research. As indicated by Matthews and Gilliland (1999), the challenges facing biological models are indeed formidable, and the issues that I have presented above serve further to highlight these challenges. However, there are reasons for assuming that significant advances are likely to be made in the biological understanding of personality.

Traditionally, biological personality research has been hindered by the necessity of constructing *hypothetical* neural mechanisms to account for laboratory findings (Pavlov's and Eysenck's notions of inhibition and excitation, and mobility of nervous processes, are obvious examples). These neural mechanisms proved largely inaccessible to direct investigation. But in recent years, this situation has radically changed. Significant advances are now being made in molecular biology and neuroscience that promise to revolutionise our understanding of biological personality processes. For example, in only the past few years, data has been accumulating which point to important molecular genetic contributions to well-established traits of personality (e.g., Benjamin, Ebstein & Belmaker, 1997; Benjamin et al., 1996; Ebstein et al., 1996; Noble et al., 1998; Staner et al., 1998). These data, though still preliminary, demonstrate at the very least that personality traits are not merely social constructions, but have direct connections with the material of genetic transmission (DNA). In addition, these associations are also beginning to point to putative causal mechanisms underlying personality traits (e.g., novelty seeking and dopamine receptors). Coupled with increasingly powerful neuroimaging techniques (e.g., anatomical and functional magnetic resonance imaging), these advances bode well for further attempts to unravel the biological complexity of human personality.

Now, whilst it would be naïve to assume that these advances will provide simple reductionist solutions to many of the problems that bedevil biological personality research, it would also seem to be equally unjustified to assume that such research will not continue to make significant contributions to our exploration of the cognitive-behavioural architecture of personality, linking DNA structure with protein processes, developmental pathways, brain physiology, and ultimately proximal functional systems that mediate positive and aversive motivation. A complementary, Janus-faced, strategy that studied, at the one end, fundamental biological processes and, at the other end, cognitive processes that mediate reactions to environmental stimuli, would seem the most prudent one to pursue. We would place ourselves at a severe disadvantage if we were to ignore such fundamental biological influences; conversely, we should not be wise to ignore the important role that cognitive accounts play in elucidating the effects of biological influences on learning and performance.

## 7. Conclusion

Matthews and Gilliland's (1999) impressive review of Gray's RST landed on the highs of predictive success as well as the lows of failure. In this commentary, I have sought to clarify a number of practical and theoretical problems and show that, in several crucial respects, RST has yet to be adequately tested; and that closer attention to operational definitions, and the mutual interplay of the BIS and BAS, may enhance its experimental precision. Although formidable challenges face RST researchers, fast-moving developments in molecular genetics and neuroscience point to significant advances being made in our understanding of the biology of human personality.

Without prejudice to the ultimate scientific status of RST, I would, at this present time, counsel caution regarding Matthews and Gilliland's (1999, p. 620) claim that "Cognitive constructs may be more appropriate than biological ones for explaining the majority of behaviours, so that explanations of the kind offered by the Eysenck and Gray theories are relevant to a restricted range of phenomena only".

## Acknowledgements

I am most grateful to Dr Gerry Matthews and Dr Alan Pickering for their valuable comments on an earlier draft of this article.

## References

Albert, M., & Bignami, G. (1968). Effects of frontal medial cortical and caudate lesions on two-way avoidance learning by rats. *Physiology and Behaviour, 3*, 141–147.

Ball, S. A., & Zuckerman, M. (1990). Sensation seeking, Eysenck's personality dimensions and reinforcement sensitivity in concept formation. *Personality and Individual Differences, 11*, 343–353.

Barratt, E. S. (1971). Psychophysiological correlates of classical differential eyelid conditioning among normal subjects selected on the basis of impulsiveness and anxiety. *Biological Psychiatry, 3*, 339–346.

Barry, H., Wagner, A. R., & Miller, N. E. (1962). Effects of alcohol and amobarbital on performance inhibited by experimental extinction. *Journal of Comparative and Physiological Psychology, 55*, 464–468.

Benjamin, J., Ebstein, R. P., & Belmaker, R. H. (1997). Personality genetics. *Israel Journal of Psychiatry and Related Sciences, 34*, 270–280.

Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., & Hamer, D. H. (1996). Population and familial association between the D4 dopamine receptor gene and measures of novelty seeking. *Nature Genetics, 12*, 81–84.

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

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments, 3*, 167–226.

Cook, E. W., Hawk, L. W., Davis, T. L., & Stevenson, V. E. (1991). Affective individual differences and startle reflex modulation. *Journal of Abnormal Psychology, 100*, 5–13.

Corr, P. J. (1994). *An investigation of the biological basis of the major dimensions of personality*. Unpublished doctoral thesis, University of London

Corr, P. J. (2000). Individual differences in reward sensitivity and impulsivity: the role of expectancies and frustrative nonreward (submitted)

Corr, P. J., & Kumari, V. (1997). Sociability/impulsivity and haloperidol-induced de-arousal: critical flicker/fusion frequency and procedural learning. *Personality and Individual Differences, 22*, 805–815.

Corr, P. J., Pickering, A. D., & Gray, J. A. (1995a). Personality and reinforcement in associative and instrumental learning. *Personality and Individual Differences, 19*, 47–71.

Corr, P. J., Pickering, A. D., & Gray, J. A. (1995b). Sociability/impulsivity and caffeine-induced arousal: critical flicker fusion frequency and procedural learning. *Personality and Individual Differences, 18*, 713–730.

Corr, P. J., Kumari, V., Wilson, G. D., Checkley, S., & Gray, J. A. (1997a). Harm avoidance and affective modulation of the startle reflex: a replication. *Personality and Individual Differences, 22*, 591–593.

Corr, P. J., Pickering, A. D., & Gray, J. A. (1997b). Personality, punishment, and procedural learning: a test of J. A. Gray anxiety theory. *Journal of Personality and Social Psychology, 73*, 337–344.

Corr, P. J., Wilson, G. D., Fotiadou, M., Kumari, V., Gray, N. S., Checkley, S., & Gray, J. A. (1995c). Personality and affective modulation of the startle reflex. *Personality and Individual Differences, 19*, 543–553.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences, 22*, 491–533.

Diaz A, & Pickering, A. D. (1993). The relationship between Gray's and Eysenck's personality spaces. *Personality and Individual Differences, 15*, 297–305.

Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., & Bemaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature Genetics, 12*, 78–80.

Evenden, J. L., & Ryan, C. M. (1996). The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology, 128*, 161–170.

Eysenck, H. J., & Eysenck, S. B. G. (1964). *Eysenck personality inventory*. London: University of London Press.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck personality questionnaire (adults)*. London: Hodder and Stoughton.

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.

Eysenck, H. J., & Eysenck, S. B. G. (1991). *Manual of the Eysenck personality scales*. London: Hodder and Stoughton.

Eysenck, H. J., & Levey, A. (1972). Conditioning, introversion–extraversion and the strength of the nervous system. In V. D. Nebylitsyn, & J. A. Gray, *Biological bases of individual behaviour* (pp. 206–220). London: Academic Press.

Gray, J. A. (1967). Disappointment and drugs in the rat. *Advancement of Science, 23*, 595–605.

Gray, J. A. (1969). Sodium amobarbital and effects of frustrative non-reward. *Journal of Comparative and Physiological Psychology, 69*, 55–64.

Gray, J. A. (1970). The psychophysiological basis of introversion–extraversion. *Behaviour Research and Therapy, 8*, 249–266.

Gray, J. A. (1976). The behavioural inhibition system: a possible substrate for anxiety. In M. P. Feldman, & A. M. Broadhurst, *Theoretical and experimental bases of behaviour modification* (pp. 3–41). London: Wiley.

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

Gray, J. A. (1987). *The psychology of fear and stress*. Cambridge: Cambridge University Press.

Gray, J. A., Kumari, V., Lawrence, N., & Young, A. M. J. (1999). Functions of the dopaminergic innervation of the nucleus accumbens. *Psychobiology, 27*, 223–235.

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

Green, R. H., Beatty, W. W., & Schwartzbaum, J. S. (1967). Comparative effects of septo-hippocampal and caudate lesions on avoidance behaviour in rats. *Journal of Comparative Neurology and Psychology, 64*, 444–452.

Grossen, N. E., Kostansek, D. J., & Bolles, R. W. (1969). Effects of appetitive discriminative stimuli on avoidance behaviour. *Journal of Experimental Psychology, 81*, 340–343.

Gupta, B. S. (1976). Extraversion and reinforcement in verbal operant conditioning. *British Journal of Psychology, 67*, 47–52.

Gupta, B. S., & Nagpal, M. (1978). Impulsivity/sociability and reinforcement in verbal operant conditioning. *British Journal of Psychology, 68*, 203–206.

Hernaiz, H. (1991). *Human reactions to arousal and reinforcement: a test of the relationship between Eysenck's and Gray's theories of personality*. Unpublished doctoral thesis, University of London

Inoue, A., French, C. C., Pring, L., & Corr, P. J. (2000). Trait anxiety effects of affective modulation on the startle reflex, and decision and movement time components of the emotional Stroop task (submitted).

Isaacson, R. L., Douglas, R. J., & Moore, R. Y. (1961). The effect of radical hippocampal ablation on acquisition of avoidance response. *Journal of Comparative and Physiological Psychology, 54*, 625–628.

Ison, J. R., & Pennes, E. S. (1969). Interaction of amobarbital and reinforcement schedule in determining resistance to extinction of an instrumental running response. *Journal of Comparative and Physiological Psychology, 68*, 215–219.

Kamano, D. K., Martin, L. K., & Powell, B. J. (1966). Avoidance response acquisition and amobarbital dosage levels. *Psychopharmacologia, 8*, 319–323.

Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.

Konorski, J., & Szwejkowska, G. (1952). Chronic extinction and restoration of conditioned reflexes: IV. The dependence of the course of extinction and restoration of conditioned reflexes on the 'history' of the conditioned stimulus (the principle of the primary of first training). *Acta Biologiae Experimentalis, 17*, 141–165.

Kumari, V., Corr, P. J., Wilson, G. D., Kaviani, H., Thornton, J. C., Checkley, S. A., & Gray, J. A. (1996). Personality and modulation of the startle reflex by emotionally-toned filmclips. *Personality and Individual Differences, 21*, 1029–1041.

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

MacAndrew, C., & Steele, T. (1991). Gray's behavioural inhibition system: a psychometric examination. *Personality and Individual Differences, 12*, 157–171.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Clarendon Press.

Mangan, G. L. (1978). Factors of conditionability and their relationship to personality types. *Pavlovian Journal of Biological Science, 13*, 226–235.

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

Miller, N. E. (1959). Liberalization of basic S–R concepts: extensions to conflict behaviour, motivation and social learning. In S. Koch, *Psychology: a study of a science* (pp. 196–292). New York: McGraw-Hill.

Miller, N. E. (1964). The analysis of motivational effects illustrated by experiments on amylobarbitone sodium. In H. Steinberg, *Animal behaviour and drug action* (pp. 1–18). London: Churchill.

Mowrer, O. H. (1960). *Learning and behaviour*. New York: Wiley.

Nagpal, M., & Gupta, B. S. (1979). Personality, reinforcement and verbal operant conditioning. *British Journal of Psychology, 70*, 471–476.

Newman, J. P. (1987). Reaction to punishment in extraverts and psychopaths: Implications for the impulsive behaviour of disinhibited individuals. *Journal of Research in Personality, 21*, 464–480.

Noble, E. P., Ozkaragoz, T. Z., Ritchie, T. L., Zhang, X. X., Belin, T. R., & Sparkes, R. S. (1998). D-2 and D-4 dopamine receptor polymorphisms and personality. *American Journal of Medical Genetics, 81*, 257–267.

Olds, N. E., & Fobes, J. C. (1981). The central basis of motivation: intracranial self-stimulation studies. In M. K. Rozenzweig, G. Porter, & C. W. Porter, *Annual review of psychology, Vol. 32* (pp. 523–574). Palo Alto: Annual Reviews Inc.

Olton, D. S., & Isaacson, R. L. (1968). Hippocampal lesions and active avoidance. *Physiology and Behaviour, 3*, 719–724.

Ostaszewski, P. (1996). The relation between temperament and rate of temporal discounting. *European Journal of Personality, 10*, 161–172.

Paisey, T. J. H., & Mangan, G. L. (1988). Personality and conditioning with appetitive and aversive stimuli. *Personality and Individual Differences, 9*, 69–78.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., Williams, S. C. R., Bullmore, E. T., Brammer, M., & Gray, J. A. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London Series B: Biological Sciences, 265*, 1809–1817.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389*, 495–498.

Pickering, A. D. (1997). The conceptual nervous system and personality: from Pavlov to neural networks. *European Psychologist, 2*, 139–163.

Pickering, A. D., Corr, P. J., & Gray, J. A. (1999). Interactions and reinforcement sensitivity theory: a theoretical analysis of Rusting and Larsen (1997). *Personality and Individual Differences, 26*, 357–365.

Pickering, A. D., Corr, P. J., Powell, J. H., Kumari, V., Thornton, J. C., & Gray, J. A. (1997). Individual differences in reactions to reinforcing stimuli are neither black nor white: to what extent are they gray?. In H. Nyborg, *The scientific study of human nature: tribute to Hans J. Eysenck at eighty* (pp. 36–67). London: Elsevier Sciences.

Rabe, A., & Haddad, A. K. (1969). Acquisition of two-way shuttle-box avoidance after selective hippocampal lesions. *Physiology and Behaviour, 4*, 319–323.

Schweitzer, J. B., & Sulzerazaroff, B. (1995). Self-control in boys with attention-deficit hyperactivity disorder: effects of added stimulation and time. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*, 671–686.

Sonugabarke, J. S., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion 1: the effect of delay on choice. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*, 387–398.

Spielberger, C. D., Gorsuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Staner, L., Hilger, C., Hentges, F., Monreal, J., Hoffmann, A., Couturier, M., BeBon, O., Stefos, G., Souery, D., & Mendlewicz, J. (1998). Association between novelty-seeking and the dopamine D3 receptor gene in bipolar patients: a preliminary report. *American Journal of Medical Genetics, 81*, 192–194.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford: Oxford University Press.

Torrubia, R., & Tobena, A. (1984). A scale for the measurement of 'susceptibility to punishment' as a measure of anxiety: preliminary results. *Personality and Individual Differences, 5*, 371–373.

Wagner, A. R. (1963). Sodium amytal and partially reinforced runway performance. *Journal of Experimental Psychology, 65*, 474–477.

Wilson, G. D., Barrett, P. T., & Gray, J. A. (1989). Human reactions to reward and punishment: a questionnaire measure of Gray's personality space. *British Journal of Psychology, 80*, 509–515.

Wilson, G. D., Kumari, V., Gray, J. A., & Corr, P. J. (2000). The role of neuroticism in reflex reactions to fearful and disgusting stimuli. *Personality and Individual Differences*, in press

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

Zinbarg, R., & Revelle, R. (1989). Personality and conditioning: a test of four models. *Journal of Personality and Social Psychology, 57*, 301–314.