Approach and Avoidance Behaviour: Multiple Systems and their Interactions

Philip J. Corr
Department of Psychology, City University London, UK

Abstract

Approach–avoidance theories describe the major systems that motivate behaviours in reaction to classes of appetitive (rewarding) and aversive (punishing) stimuli. The literature points to two major "avoidance" systems, one related to pure avoidance and escape of aversive stimuli, and a second, to behavioural inhibition induced by the detection of goal conflict (in addition, there is evidence for nonaffective behavioural constraint). A third major system, responsible for approach behaviour, is reactive to appetitive stimuli, and has several subcomponents. A number of combined effects of these systems are outlined. Finally, the hierarchical nature of behavioural control is delineated, including the role played by conscious awareness in behavioural inhibition.

Keywords

approach, avoidance, behavioural inhibition, constraint, emotion, motivation, personality

Approach–avoidance theories aim to describe the major systems that motivate behaviours in reaction to classes of appetitive (rewarding) and aversive (punishing) stimuli, and to explain consistent patterns of individual differences in these behaviours. Current theories trace their origins to early researchers, especially those learning theorists who rejected the Hullian notion of a single drive underlying reinforcement-based behaviour (Hull, 1952). In place of this single-factor theory, they pointed to two central state (emotion) processes. Mowrer (1960), in particular, was a pioneer in this field, and the work of others (e.g., Konorski, 1967) made further significant progress. An important development of this work, especially in terms of its significance for human emotion, motivation and learning, came through the work of Jeffrey Gray (1970, 1972, 1975, 1982). This work had a major influence on current-day agreement that there are a small number of state systems that mediate reactions to different classes of reinforcing stimuli (i.e., rewards and punishments of various kinds) and which generate emotion and shape ("motivate") approach–avoidance behaviour. This work also led to important links being made between on the one hand, state systems and traits, and on the other hand, psychopathology (e.g., anxiety and depression), as well as all varieties of everyday behaviour (e.g., gambling, substance abuse, economic decisions; for a summary, see Corr, 2008a).

Gray's early work led directly to the reinforcement sensitivity theory (RST; Gray & McNaughton, 2000; see also Corr,

2008a), which postulates three major neuropsychological systems: one positive, the *behavioural approach system* (BAS); and two negative, the *fight–flight–freeze system* (FFFS) and the *behavioural inhibition system* (BIS). The BAS is activated by appetitive stimuli (e.g., food and sexual partners); the FFFS by aversive stimuli (e.g., predators); and the BIS by conflicting stimuli (e.g., coactivation of FFFS and BAS when their motivational tendencies are opposing, as in avoidance–approach conflict seen in many social situations).

In terms of definition and operational procedures, the appetitive and aversive stimuli which activate these approach–avoidance systems are defined independently of the individual under test: They are defined in terms of typical reactions of the vast majority of individuals in the population (i.e., in terms of the "direction" of behaviour, namely, to avoid in the case of "aversive" stimuli, and approach in the case of "appetitive" stimuli). However, there are considerable individual differences in these reactions, and these differences comprise the personality bases of approach and avoidance motivations (principally, but not exclusively, related to extraversion and neuroticism). These reactions are uncomplicated with the use of unambiguous stimuli (e.g., a pain-inducing shock). However, it should be noted that some stimuli (e.g., a conspecific) may elicit different reactions in different people depending on their evaluation of the stimulus, for example producing avoidance in some people but approach in others. In such cases, it is important to determine

Corresponding author: Philip J. Corr, Department of Psychology, City University London, Northampton Square, London EC1V 0HB, UK. *Email*: philip.corr@btopenworld.com

the typical "direction" of behaviour induced by such stimuli and the prevalence of contrary reactions. In other words, it is important to determine that "aversive" stimuli are actually punishing and "appetitive" stimuli are actually rewarding. This requirement is especially important when extending animal-based models to human behaviour, where perception and valuation determine in many cases so-called "appetitive" and "aversive" stimuli: These stimuli need to be experimentally validated and not just theoretically assumed.

In addition to these considerations, the specific forms of approach–avoidance behaviour will depend on context and the environment. Some environments afford certain forms of behaviour. For example, simple avoidance of a threatening stimulus may be possible in one environment (e.g., a park with lots of barking dogs), but not in another environment which may call forth a different form of defensive reaction (e.g., fleeing or fighting—being approached at night by an aggressive individual). Context, too, is important. For example, in most people an aggressive boss would elicit a different reaction than an aggressive stranger; and very often we have to inhibit automatic reactions depending on the situation (e.g., fleeing from the sound of the dentist's drill, or getting flustered when giving an oral presentation in front of people). Therefore, the rather clinical-sounding terms "approach" and "avoidance" need to be seen in the light of the affordances and constraints of specific situations. This is especially important when assigning motivational functions to these behaviours: We cannot simply "read-off" functions from them without considering the specific context in which behaviour is elicited and observed.

As discussed in what follows, the general state systems of the FFFS, BIS, and BAS are rather well established in the broader research literature—including animal neurophysiology, human experimental, and personality. It is these state systems, their interactions, and how they differ between individuals, that may be seen to form the foundations of the whole family of approach–avoidance theories. However, at this stage of knowledge, it needs to be recognised that, although Gray's approach may be a good starting point from which to work towards a consensual model, much more research is needed before any firm conclusions can be reached.

Foundations of Approach–Avoidance Behaviour

The approach adumbrated by Gray during the 1970s was founded on several tenets. First, that emotions (e.g., fear and hope) are central states activated by reinforcing stimuli (generally called "punishment" and "reward"); and second, that two major systems underpin the activation of these central states, one related to sensitivity/reactivity to "punishment" and another to "reward." At least in part, individual differences, as expressed in personality traits (e.g., extraversion and neuroticism), reflect long-term stabilities in the operation of these state systems. (Space prevents adequate discussion of the personality side of approach–avoidance motivation; for discussion, see Corr, DeYoung, & McNaughton, 2013.)

Now, if we define reward and punishment as stimuli that evoke emotional responses then we are getting very close to inferential hazard. However, this danger may be mitigated: (a) if the definition of "reward" includes the termination or omission of punishment (this is the "hope = relief" hypothesis), and the definition of "punishment" includes the termination or omission of reward (this is the "fear = frustration" hypothesis); and (b) if drugs (e.g., affecting dopamine and serotonin systems) are used to dissect the classes of behaviour related to these "reward" and "punishment" systems. By these means "a theory of the central states relating to approach–avoidance behaviour can be built and, then, empirically tested" (see Fowles, 2006).

A good example of the potential of falsification of this approach is seen in revised RST: In 2000, Gray and McNaughton rejected the theory of the BIS as one related to conditioned aversive stimuli (e.g., a signal of the presentation of a pain-inducing stimulus), replacing it with a more general one that is sensitive to goal conflicts of all kinds (e.g., between approach and avoidance motivation when they are in opposition). In addition, this previous version of BIS (Gray, 1982) defined its adequate inputs in terms of reactions to stimuli that were preferentially affected by antianxiety drugs (mainly barbiturates, alcohol, and benzodiazepines). This association verged on tautology. However, mitigation of this specific inferential hazard came with the discovery that anxiety was reduced by the newer classes of antianxiety drugs (e.g., selective serotonin reuptake inhibitors [SSRIs]), which also acted on the septo-hippocampal system, as originally proposed, and which also affected the electrocortical signature of BIS activity, namely the theta rhythm (Gray & McNaughton, 2000; see also Mitchell, McNaughton, Flanagan, & Kirk, 2008).

Approach–Avoidance Systems and their Interactions

In general terms, "reward" stimuli motivate approach behaviour, and "punishment" stimuli motivate avoidance/escape behaviour (Gray, 1975). But here there are some complexities that any theory must consider. At the state level, reward and punishment motivations (approach–avoidance tendencies) subtract from each other, and have different goal gradients (Miller, 1944). And, in addition to these two systems, there is a third system of "avoidance": Over and above these subtractive effects, the inhibition of approach by approach–avoidance *conflict* is neurally and psychopharmacologically distinct from simple avoidance/escape (Gray, 1982; Gray & McNaughton, 2000). In RST parlance, these two "avoidance" motivations are controlled by two parallel processes: FFFS and BIS. Whereas the BIS is generally sensitive to anxiolytic drugs, the FFFS is *relatively* insensitive to anxiolytic drugs but sensitive to panicolytic ones (Perkins et al., 2009).

Such pharmacological data add support to a crucial point: In plain English, "behavioural inhibition," if this means a reduction in behaviour, is not necessarily dependent on the BIS. When reward and punishment are not approximately equal in value and are not in conflict, they subtract from each other

(Gray & Smith, 1969), and resulting reduction in behaviour is specifically *not* affected by anxiolytic drugs (McNaughton & Gray, 1983); but when they are approximately equal and in conflict, then the BIS is activated and anxiolytic drugs affect it. In addition, there may be a form of "nonaffective constraint" on behaviour reflecting the cortical inhibition of behaviour which is independent of the aforementioned two forms of behavioural reduction/inhibition (Carver, Johnson, & Joormann, 2008).

Therefore, the processing of conflict and the resultant behavioural inhibition is not the same as simple (pure) avoidance, although in both cases it appears that behaviour is inhibited (Gray & McNaughton, 2000). These different types of avoidance are often in opposition to each other: Freezing, fighting, and fleeing are all forms of avoidance, whereas behavioural inhibition allows cautious approach (or withholding of entry) to a dangerous place. We see this opposition in the everyday example of the motivation to flee from the dentist's chair because of the potential pain involved in the procedure and the behavioural inhibition, and resulting anxiety, induced by the conflict, namely the reward of getting one's teeth repaired/improved at the same time as the unpleasantness of the situation (here, too, we witness the control of behaviour by higher-level cognitive processes which allow us to endure the situation because we know the outcome will be beneficial). An important implication of this scheme is that, in order to study this conflict-related behavioural inhibition proper, it is important first to characterise and measure simple approach and avoidance and, then, to compare the effects of behavioural inhibition superimposed on these pure forms of approach and avoidance.

Critically, then, when theorists talk about "avoidance," they need to be clear whether it is active avoidance (pure avoidance/ escape) or passive avoidance (conflict-related avoidance); and, on top of this, there seems to be a nonaffective form of constraint on behaviour.

Separable and joint effects. As noted before, an obvious interaction of reward/approach and punishment/avoidance is the subtractive effect on choice behaviour. One incarnation of this subtractive rule is seen in the case of trait measures of approach–avoidance behaviour. Based on the original work of Gray and Smith (1969), there is a growing literature on the "joint subsystems hypothesis" of reward and punishment systems (for a summary of empirical studies, see Corr, 2004). Studies indicate that, in predicting human reactions to various types of reward and punishment, consideration should be given to their combined effects (sometimes revealed by two main effects, at other times by statistical interactions). Combined effects are based on two assumptions: (a) Most forms of behaviour are not pure in the sense that there are varying degrees of aversive and appetitive stimuli in the ambient environment; and (b) behaviour is often the result of the activation of multiple state systems (i.e., systems work together to shape the final behaviour shown). An everyday example would be performance during a job interview which, typically, will be jointly influenced by the positive motivation to impress as well as the fear/anxiety of saying the wrong thing and looking foolish or incompetent. However, where pure behaviour can be assured (e.g., an intense level of threat), then separable effects are to be expected. The subtractive effect discussed before provides the major theoretical rationale for such combined effects of rewarding and punishing stimuli, but there are other possible forms of interaction, as discussed below.

Thus, basic animal studies suggest: (a) There are, at least, two "avoidance" systems, one for simple active avoidance/ escape (FFFS) and one for goal conflict (passive avoidance; BIS); and (b) these systems often interact with the reward system in a number of ways (e.g., FFFS–BAS, in a subtractive fashion; and BIS–BAS in an inhibitory fashion)—although there are a number of possible other interactions, some additive, between these systems (e.g., FFFS-related flight to a place of safety would also entail the BAS in a unified action; for a discussion of these relationships, see Boureau & Dayan, 2011).

Approach Behaviour and its Components

Approach behaviour, which is initiated by the presentation of a stimulus that is perceived to be rewarding, has received much less theoretical attention than the defensive systems discussed before, and there is much still to be done to clarify the component processes. Nonetheless, there is good reason to believe that approach behaviour is multidimensional (Carver & White, 1994).

The primary function of the system controlling approach behaviour is to move the animal up the temporo-spatial gradient, from a start state (e.g., the idea of, or the physical distance to a source of food), towards the final biological reinforcer (e.g., consumption of food). To move along the temporo-spatial gradient to the final biological reinforcer, some form of "subgoal scaffolding" is needed (Corr, 2008b). This process consists of (a) identifying the biological reinforcer, (b) planning behaviour, and (c) executing the plan (i.e., "problem solving") at each stage of the temporo-spatial gradient. This approach behaviour entails a series of subprocesses, some of which oppose each other. For example, behaviour restraint and planning are often required to achieve approach goals (e.g., making arrangements for a holiday), but not at the final point of capture of the biological reinforcer (having fun on holiday), where nonplanning and fast reactions (i.e., impulsivity) are more appropriate. As aptly noted by Carver (2005, p. 312), "unfettered impulse can interfere with the attainment of longer term goals."

Is there any evidence for the aforementioned claims? Well, at the psychometric level, there are separate components to BAS approach behaviour. In replicated samples, Corr and Cooper (2013) found evidence for four subfactors, comprising "reward interest" and "goal-drive persistence," that characterise the early stages of approach, and "reward reactivity" and "impulsivity," that characterise the behavioural and emotional excitement as the final biological reinforcer is reached. Emotion in the former case may be termed "anticipatory pleasure" (or "hope"); in the latter case something akin to an "excitement attack" of high pleasure/joy. This four-factor model updates Carver and White's (1994) three-factor model of trait approach behaviour. How

these separate components of the BAS relate, in terms of subtractive, additive, and inhibitory effects, to the FFFS and BIS is as yet unknown.

Hierarchical Control of Automatic and Controlled Processes

Approach–avoidance behaviour is controlled by a hierarchical system of neuropsychological processes, ranging from the reflexive-automatic to the reflective-controlled (including conscious awareness). In an attempt to tackle this issue, Corr (2010) developed a model, based on a neuropsychological model of the functions of consciousness (Gray, 2004), which postulates that all behaviours (and related thoughts, feeling, etc.) are automatically organised and executed, without the *immediate* control by higher-level controlled processes (and certainly not conscious processing, which simply takes too long to be generated by the brain to have immediate control over the events it *re*presents).

The model states that when everything is "going to plan" (i.e., things are as expected), we are not generally aware of ongoing events (however, events and stimuli that are particularly important for ongoing goals do attract controlled processing). It is only at critical junctures (i.e., the expected does not happen) that the outputs of processing attract conscious awareness, and these outputs tend to entail (goal conflict) error, usually in the form of actual states of the world departing from expected states. For example, whilst driving a car we may find ourselves braking hard and only *then* realise why we braked—that is, we are conscious of the error only after it has occurred and only after the brain has executed the appropriate (reflexive-automatic) response (Figure 1).

The proximal–distal aspects of threat are of importance too, as is the level of intensity. High-intensity threat in the context of goal conflict would quickly resolve itself in the form of FFFS-related behaviour. But when threats are less intense or perceived to be distant in terms of space or time, then BIS-related processing allows the individual to engage in approach behaviour but in a much more cautious and risk-aversive manner. Thus, when threats are intense and immediate, automatic processing dominates; but when threats are less intense and not immediate, then controlled processes are activated to risk assess the problem
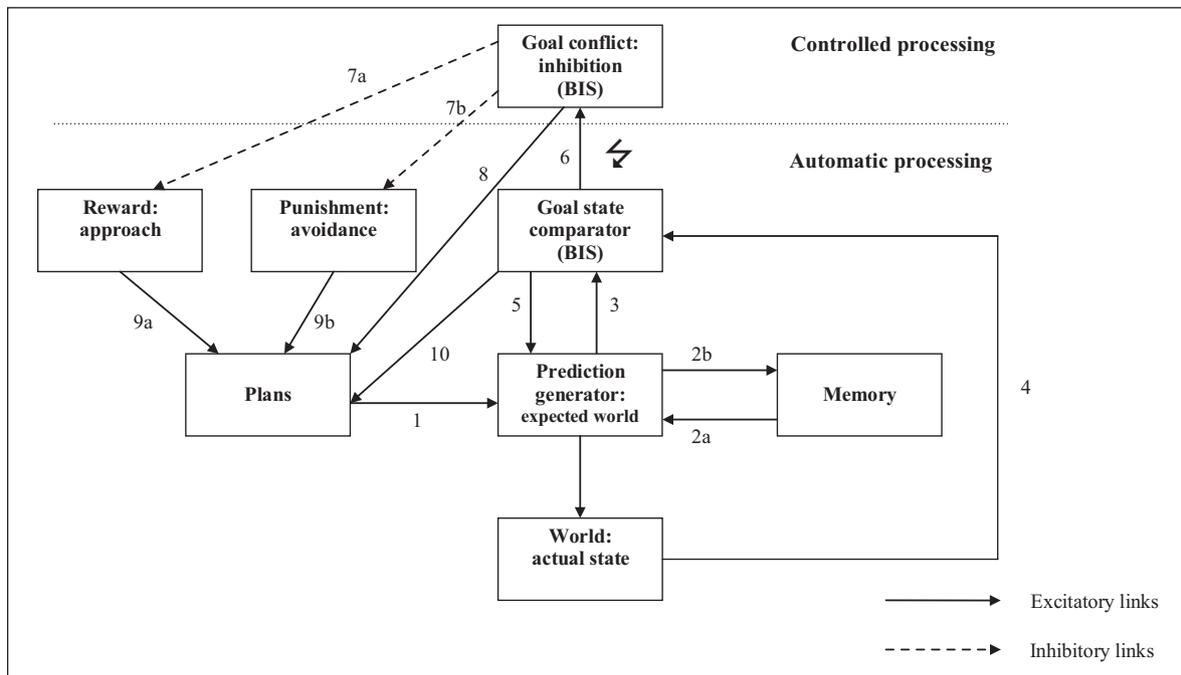

Figure 1. Information processing diagram of the functioning of the behavioural inhibition system (BIS) in automatic and controlled processing modes, which contains basic reward:approach and punishment:avoidance processes, as well as goal conflict device related to the BIS. Behavioural plans (plans) lead to predictions (prediction generator; 1) of future states of the world, which receives input from (2a), and sends output to (2b), stored previous experience (memory). The BIS (goal state comparator) receives input from the prediction generator (3), and then compares the response reinforcement outcomes (world: actual state) with predictions (4), and then one of two things happen: (a) "everything is going to plan," and the BIS goal state comparator sends input to the prediction generator to continue the motor program ("just checking mode"; 5); or (b) the BIS goal state generator detects a mismatch between prediction and outcome and generates an error signal (⚡), which leads to activation of the BIS and controlled processes (6).

Once the BIS is activated, there is inhibition of the reward:approach system (BAS; 7a) and the punishment:avoidance system (FFFS; 7b); and at this time the BIS initiates cautious behaviour and risk assessment, which then informs plans (8), which simultaneously receives input, about current states, from the BAS and FFFS (9a, b), as well as input, about the nature of the conflict, from the BIS goal state comparator (10). Plans initiate appropriate behaviour and the described cycle is repeated, until behavioural resolution is achieved in the form of punishment-related avoidance/escape or reward-related approach.

situation. It is assumed that stimuli associated with error enter conscious awareness where they are replayed for detailed analysis; and, after this analysis, the automatic neural-behavioural machinery that controls behaviour at any given moment is (re)adjusted so that future behaviour is more appropriate when the same set of stimuli, which led to the error signal, are encountered again. By this means, we learn from our mistakes and the machinery that controls our automatic behaviour is better prepared when it encounters a similar situation next time (e.g., we no longer flinch so readily at the sight of an attacking shark at a 3-D movie). A failure to adjust our behaviour on the basis of past experience is generally seen as highly maladaptive and indicative of a problem of control (e.g., a losing gambler chasing their losses).

Fight/Anger

The previous model of behavioural control may help to explain the problem of how fight/anger relates to the approach–avoidance systems. Basic animal research indicates that predatory, or instrumental, fight should be associated with the BAS; however, when it is of a defensive nature it should be associated with the FFFS as its function is to remove the animal from the source of a high-intensity and immediate threat (e.g., cat to rat), especially when other forms of escape are not available. When cornered by a predator, animals do fight back, and the same is often seen in human *behavioural* reactions in a comparable situation. Thus, conceptually, and in behavioural terms, defensive fight should be expected to be part of the FFFS. But there is now a growing literature to show that fight/anger is more associated with the BAS (Carver & Harmon-Jones, 2009; Harmon-Jones, 2003). This literature seems to point to several relevant issues in extending animal-based models to human research, as well as highlighting the importance of the general form of the behavioural control model shown in Figure 1.

First, it may be difficult to distinguish reactive, defensive aggression (controlled by FFFS) from instrumental aggression (controlled by the BAS)—language may simply fail to differentiate the psychological states of each type. Second, because aggression involves behavioural activation, even when defensive, this is a case in which the sensitivity of multiple systems is likely to be important, and thus the fight component of the FFFS is likely to be potentiated by the BAS. Third, for defensive fight, low base rates may be a problem (at least in the population who tend to take part in psychology studies), if items that describe behaviours are manifested very infrequently in normal human life. Finally, measurement of fight might be best achieved by a behavioural measure, not a questionnaire one, especially if it represents a form of automatic, prepotent, reaction to a high-intensity, and inescapable, threat. This last possibility is likely to be of considerable importance across the whole range of approach–avoidance behaviours. For this reason, all forms of approach–avoidance behaviour should be defined in terms of overt reactions to experimentally controlled stimuli in addition to any attempt to measure them by questionnaire.

Conclusions

Research over many decades points to at least two major system of avoidance: pure avoidance/escape and conflict-related behavioural inhibition (and possibly a third, nonaffective, form of constraint); and, on the reward side, a multicomponent incentive system that is responsible for mediating the complex cascade of responses from speculative appetitive exploration to capture of the biological reward. These systems interact in a number of sometimes complex ways to influence behaviour. Long-term stabilities in these systems comprise the foundations of personality traits. Important too are hierarchical behavioural control processes, including the exotica of conscious awareness. It may be this wider panorama that is of ultimate importance, especially in our attempts to integrate approach–avoidance theories within the larger psychological landscape.

References

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, *36*, 74–97. doi: 10.1038/npp.2010.151

Carver, C. S. (2005). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review*, *9*, 312–333. doi: 10.1207/s15327957pspr0904_2

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, *135*, 183–204. doi: 10.1037/a0013965

Carver, C. S., Johnson, S. L., & Joormann, J. (2008). Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: What depression has in common with impulsive aggression. *Psychological Bulletin*, *134*, 912–943. doi: 10.1037/a0013740

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. doi: 10.1037/0022–3514.67.2.319

Corr, P. J. (2004). Reinforcement sensitivity theory and personality. *Neuroscience and Biobehavioral Reviews*, *28*, 317–332. doi:10.1016/j.neubiorev.2004.01.005

Corr, P. J. (2008a). *The reinforcement sensitivity theory of personality*. Cambridge, UK: Cambridge University Press.

Corr, P. J. (2008b). Reinforcement sensitivity theory (RST): Introduction. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp. 1–43). Cambridge, UK: Cambridge University Press.

Corr, P. J. (2010). Automatic and controlled processes in behavioral control: Implications for personality psychology. *European Journal of Personality*, *24*, 376–403. doi: 10.1002/per.779

Corr, P. J., & Cooper, A. (2013). *The reinforcement sensitivity theory of personality questionnaire (RST-PQ): Development and validation*. Unpublished manuscript.

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

Fowles, D. C. (2006). Jeffrey Gray's contributions to theories of anxiety, personality, and psychopathology. In T. Canli (Ed.), *Biological basis of personality and individual differences* (pp. 7–34). New York, NY: Guilford Press.

Gray, J. A. (1970). The psychophysiological basis of introversion–extraversion. *Behavior Research and Therapy*, *8*, 249–266. doi: 10.1016/0005–7967(70)90069–0

Gray, J. A. (1972). Learning theory, the conceptual nervous system and personality. In V. D. Nebylitsyn & J. A. Gray (Eds.), *The biological bases of individual behavior* (pp. 372–99). New York, NY: Academic Press.

Gray, J. A. (1975). *Elements of a two-process theory of learning*. London, UK: Academic Press.

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

Gray, J. A. (2004). *Consciousness: Creeping up on the hard problem*. Oxford, UK: Oxford University Press.

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

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

Harmon-Jones, E. (2003). Anger and the behavioral approach system. *Personality and Individual Differences*, *35*, 995–1005. doi: 10.1016/S0191–8869(02)00313–6

Hull, C. L. (1952). *A behavior system*. New Haven, CT: Yale University Press.

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

McNaughton, N., & Gray, J. A. (1983). Pavlovian counterconditioning is unchanged by chlordiazepoxide or by septal lesions. *Quarterly Journal of Experimental Psychology*, *35B*, 221–233. doi: 10.1080/14640748308400907

Miller, N. E. (1944). Experimental studies of conflict. In J. M. Hunt (Ed.), *Personality and the behavioral disorders* (pp. 431–465). New York, NY: Ronald Press.

Mitchell, D. J., McNaughton, N., Flanagan, D., & Kirk, I. J. (2008). Frontal midline theta from the perspective of hippocampal "theta." *Progress in Neurobiology*, *86*, 156–185. doi: 10.1016/j.pneurobio.2008.09.005

Mowrer, O. H. (1960). *Learning theory and behavior*. New York, NY: Wiley.

Perkins, A. M., Ettinger, U., Davis, R., Foster, R., Williams, S. C. R., & Corr, P. J. (2009). Effects of lorazepam and citalopram on human defensive reactions: Ethopharmacological differentiation of fear and anxiety. *Journal of Neuroscience, 29*, 12617 -12624.