Applying a cognitive neuroscience perspective to the disorder of psychopathy

R. J. R. BLAIR
Mood and Anxiety Disorders Program, National Institute of Mental Health

Abstract

Four models of psychopathy (frontal lobe dysfunction, response set modulation, fear dysfunction, and violence inhibition mechanism hypotheses) are reviewed from the perspective of cognitive neuroscience. Each model is considered both with respect to the psychopathy data and, more importantly, for the present purposes, with respect to the broader cognitive neuroscience fields to which the model refers (e.g., models of attention with respect to the response set modulation account and models of emotion with respect to the fear dysfunction and violence inhibition mechanism models). The paper concludes with an articulation of the more recent integrated emotion systems model, an account inspired both by recent findings in affective cognitive neuroscience as well as in the study of psychopathy. Some directions for future work are considered.

According to the website of the Cognitive Neuroscience Society, the term cognitive neuroscience emerged in the late 1970s in the back seat of a New York City taxi. The neuroscientist Michael Gazzaniga and the cognitive psychologist George A. Miller were traveling to a dinner organized to consider how the brain enables the mind, a subject in need of a name. The term cognitive neuroscience had emerged by the end of the ride.

The cognitive neuroscientific approach is a linguistic formalization of one form of inter-discipline integration that is occurring. The importance of a multiple levels of analysis approach is clear (Cicchetti & Dawson, 2002; Kopnisky, Cowan, & Hyman, 2002). For example, to understand a disorder we need to specify its behavioral profile (i.e., its clinical description), the functional impairments that give rise to this behavioral profile (i.e., cogni-

tive psychology), the neural systems that mediate these functions (i.e., systems neuroscience), the molecular level factors that are impacting on the neural systems (i.e., molecular neuroscience), and the genetic bases of these molecular level factors. Cognitive neuroscience is the relatively formalized combination of the fields of cognitive psychology and systems neuroscience. In short, cognitive neuroscience identifies an interdisciplinary approach to understanding the nature of thought. Sometimes the term affective neuroscience is used to identify an interdisciplinary approach to understanding the nature of emotion (Davidson, 2003; Panksepp, 1998). However, for the purposes of this paper, affective processing will be considered as simply another form of cognitive processing.

Cognitive psychology is concerned with the functional architecture of the mind. It need not be concerned with determining the regions of the brain that implement this architecture. A cognitive psychologist might be interested in a particular cognitive function and could have no position on what brain regions mediate this function. Neuroscience is concerned with the neural architecture of

Address correspondence and reprint requests to: James Blair, Unit on Affective Cognitive Neuroscience, Mood and Anxiety Disorders Program, National Institute of Mental Health, 15K North Drive, Room 206, MSC 2670, Bethesda, MD 20892-2670; E-mail: blairj@intra.nimh.nih.gov.

the brain. It need not be concerned with the specific functions that the individual systems implement. A neuroscientist investigating a particular disorder might be interested in a particular neural system and have no position on what functions this system is involved in mediating; the individual might use "probe tasks" to examine whether there is anomalous activity in the region of interest in the population under study.

To illustrate further with respect to aversive conditioning: a cognitive psychological account of aversive conditioning might make reference to the formation of stimulus–reinforcement associations. A cognitive psychology account of a disorder marked by impairment in aversive conditioning might consider the developmental implications of this impairment and perhaps suggest, to ameliorate the problem, the use of specific compensatory strategies for the patient. A neuroscience account of aversive conditioning might refer to the amygdala and temporal cortex and perhaps the role of neurotransmitters such as norepinephrine. A neuroscience account of a disorder marked by impairment in aversive conditioning might consider these neural/ neurotransmitter systems and perhaps suggest, as a potential treatment, pharmacological agents that might modulate the neurotransmitter level. Cognitive neuroscience, by definition, is concerned with the functional architecture of the neural architecture; how components of brain regions interact to achieve particular tasks. For example, a cognitive neuroscience account of aversive conditioning might make reference to the role of the amygdala in allowing the formation of stimulus–reinforcement associations, temporal cortex in representing the sensory qualities of the presented stimuli, and norepinephrine in accentuating the salience of those stimulus features associated with reinforcement. A cognitive neuroscience account of a disorder marked by impairment in aversive conditioning might consider whether it was attributable to an inability to form stimulus–reinforcement associations, for example. Treatments inspired by cognitive neuroscience might either focus on pharmacological means to improve the impairment in stimulus–reinforcement as-

sociations or cognitive therapies that might teach the patient methods that might bypass their impairment.

Before I continue into the main body of the paper, I briefly consider the advantages of the cognitive neuroscience approach. The neuroscience element of the cognitive neuroscience approach involves the specification of the neural systems involved in a particular type of processing. From a developmental psychopathology perspective, it also involves speculation regarding which neural systems are dysfunctional in a particular disorder. This is important with respect to potential treatments because it allows principled hypotheses regarding which pharmacological treatments might work. If we believe regions X and Y are dysfunctional in patients with a particular disorder and we know that neurotransmitter innervates these regions, a pharmacological agent that increases the levels of the neurotransmitter might be a useful candidate treatment.

The cognitive element of the cognitive neuroscience approach involves the specification of the cognitive functions that are reliant on these neural systems. It is important to note here that there is not a one to one mapping of cognitive function to neural region. Cognitive functions are achieved through different neural regions interacting with one another. Which regions a region interacts with will determine the cognitive function performed. For example, specific forms of instrumental learning rely on the amygdala interacting appropriately with the caudate while aversive conditioning requires the interaction of the amygdala and specific brainstem regions (see below).

The absence of a one to one mapping of cognitive function to neural region means that we must take care with respect to the pharmacological treatment speculations offered above. For example, the orbital frontal and ventrolateral frontal cortex are involved in response gating as a function of contingency change (i.e., in response reversal tasks) and task demands (i.e., response control tasks such as Go/No-Go and Stop-Signal tasks; Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Blair, 2004; Casey, Forman, Franzen,

Berkowitz, Braver, Nystrom, Thomas, & Noll, 2001; Cools, Clark, Owen, & Robbins, 2002; Rogers, Owen, Middleton, Williams, Pickard, Sahakian, & Robbins, 1999). Manipulations that reduce serotonin levels have been reported to disrupt performance on response reversal tasks (Rogers, Blackshaw, et al., 1999) but not response control tasks (Crean, Richards, & de Wit, 2002; Rowley, Van, Mortimore, & Connell, 1997; Schmitt, Jorissen, Sobczak, van Boxtel, Hogervorst, Deutz, & Riedel, 2000). Serotonergic manipulations do alter the neural response within ventrolateral prefrontal cortex during the performance of response control tasks. However, they do not alter behavioral performance (Anderson, Clark, Elliott, Kulkarni, Williams, & Deakin, 2002). In other words, pharmacological agents may affect the functioning of a region such as ventrolateral prefrontal cortex that is involved in more than one cognitive function (e.g., response reversal and response control; see Blair, 2004) but only affect the operation of one of these cognitive functions. This means, on the basis of the evidence presented above, that it would be unwise to offer a serotonergic treatment for a disorder associated with impairment in response control but not response reversal even if we suspected disturbance in ventrolateral prefrontal cortex in that disorder.

In addition to its implications for treatment, the cognitive neuroscience approach also has implications for diagnosis. Current psychiatric diagnoses, as articulated by *DSM-IV* (American Psychiatric Association, 1994), are behavioral syndromes, which are clusters of atypical behaviors that appear to group together. This syndromal approach was once also common in neurology. However, neurology now, for the most part, approaches disorders with respect to their putative causality (e.g., Huntington disease, caused by a faulty gene on chromosome 4).

A major problem with the behavioral approach is that it increases the risk that different disorders will be grouped together because they all appear to increase the probability of a particular behavioral problem. The *DSM-IV* diagnoses of conduct disorder (CD) and antisocial personality disorder (APD) group individuals together who are associated with an

increased risk of aggression. However, there are two main forms of aggression: instrumental and reactive. Instrumental aggression (also referred to as proactive aggression) is aggression that is purposeful and goal directed. The aggression is used instrumentally to achieve a specific desired goal such as obtaining the victim's possessions or to increase status within a group hierarchy (Berkowitz, 1993). In contrast, reactive aggression (also referred to as affective aggression) occurs when a frustrating or threatening event triggers the aggressive act and frequently also induces anger (Barratt, Stanford, Dowdy, Liebman, & Kent, 1999; Barratt, Stanford, Kent, & Felthous, 1997; Berkowitz, 1993; Crick & Dodge, 1996; Linnoila, Virkkunen, Scheinin, Nuutila, Rimon, & Goodwin, 1983). These forms of aggression are mediated by different neurocognitive systems that can become dysfunctional in a variety of ways (Blair, 2001). In other words, there is a wide variety of different pathologies that can give rise to increased aggression and, in consequence, a diagnosis of CD/APD. However, using the same treatment for all these different pathologies because they all received the same diagnosis is only going to mean that many patients do not receive a treatment that might actually be beneficial to them. In contrast, by understanding the nature of the neurocognitive systems that allow successful functioning and how these systems can become dysfunctional we can design treatments to benefit the individual.

The goal of this paper is to consider the implications of applying a cognitive neuroscience perspective to the understanding of the developmental disorder of psychopathy. Psychopathy is a developmental disorder that presents across the lifespan (Harpur & Hare, 1994); children and adults with the disorder present with similar symptomatology. The syndrome is marked by both affective–interpersonal (e.g., such as lack of empathy and guilt) and behavioral components (e.g., criminal activity and poor behavioral controls; Frick, O'Brien, Wootton, & McBurnett, 1994; Harpur, Hare, & Hakstian, 1989). The lack of empathy for victims and guilt for the transgressions committed is the really remarkable feature of psychopathy. As will be con-

sidered below, there are several developmental routes to conditions that are associated with an increased propensity for aggression. However, there is no other psychiatric condition that is associated with a markedly reduced level of guilt following the committing of transgressions that harm others. With respect to the behavioral component, it is worth noting that individuals with psychopathy present with elevated levels of goal-directed instrumental aggression and frustration/threat based reactive aggression (Cornell, Warren, Hawk, Stafford, Oram, & Pine, 1996; Williamson, Hare, & Wong, 1987). This is in contrast to many other clinical conditions, for example, childhood bipolar disorder and intermittent explosive disorder, which are only associated with an increased risk for reactive aggression (Coccaro, 1998; Leibenluft, Blair, Charney, & Pine, 2003). In childhood and adolescence, psychopathic tendencies are identified principally by either the use of the Antisocial Process Screening Device (Frick & Hare, 2001) or by the Psychopathy Checklist: Youth Version (Forth, Kosson, & Hare, in press; Kosson, Cyterski, Steuerwald, Neumann, & Walker–Matthews, 2002). In adulthood, psychopathy is identified though use of the Psychopathy Checklist—Revised (Hare, 1991).

This paper will begin by reviewing four of the earlier models of psychopathy: the frontal lobe dysfunction, the response set modulation (RM), the fear dysfunction and the violence inhibition mechanism (VIM) positions. The majority of these positions are either cognitive or neuroscience accounts; the level of description is at either the cognitive or neural level, and little more than cursory attention is turned to the other level. Some of the shortcomings of these positions, at either the cognitive or neuroscience level, will be considered. Following this, the perspective from a more integrated, cognitive neuroscience perspective, the integrated emotion systems (IES) model, will be considered.

The Frontal Lobe Dysfunction Hypothesis

Frontal lobe and consequent executive dysfunction have long been related to antisocial behavior (Elliot, 1978; Gorenstein, 1982; Mof-

fitt, 1993a; Raine, 1997, 2002). This has led to suggestions that either psychopathy in particular or antisocial behavior more generally is due to frontal lobe dysfunction (Gorenstein, 1982; Moffitt, 1993a; Raine, 2002). In line with the general position, there are considerable data indicating that individuals with antisocial behavior show impaired performance on measures of executive functioning (for reviews of this literature, see Kandel & Freed, 1989; Moffitt, 1993b; Morgan & Lilienfield, 2000). In addition, and also in line with the position, neuroimaging data indicates that aggressive individuals are marked by reduced frontal functioning (Goyer, Andreason, Semple, Clayton, King, Compton–Toth, Schulz, & Cohen, 1994; Raine, Buchsbaum, & LaCasse, 1997; Raine, Buchsbaum, Stanley, Lottenberg, Abel, & Stoddard, 1994; Raine, Meloy, Birhle, Stoddard, LaCasse, & Buchsbaum, 1998; Volkow & Tancredi, 1987; Volkow, Tancredi, Grant, Gillespie, Valentine, Mullani, Wang, & Hollister, 1995). Finally, and also in line with the position, there are considerable data that patients with lesions of frontal cortex, whether these occur early in life or adulthood, present with a heightened risk for aggression (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Blair & Cipolotti, 2000; Burgess & Wood, 1990; Grafman, Schwab, Warden, Pridgen, & Brown, 1996; Pennington & Bennetto, 1993). In short, it appears that here is a situation where individuals with antisocial behavior present with reduced frontal lobe functioning, problems on measures of frontal lobe functioning and there are data that acquired frontal lobe damage leads to a heightened risk of aggression.

The position described above is a neuroscience position; that is, the basic claim is that damage to region X (frontal lobes) leads to a particular form of problematic behavior (antisocial behavior). There have been cognitive neuroscience variants of this position; that is, it has been suggested that frontal lobe damage leads to problems with inhibition (Barratt, 1994; Krakowski, Czobor, Carpenter, Libiger, Kunz, Papezova, Parker, Schmader, & Abad, 1997) or working memory (Pennington & Ozonoff, 1996), and because of this, increases the risk for aggression. However, a tight func-

tional connection between the region, form of processing and the problem behavior has rarely been articulated. This is perhaps not surprising. The scope of the frontal lobe dysfunction position is extremely broad. Frontal cortex corresponds to almost half of the cortex (Fuster, 1980) and has been implicated in a wide variety of putative processes (Baddeley & Della Sala, 1998; Burgess & Shallice, 1996; Luria, 1966; Pennington & Ozonoff, 1996; Roberts, Robbins, & Weiskrantz, 1998).

At the neural level, divisions are made between the dorsolateral, ventrolateral, orbital, and medial frontal cortex. According to the neurological literature, only lesions of orbital and ventrolateral frontal cortex are associated with increased risk of aggression. Lesions of dorsolateral frontal cortex are not (Grafman et al., 1996). This might suggest that a closer look at the neuroimaging literature should reveal reduced activation in orbital and ventrolateral frontal cortex in aggressive individuals but not reduced activation of dorsolateral prefrontal cortex. Unfortunately, the majority of the neuroimaging studies on antisocial individuals have not distinguished between regions of frontal cortex (Raine et al., 1994, 1997, 1998; Volkow et al., 1995; Volkow & Tancredi, 1987). However, one study did distinguish between the regions of frontal cortex. Interestingly, and in line with the neurological literature, this study did identify reduced neural activity in ventrolateral frontal cortex in their aggressive population (Goyer et al., 1994).

The situation is more problematic with respect to the literature examining executive functioning in individuals with antisocial behavior. This literature has tended to concentrate on the use of tasks that index executive functions commonly linked to dorsolateral prefrontal cortex; for example, the Wisconsin Card Sorting Task and the Controlled Oral Word Association Test. In other words, this literature has examined the relationship between aggression and performance on tasks that index the integrity of a neural region, dorsolateral prefrontal cortex, that the neurological literature indicates is irrelevant to an increased risk for aggression. The fact that a relationship between aggression and performance on tasks indexing dorsolateral

prefrontal cortex functioning has been observed is probably due to two reasons. First, CD and attention-deficit/hyperactivity disorder (ADHD) are highly comorbid (Biederman, Newcorn, & Sprich, 1991; Hinshaw, 1987; Taylor, Schachar, Thorley, & Wieselberg, 1986). ADHD is associated with dysfunction of right-sided dorsolateral prefrontal cortex–striatal systems (Giedd, Blumenthal, Molloy, & Castellanos, 2001) and executive dysfunction (Barkley, 1999; Pennington & Ozonoff, 1996). ADHD may be a risk factor for dysfunction that leads to antisocial behavior even if the pathology associated with ADHD itself does not lead to antisocial behavior. Indeed, in line with this suggestion, Pennington and Ozonoff (1996) noted in their review that individuals with CD who were not comorbid for ADHD presented with no indications of executive dysfunction. Second, the degree of executive/dorsolateral prefrontal cortex dysfunction likely predicts the degree of dysfunction beyond dorsolateral prefrontal cortex (i.e., orbital and ventrolateral prefrontal cortex dysfunction).

An additional difficulty for the broad frontal lobe dysfunction positions is that they have not specified the form of the aggression associated with frontal lobe dysfunction; that is, they have not specified whether the model applies to reactive or instrumental aggression or both. However, the neurological data is clear. Frontal lobe damage increases the risk of reactive aggression (Anderson et al., 1999; Blair & Cipolotti, 2000; Burgess & Wood, 1990; Grafman et al., 1996; Pennington & Bennetto, 1993). Currently, there have been no cases where acquired frontal lobe damage has increased the risk of goal-directed instrumental aggression. This issue is important because although some psychiatric disorders are only associated with an increased risk of reactive aggression (e.g., childhood bipolar disorder; Leibenluft et al., 2003) others, such as psychopathy, are associated with an increased risk for both reactive and instrumental aggression (Cornell et al., 1996; Williamson et al., 1987). As frontal lobe dysfunction is only associated with reactive aggression, the frontal lobe dysfunction positions can only be considered relevant to psychiatric conditions associated with

increased reactive aggression and, potentially, the explanation of the increased reactive aggression seen in psychopathy. They cannot be considered a full account of psychopathy, however, as they cannot provide an account of the increased instrumental aggression seen in this population.

Summary

In this section, it has been suggested that the neuroscience position, that frontal lobe dysfunction leads to increased aggression, needs greater specification. The existing data suggests that orbital and ventrolateral frontal cortex dysfunction leads to an increased risk of reactive aggression (Blair, 2001). Of course, this remains a neuroscience position. No claims have been made regarding the functions of orbital/ventrolateral frontal cortex which, when dysfunctional, might increase the risk for reactive aggression. A development of the position, to make the position a cognitive neuroscience account, will be articulated below.

RM Hypothesis

The RM hypothesis has been an influential model of psychopathy (Newman, 1998; Patterson & Newman, 1993). RM involves "a rapid and relatively automatic (i.e., noneffortful or involuntary) shift of attention from the effortful organization and implementation of goal-directed behavior to its evaluation" (Newman, Schmitt, & Voss, 1997). This "brief and highly automatic shift of attention . . . enables individuals to monitor and, if relevant, use information that is peripheral to their dominant response set (i.e., deliberate focus of attention)" (Lorenz & Newman, 2002). This proposed reduced automatic processing in individuals with psychopathy is at the core of Newman's model. Thus, "Whereas most people automatically anticipate the consequences of their actions, automatically feel shame for unkind deeds, automatically understand why they should persist in the face of frustration, automatically distrust propositions that seem too good to be true, and are automatically aware of their commitments to others, psycho-

paths may only become aware of such factors with effort" (Newman, 1998). Newman argues that it is not that individuals with psychopathy are incapable of regulating their behavior, only that self-regulation is more effortful for psychopaths because of the lack of these "relatively automatic processes" to guide actions. The RM hypothesis is an attention-based model. According to the model, "the impulsivity, poor passive avoidance, and emotion-processing deficits of individuals with psychopathy may all be understood as a failure to process the meaning of information that is peripheral or incidental to their deliberate focus of attention" (Lorenz & Newman, 2002).

The RM hypothesis has been associated with an assortment of interesting paradigms. For example, several Stroop type tasks have been developed, some of which have revealed *reduced* interference in individuals with psychopathy (Hiatt, Schmitt, & Newman, 2004; Newman et al., 1997). According to the RM hypothesis, reduced interference in individuals with psychopathy is due to their inability to use the peripheral information when processing the target stimulus. In addition, in recent work with the lexical decision task, where participants are presented with letter strings and must determine whether the string is a word or not, although healthy individuals are faster to respond to emotional/high-frequency words rather than neutral/low-frequency words, individuals with psychopathy are less so (Lorenz & Newman, 2002). According to the RM hypothesis, the absence of such frequency effects on lexical decision performance in individuals with psychopathy is due to their inability to use the peripheral frequency information because of their focus of attention on the dominant response set (deciding whether the stimulus was a word or not). In addition, the RM hypothesis has been used to explain impairment in individuals with psychopathy on some tasks involving reward and punishment (Newman & Kosson, 1986; Newman, Patterson, & Kosson, 1987). For example, in the passive avoidance task, the participant must learn to respond to stimuli associated with reward and avoid stimuli associated with punishment. Individuals with

psychopathy are impaired in learning to avoid the stimuli associated with punishment (Newman & Kosson, 1986). According to the RM hypothesis, the poor performance of individuals with psychopathy on passive avoidance tasks is due to their inability to shift their attention from their goal of responding to gain reward to the peripheral punishment information.

The problem that the RM hypothesis faces is the problem that all models of psychopathology face. Such models are not just accounts of the impairment presented by the patient population being considered. They are also accounts of functioning in healthy individuals. As such, the RM hypothesis, an account of impaired attention in individuals with psychopathy needs to be, at the very least, compatible with contemporary models of attention. The RM hypothesis appears to consider a "spotlight" form of attention model. Attention is directed at (spotlights) task-relevant information but also following "brief and automatic" shifts of attention in healthy individuals spotlights task-irrelevant information. Examples of task irrelevant information include, for a word, its frequency and emotional valence. I would argue that such a view of attention is not clearly consistent with the attention literature (Berger & Posner, 2000; Desimone & Duncan, 1995; Duncan, 1998). This argument has been detailed elsewhere (Blair, Mitchell, & Blair, 2005).

Summary

In this section, it has been suggested that the RM hypothesis, as an attention-based account, needs either to be compatible with current models of attention or an alternative to them. However, when the RM hypothesis is interpreted within current models of attention, it becomes clear that the processing deficits, if they are attentional in nature, appear to be of a variety of different forms and not due to a unitary deficit.

The Fear Dysfunction Hypotheses

One of the main positions regarding the emotional impairment shared by individuals with psychopathy is that there is impairment in the neurophysiological systems modulating fear behavior (Eysenck, 1964; Fowles, 1988; Gray, 1987; Lykken, 1995; Patrick, 1994; Trasler, 1978). For example, Cleckley (1976, p. 340) wrote: "Within himself he appears almost as incapable of anxiety as of profound remorse." The dysfunctional fear positions all assume that moral socialization is achieved through the use of punishment (Eysenck, 1964; Eysenck & Gudjonsson, 1989; Trasler, 1978). In essence, they assume that the healthy individual is frightened by punishment and associates this fear with the action that resulted in the punishment, thus making the individual less likely to engage in the action in the future. The suggestion is that individuals with psychopathy, because they are less aversively aroused by punishment, make weaker associations, and thus are more likely to engage in the punished action in the future than healthy individuals.

In many respects the fear dysfunction hypotheses can be considered the first cognitive neuroscience perspectives on the development of psychopathy. The perspectives of at least several authors existed at both the neural and cognitive levels. For example, Patrick was the first to consider that amygdala dysfunction might be linked to the development of psychopathy (Patrick, 1994). Other theorists, inspired by Gray's work on the Behavioral Inhibition System (BIS; Gray, 1987), considered that septal hippocampal regions might be dysfunctional in individuals with psychopathy (Fowles, 1988; Gray, 1987). At the cognitive level, all made reference to systems modulating fear behavior however few articulated the specifics of these systems in detail. A notable exception to this lack of detail was Gray's BIS model (Gray, 1987; McNaughton & Gray, 2000). This suggested the existence of a unitary fear system, the BIS, which, when activated by signals of punishment and nonreward, novel stimuli and innate fear stimuli, gives rise to behavioral inhibition, increased arousal and increased attention.

The variants of the fear dysfunction hypothesis have generated a considerable body of empirical literature. Indeed, the earliest formal experimental investigations of psychopathy were based around the fear dysfunction

hypothesis (Lykken, 1957). Functions thought to be mediated by the fear systems include: aversive conditioning, generating autonomic responses to anticipated threat, the augmentation of the startle reflex following the presentation of visual threat primes, passive avoidance learning, and response reversal. Individuals with psychopathy present with impairment in aversive conditioning (Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002; Lykken, 1957), generating autonomic responses to anticipated threat (Hare, 1982; Ogloff & Wong, 1990), the augmentation of the startle reflex following the presentation of visual threat primes (Levenston, Patrick, Bradley, & Lang, 2000), passive avoidance learning (Lykken, 1957; Newman & Kosson, 1986), and response reversal/extinction (Mitchell, Colledge, Leonard, & Blair, 2002; Newman et al., 1987).

In addition to these data indicating dysfunction in fear processing in individuals with psychopathy, Kochanska has stressed the role of fearfulness as an important temperamental factor in moral/conscience development (Kochanska, 1993, 1997). Indeed, she and others have found fearful children to show higher levels of moral development/conscience using a variety of measures (Asendorpf & Nunner–Winkler, 1992; Kochanska, 1997; Kochanska, De Vet, Goldman, Murray, & Putman, 1994; Rothbart, Ahadi, & Hershey, 1994).

The difficulty for a cognitive neuroscience account is that it is vulnerable to neural and cognitive level data. With the exception of Patrick's (1994) consideration of the amygdala, most fear dysfunction theorists followed Gray's stress on the importance of the septal hippocampal system. Although the septum and hippocampus play roles in some fear-related functions, in particular in allowing the representation of potential spatial context conditioned stimuli (CS), many of the functions attributed to them by the BIS model have not been empirically supported (LeDoux, 1998; O'Keefe, 1991). For example, most forms of aversive conditioning are not reliant on the integrity of the hippocampus, only on the integrity of the amygdala (Bechara, Tranel, Damasio, Adolphs, Rockland, & Damasio, 1995; LeDoux, 1998).

At the cognitive level, the claim was of a unitary "fear" system. However, such a position has also not been empirically supported. Instead, there are a series of at least partially separable neural systems that are engaged in specific forms of processing that can be subsumed under the umbrella term fear (Blair, 2004). For example, aversive conditioning and instrumental learning are two forms of processing in which the fear system is thought to be involved (Lykken, 1995; Patrick, 1994). Yet, the neural circuitry to achieve aversive conditioning and instrumental learning are doubly dissociable (Killcross, Robbins, & Everitt, 1997). Thus, a lesion to the central nucleus (CeN) of the amygdala will prevent aversive conditioning but still allow instrumental learning to occur. In contrast, a lesion to the basolateral (BLA) nucleus of the amygdala will prevent instrumental learning but still allow aversive conditioning to occur. Moreover, early amygdala lesions result in a massive reduction of neophobia; the infant monkey is no longer fearful of a novel objects. However, the same infant monkeys with amygdala lesions show heightened social phobia; that is, their fear response to another infant monkey is actually heightened (Amaral, 2001). These findings strongly suggest partially separable "fear" systems: for aversive conditioning/instrumental learning and for social threats (Blair, 2004).

Finally, the core developmental assumption of the fear dysfunction positions, that moral socialization is achieved through the use of punishment and the formation of conditioned fear responses has been questioned (Blackburn, 1988; Blair, 2001; Blair, Jones, Clark, & Smith, 1995). Thus, the developmental literature indicates that moral socialization is not achieved through the formation of conditioned fear responses but rather through the induction and fostering of empathy (Hoffman, 1984). Studies have shown, for example, that moral socialization is better achieved through the use of induction (reasoning that draws children's attention to the effects of their misdemeanors on others and increases empathy) than through harsh authoritarian or power assertive parenting practices which rely on the use of punishment (Baumrind, 1971, 1983;

Hoffman & Saltzstein, 1967). Indeed, there have been suggestions that although empathy facilitates moral socialization, fear actually hinders it (Hoffman, 1994). Thus, in a review of a large number of studies of disciplinary methods, it was concluded that punishment-based power assertion had an adverse effect on moral socialization regardless of age (Brody & Shaffer, 1982). Indeed, it has been suggested that the primary utility of power assertion is to prevent the parent from being ignored while the child is transgressing (Hoffman, 1988).

Summary

In this section, it has been suggested that the fear dysfunction hypotheses can be considered the first cognitive neuroscience perspectives on the development of psychopathy. They have been highly successful in generating a considerable body of data. However, data from the broader literature indicates two main ways in which many were incorrect. First, those variants of the hypothesis that stressed the importance of the septal hippocampal system in fear dysfunction were wrong. The septum and hippocampus have been shown to not mediate the functions attributed to them by the older fear positions (LeDoux, 1998; O'Keefe, 1991). Second, those variants that suggested a unitary fear system were also wrong. Data strongly suggest partially separable systems mediating specific functions that fall under the umbrella term fear. Third, they can be considered suspect following their stress on the role of fear in moral socialization.

VIM Model

The importance of empathy for moral socialization was one of the reasons for the development of the original VIM model of psychopathy (Blair, 1995; Blair, Jones, Clark, & Smith, 1997). The VIM model was a cognitive account that was originally designed with respect to typical and atypical moral development. The basic suggestion was the existence of an early developing system, the VIM, which, when activated by distress cues, generates an aversive emotional reaction in observers. It suggested that this system was necessary for moral socialization, and

that this system was dysfunctional in individuals with psychopathy.

There is a body of data consistent with the position. Many social animals, including humans, find the experience of the distress of conspecifics aversive. Thus, both rats and monkeys will learn to make instrumental responses (press levers/pulling chains) that terminate unpleasant occurrences to conspecifics (Church, 1959; Masserman, Wechkin, & Terris, 1964; Rice, 1965; Rice & Gainer, 1962). For example, if a rat learns that pressing a bar will lower another, suspended, rat to the ground (a distressing experience for the suspended rat), the rat will press the lever (Rice & Gainer, 1962). Alternatively, if a rhesus monkey learns that a lever previously associated with greater food reward now also causes another monkey in sight of the test animal to receive an electric shock, the monkey will terminate responding to the lever in favor of other, less rewarded, response options (Masserman et al., 1964).

The distress of another individual is considered aversive by most humans (Bandura & Rosenthal, 1966). Moreover, the presentation of cues indicating another individual's sadness or fear during aggression reduces the probability of future physical aggression (Perry & Perry, 1974), disputes over property ownership (Camras, 1977), and aggressive sexual activity (Chaplin, Rice, & Harris, 1995).

At its simplest, the VIM is thought to be a system that when activated by distress cues, the sad and fearful expressions of other individual, results in increased autonomic activity, attention, and activation of the brainstem threat response system (usually resulting in freezing; Blair, 1995). Through association these representations of moral transgressions become triggers for the mechanism. The appropriately developing child thus initially finds the pain of other individuals aversive, and then, through socialization, thoughts of acts that cause pain to other individuals aversive also. The VIM account allowed an explanation for the moral/conventional distinction; the distinction made by typically developing children between moral (victim based) and conventional (social order based) transgressions from the age of 36 months (Smetana, 1981, 1985, 1993); only moral transgressions

would be associated with the activation of the VIM. The suggestion was that the development of psychopathy is marked by dysfunction within the VIM. This predicted that individuals with psychopathy would be less able to distinguish between moral and conventional transgressions, a prediction that has been confirmed in individuals with psychopathy and related CDs (Arsenio & Fleiss, 1996; Blair, 1995, 1997; Blair et al., 1995; Blair, Monson, & Frederickson, 2001; Blair, Newman, Mitchell, Peschardt, & Leonard, 2005; Nucci & Herman, 1982). The model also predicted the reduced responsiveness of individuals with psychopathy to sad and fearful facial expressions. Research has shown that psychopathic individuals show reduced autonomic responses to the distress of other individuals (Aniskiewicz, 1979; Blair, 1999; Blair et al., 1997; House & Milligan, 1976). Moreover, it has been shown that although "victim" scenes prime up the threat system in healthy individuals such that after this prime, the participant presents with an augmented startle response, this is not the case in individuals with psychopathy (Levenston et al., 2000). In addition, individuals with psychopathy, in child and adulthood, present with impairment even in the naming of sad and particularly fearful facial and vocal affect (Blair, Colledge, Murray, & Mitchell, 2001; Blair, Mitchell, Richell, Kelly, Leonard, Newman, & Scott, 2002; Stevens, Charman, & Blair, 2001).

Although the original VIM model could provide an account of the emergence of instrumental antisocial behavior in individuals with psychopathy, and it did generate a variety of predictions that have been empirically confirmed, it faces a serious difficulty. At best, it can be considered only a very incomplete account of the disorder. Even at the cognitive level, the VIM account could not explain the results generated by the RM and fear hypotheses. It made no neural level predictions at all.

Summary

In short, the VIM model provides an account of the emergence of the instrumental antisocial behavior displayed by individuals with psychopathy. However, it does not provide a complete account of the range of impairments shown by individuals with psychopathy.

The IES Model

None of the positions briefly described above, with the partial exception of some variants of the fear dysfunction position, were cognitive neuroscience accounts. These positions either couched their explanation primarily either at the neural (the frontal lobe dysfunction position) or cognitive levels (RM and the RM position). The IES model, in contrast, was designed as a cognitive neuroscience position from the beginning (Blair, 2004). This model will be described from the specific perspective of an account of psychopathy.

The Amygdala

At the neural level, the basic argument is that specific forms of amygdala dysfunction are the central focus of the pathology associated with psychopathy (Blair, 2004; Blair, Morris, Frith, Perrett, & Dolan, 1999; cf. Patrick, 1994). The amygdala is frequently divided into two parts: the BLA and CeN (Everitt, Cardinal, Parkinson, & Robbins, 2003; LeDoux, 1998). The amygdala is one of the most crucial regions in the neural circuitry that processes emotion, and is at the center of what Joe LeDoux termed "The Emotional Brain" (LeDoux, 1998).

There are considerable indications of amygdala dysfunction in individuals with psychopathy (Blair, 2003b). Functional imaging studies have shown that adults with the disorder present with reduced amygdala activation during emotional memory (Kiehl, Smith, Hare, Mendrek, Forster, Brink, & Liddle, 2001) and aversive conditioning tasks (Veit, Flor, Erb, Hermann, Lotze, Grodd, & Birbaumer, 2002). In addition, individuals with psychopathy present with impairment on a series of tasks that require the functional integrity of the amygdala. Thus, lesions of the amygdala disrupt aversive conditioning (Bechara et al., 1995; LaBar, LeDoux, Spencer, & Phelps, 1995), the augmentation of the startle reflex by visual threat primes (Angrilli, Mauri, Palomba, Flor, Birhaumer, Sartori, & di Paola,

1996), passive avoidance learning (Ambrogi Lorenzini, Baldi, Bucherelli, Sacchetti, & Tassoni, 1999), and fearful expression recognition (Adolphs, 2002; Blair, 2003a). Individuals with psychopathy, show impairment in aversive conditioning (Flor et al., 2002), the augmentation of the startle reflex by visual threat primes (Levenston et al., 2000), passive avoidance learning (Newman & Kosson, 1986), and fearful expression recognition (Blair, Colledge, Murray, & Mitchell, 2001).

The amygdala allows the formation of three types of appetitive and aversive CS association (Everitt et al., 2003):

1. CS–unconditioned response (UR) associations (e.g., salivation to a tone previously associated with food): The CeN, but not the BLA, is necessary for the formation of CS-UR associations (Everitt et al., 2003; Killcross et al., 1997).
2. CS–affect representation associations: The suggestion is one of "an emotional 'tone' . . . *[fear or the expectation of reward]* . . . that is tagged to a stimulus" (Everitt et al., 2003; p. 234, italics added). The BLA, but not the CeN, is necessary for the formation of these associations.
3. CS–valenced sensory properties of the unconditioned stimulus (US) associations: Here, the CS is associated with specific sensory (e.g., visual appearance, sound and smell) and "consumatory" (e.g., taste) of the US. The BLA, but not the CeN, is necessary for the formation of these associations. Reinforcer devaluation studies indicate that these associations are not stored in the amygdala (Pickens, Saddoris, Setlow, Gallagher, Holland, & Schoenbaum, 2003). The suggestion here is that they are stored within the insula (e.g., Blair et al., 2005).

At the cognitive level, the basic claim is that individuals with psychopathy are impaired in the formation of the three types of CS association described above; referred to henceforth as stimulus–reinforcement associations (Blair, 2003b, 2004). Importantly, this position allows an integration of the earlier fear and empathy (VIM) dysfunction posi-

tions. The claim that individuals with psychopathy are impaired in aversive CS association is clearly compatible with the fear dysfunction position. Impairment in the formation of aversive stimulus–reinforcement associations would give rise to the observed deficits in individuals with psychopathy in aversive conditioning (Flor et al., 2002; Lykken, 1957), the augmentation of the startle reflex following the presentation of visual threat primes (Levenston et al., 2000), and passive avoidance learning (Lykken, 1957; Newman & Kosson, 1986; see Blair, 2004).

With respect to the empathy/VIM model, the argument is that one class of aversive stimuli is the distress of other individuals; the expressions of fear and sadness (Blair, 1995, 2003b). The argument has been made that these expressions serve as social US allowing conspecifics to teach the societal valence of objects and actions to the developing individual (Blair, 2003a). Due to their impairment in the formation of aversive stimulus–reinforcement associations, individuals with psychopathy are less able to take advantage of this "moral" social referencing; they are more difficult to socialize. In short, the position also allows the explanation of impairment in individuals with psychopathy in responsiveness to distress cues (Aniskiewicz, 1979; Blair, 1999; Blair et al., 1997; House & Milligan, 1976), fearful facial and vocal expression recognition (Blair et al., 2002; Blair, Colledge, Murray, & Mitchell, 2001), and the processing of the moral/conventional distinction (Arsenio & Fleiss, 1996; Blair, 1995, 1997; Blair et al., 1995, 2005; Blair, Monson, & Fredrickson, 2001; Nucci & Herman, 1982). Moreover, it becomes easier to understand findings indicating the importance of "fearfulness" as an important temperamental factor in socialization (Kochanska, 1993, 1997). It is not clear what role fear has in moral socialization; certainly, and in contrast, to early positions (Eysenck, 1964), aversive conditioning has no obvious role in socialization (Brody & Shaffer, 1982; Hoffman, 1994). However, if we understand the temperamental variable "fearfulness" as an index of the integrity of the amygdala (Blair, 2003c), its role in socialization becomes clear. Fearfulness indexes the integrity of the neural

system necessary for empathy induction and the moral social referencing described above.

Concerning the earlier models, the IES account can be considered an extension of the fear dysfunction and VIM models. Indeed, the VIM model has developed into the IES account. However, there are three important differences between the IES and fear dysfunction positions. First, within the IES account there is no unitary "fear" system, but rather a series of systems that can function in integrated ways to achieve particular types of processing goals. For example, there is a clear dissociation between the role of the CeN in CS-UR association formation as opposed to the role of the BLA in CS-affect representation association formation (Everitt et al., 2003).

Second, and a related point to the first, the IES account does not predict that all punishment-based processing in individuals with psychopathy is disrupted. Due to the neuroscience roots of the account a dichotomy is made between aversive stimulus reinforcement association formation and stimulus–response association formation on the basis of punishment information. Importantly, although the functional integrity of the amygdala is necessary for the formation of stimulus–reinforcement associations, it is *not* necessary for the formation of stimulus–response (CS-CR) associations; that is, the individual learns a response to a stimulus as a function of reward/punishment information. Lesions of the amygdala do not disrupt the formation of stimulus response associations (Baxter & Murray, 2002).

Some instrumental learning tasks are reliant on the formation of stimulus–reinforcement associations; for example, passive avoidance learning. In passive avoidance learning, the participant is presented with stimuli. Some stimuli, if responded to, engender reward while others, if responded to, engender punishment. The participant's task is to learn to respond to the "good" stimuli and avoid responding to the "bad" stimuli. If the individual has formed a CS-positive affect association, the individual will approach (respond to) this stimulus. If the individual has formed a CS-negative affect association, the individual will avoid (fail to respond to) this stimulus. In line with this

position, individuals with psychopathy are impaired in passive avoidance learning (Lykken, 1957; Newman & Kosson, 1986).

Other instrumental learning tasks cannot be solved through stimulus–reinforcement association formation and must be solved through the formation of stimulus–response associations. For example, object discrimination learning involves learning to respond to one of two objects (one rewarded and one not rewarded) repeatedly presented in a pairwise fashion over a series of trials. In other words, the participant must learn that when Stimulus A and Stimulus B are present they should respond towards Stimulus A (Baxter & Murray, 2002). In object discrimination tasks, and unlike passive avoidance learning tasks, the participant cannot learn that some of the stimuli are good or bad and should therefore be approached or avoided. In object learning tasks, the compound stimulus (A plus B) can be good or bad; what determines whether it is or not is not the quality of the stimulus (this is always repeated) but the quality of the response made to the stimulus. In line with this, individuals with psychopathy show no difficulty on object discrimination learning tasks (Mitchell et al., 2002).

It is interesting that some data that were highly critical of the fear dysfunction position become easily explained through reference to the IES account. Newman (1998) developed a punishment-only version of the passive avoidance paradigm. In punishment-only versions of the passive avoidance task, some stimuli, if *not* responded to, engender punishment whereas others, if responded to, engender punishment. Newman showed that individuals with psychopathy were intact on this task and correctly argued that this result is not compatible with the older fear dysfunction positions (Newman, 1998). However, it is important to note that punishment-only task variants of the passive avoidance task cannot be solved through the formation of CS-affect representation associations. In these variants of the passive avoidance task, there are no good or bad stimuli; both $S + s$ and $S - s$ can give rise to punishment. Instead of forming a stimulus–reinforcement association, the participant must form a stimulus–response association; If $S+$,

do R1 (respond to avoid punishment); if S−, do R2 (respond differently to avoid punishment). In short, given the reliance of punishment-only versions of the task on stimulus–response associations, they should be from the IES perspective (and are) solvable by individuals with psychopathy.

Third, a difference between the IES and earlier fear dysfunction accounts is that the fear dysfunction accounts assumed either that the processing of reward information was intact or even that the processing of reward information might be enhanced in individuals with psychopathy (Fowles, 1988; Levenston, Patrick, Bradley, & Lang, 1996). In contrast, the cognitive neuroscience-based IES account might assume at least some impairment in the formation of appetitive stimulus–reinforcement associations in individuals with psychopathy. This is because the amygdala is known to be involved in both the processing of reward and punishment information (Baxter & Murray, 2002; Everitt et al., 2003).

Although it is clear that individuals with psychopathy do present with impairment in the formation of stimulus–punishment associations/reduced representation of aversive stimuli, the extent to which individuals with psychopathy present with impairment in the formation of stimulus–reward associations/reduced representation of appetitive stimuli is less clear (Levenston et al., 2000; Peschardt, Leonard, Morton, & Blair, 2005; Peschardt, Morton, & Blair, 2005). Individuals with psychopathy show appropriate suppression of the startle reflex following the presentation of positive visual primes but reduced augmentation of the startle reflex following the presentation of negative visual primes (Levenston et al., 2000; Pastor, Molto, Vila, & Lang, 2003; Patrick, Bradley, & Lang, 1993). This suggests that individuals with psychopathy are unimpaired in processing positive material. However, in lexical decision-making tasks where participants must identify words versus nonwords, comparison individuals are faster to identify positive and negative emotional words than neutral ones, but individuals with psychopathy show a significantly reduced emotional advantage (Lorenz & Newman, 2002; Williamson, Harpur, & Hare,

1991). In addition, Verona and colleagues reported reduced skin conductance responses to both positive and negative auditory stimuli in individuals with psychopathy (Verona, Patrick, Curtin, Bradley, & Lang, 2004). Finally, in recent work within our own group, using both affective priming (Peschardt, Morton, et al., 2005) and decision-making paradigms (Peschardt, Leonard, et al., 2005), we have found impaired processing of both positive and negative material, but that this impairment is particularly severe for negative material. Our assumption is that appetitive stimulus–reinforcement association formation is impaired but less impaired than aversive stimulus–reinforcement association formation (Blair, 2004; Peschardt, Leonard, et al., 2005). This difference may reflect fundamental neurotransmitter impairments in the pathology of the disorder (Blair, 2003b; Peschardt, Leonard, et al., 2005). However, this issue will not be considered further in this paper. It is interesting that, given this claim that stimulus–reward association formation is less impaired in individuals with psychopathy than stimulus–punishment association formation, Kochanska has reported data indicating that conscience development in "fearless" children, is best achieved by socialization practices that presumably capitalizing on mother–child positive orientation (secure attachment, maternal responsiveness; Kochanska, 1997).

It is important to note here that the IES model is a cognitive neuroscience position. The claim being made here is not that the amygdala is dysfunctional in individuals with psychopathy. Rather, it is that the formation of (aversive and to a less extent appetitive) stimulus–reinforcement associations, a function reliant on the amygdala, is impaired in individuals with psychopathy. The model predicts that individuals with psychopathy should present with impairment on any task reliant on the amygdala's role in the formation of stimulus–reinforcement associations. It does not necessarily predict that individuals with psychopathy will show impairment on any task that requires the functional integrity of the amygdala.

Additional functions to require the functional integrity of the amygdala relate to

other aspects of social cognition, in particular, affect-related judgments of facial stimuli (Adolphs, 2003; Baron–Cohen, Ring, Bullmore, Wheelwright, Ashwin, & Williams, 2000). Data indicate that patients with amygdala lesions present with different judgments of the trustworthiness of other's faces from healthy individuals (Adolphs, Tranel, & Damasio, 1998). In addition, a recent neuroimaging work indicates that healthy individuals show greater amygdala activation to faces judged to be untrustworthy relative to faces judged to be trustworthy (Winston, Strange, O'Doherty, & Dolan, 2002). Patients with amygdala lesions also show impairment when determining which complex social emotion is being displayed by an individual when they have information from the eye region only (Adolphs, Baron–Cohen, & Tranel, 2002; Stone, Baron–Cohen, Calder, Keane, & Young, 2003). Neuroimaging work compliments these findings by demonstrating amygdala activation during the performance of this task (Baron–Cohen, Ring, Wheelwright, Bullmore, Brammer, Simmons, & Williams, 1999). However, despite this apparent role of the amygdala in these two aspects of social cognition, individuals with psychopathy do not present with impairment in either the making of trustworthiness judgments (Richell, Mitchell, Newman, Leonard, Baron–Cohen, & Blair, 2003) or the judging of complex social emotions from the eyes (Richell, Mitchell, Peschardt, Winston, Leonard, Dolan, & Blair, 2005).

At present, no formal model of these aspects of social cognition has been proposed. There have been no attempts to tie these aspects of social cognition to the known role of the amygdala in the formation of stimulus–reward and stimulus–punishment associations. It is even conceivable that these aspects of social cognition do not involve the amygdala per se, but rather cortex adjacent to the amygdala or even fiber tracts that pass through the amygdala. However, if we assume that they do involve the amygdala, then it is clear that not all aspects of amygdala functioning are impaired in individuals with psychopathy. The current data only indicate that the formation of stimulus–reinforcement associations are impaired in psychopathy, and even in this case,

the impairment in the formation of stimulus–punishment associations is far more marked than that for the formation of stimulus–reward associations.

Orbital and Ventrolateral Frontal Cortex

As noted above with respect to the frontal lobe dysfunction position, there are considerable data indicating that orbital and ventrolateral frontal cortex are involved in the modulation of reactive aggression. Indeed, the orbital and ventrolateral frontal cortex is considerably involved in the regulation and mediation of emotional behavior. The orbital frontal cortex regulates the neural systems that mediate the basic responses to threat stimuli (i.e., the amygdala, hypothalamus and periaqueductal gray; Gregg & Siegel, 2001; Panksepp, 1998).

The link between orbital frontal cortex damage and increased levels of aggression has led to claims that the disorder of psychopathy is due to orbital frontal cortex dysfunction (Anderson et al., 1999; Damasio, 1994). However, as noted above, patients with orbital frontal cortex damage, if they present with aggression, present only with reactive aggression whether the lesion occurs in child- or adulthood (Anderson et al., 1999; Blair & Cipolotti, 2000; Grafman et al., 1996; Pennington & Bennetto, 1993). There has never been a recorded case of a patient where orbital/ventrolateral frontal lobe damage has been associated with an increase in the incidence of instrumental aggression. Although individuals with psychopathy do present with an increased incidence of reactive aggression, their defining feature is their increased presentation of instrumental aggression (Cornell et al., 1996; Williamson et al., 1987). In addition, many of the results discussed above and interpretable with reference to amygdala pathology cannot be explained with reference or orbital/ventrolateral frontal lobe pathology (Blair, 2004). For example, individuals with psychopathy present with impaired aversive conditioning (Flor et al., 2002; Lykken, 1957). Patients with orbital/ventrolateral frontal lobe pathology, even those presenting with profound behavioral disturbance (those Damasio,

1994, termed presenting acquired sociopathy), do not present with impaired aversive conditioning (Bechara, Damasio, Damasio, & Lee, 1999).

This is not to suggest that there are no indications of orbital/ventrolateral frontal cortex dysfunction in individuals with psychopathy. Animal and human lesion studies, as well as recent functional imaging studies, all strongly indicate a role of orbital/ventrolateral frontal cortex in response reversal and extinction (Cools et al., 2002; Rolls, 1997; Rolls, Hornak, Wade, & McGrath, 1994). Both response reversal and extinction involve changing the response to a stimulus following a change in contingency. From the perspective of the IES model, ventromedial regions code expectancy of reinforcement and allow rapid decision making on the basis of this information while ventrolateral regions gate response choice following a change in contingency (Blair, 2004).

Individuals with psychopathy show marked problems in response reversal/extinction (LaPierre, Braun, & Hodgins, 1995; Mitchell et al., 2002; Newman et al., 1987). This appears to be related to a reduced sensitivity to temporal difference errors; the difference between the expected reward and received reward (O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003). The impairment shown by individuals with psychopathy becomes more marked, the more subtle the temporal difference error to be detected (Budhani & Blair, 2005). In short, there are reasons to believe that individuals with psychopathy do present with a specific functional impairment related to the integrity of orbital/ventrolateral frontal cortex.

Interestingly, adults with psychopathy are notably more impaired on response reversal and extinction tasks than children with psychopathic tendencies (Blair, Colledge, & Mitchell, 2001; Fisher & Blair, 1998; Mitchell et al., 2002; Newman et al., 1987; O'Brien & Frick, 1996). For example, while adults with psychopathy and children with psychopathic tendencies are both notably impaired on Newman's card-playing task (Newman et al., 1987) only adults with psychopathy show impairment on the reversal state of the Intradimensional–Extradimensional (ID-ED) Task (Dias, Robbins, & Roberts, 1996).

In Newman's card-playing task, the participant has to decide whether to play a card. Initially, the participant's choice to play is always reinforcing; if the participant plays the card he/she will win points or money. However, as the participant progresses through the pack of cards, their probability of reward decreases. The participant should terminate his/her responding before he/she receives greater levels of punishment than reward. In the ID-ED task (Dias et al., 1996), the participant learns that responding to one of two stimuli gains reward while responding to the other is punished. This contingency is then reversed; that is, responding to the first stimulus is no longer rewarded but punished while responding to the second is now rewarded.

We have argued that the principal difference between these two tasks is the salience of the contingency change. In the card-playing task, the probability of reinforcement decreases by 10% over every 10 trials. In the ID-ED task, the probability of reinforcement changes from 100 to 0% once the initial learning criterion has been achieved. Dayan and colleagues have stressed the importance of prediction errors in emotional learning (O'Doherty, Dayan, Schultz, Deichmann, Friston, & Dolan, 2004; Schultz, Dayan, & Montague, 1997; Sutton & Barto, 1981). The prediction error is the difference between the expected value associated with a stimulus/action and the actual value currently received with respect to that stimulus/action. In other words, unexpected rewards induce large positive prediction errors (initiating rapid learning). Absent highly expected rewards induce large negative prediction errors (initiating response reversal/extinction). Clearly, for salient contingency changes (i.e., 100–0 in the reversal phase of the ID-ED task; i.e., if the response to a stimulus is always rewarded and then never rewarded), the negative prediction error is large, and children with psychopathic tendencies can alter their behavior (although adults with psychopathy show impairment). For more subtle contingency changes (i.e., those present in the four and one pack card playing tasks), the negative temporal difference error is smaller, and both adults with psychopathy and children with psychopathic tendencies are less able to alter their behavior to this stimulus.

We have recently tested this hypothesis using a probabilistic response paradigm involving different stimulus pairs of varying contingency. We found that the impairment in adults was far greater than that seen in children with psychopathic tendencies. However, differences between both groups with psychopathic tendencies and their respective comparison groups increased as the salience of the contingency change decreased (Budhani & Blair, in press; Budhani, Richell, & Blair, 2005).

Of course, it is important to note that this difficulty with response reversal is not unique to individuals with psychopathy. Patients with intermittent explosive disorder and childhood bipolar disorder also present with difficulties in response reversal (Best, Williams, & Coccaro, 2002; Gorrindo, Blair, Budhani, Pine, & Leibenluft, 2005). Nor is an elevated risk of reactive aggression unique to psychopathy (although an increased risk of instrumental aggression is). Patients with intermittent explosive disorder and childhood bipolar disorder also present with significantly increased incidences of reactive aggression (Coccaro, 1998; Leibenluft et al., 2003).

Frustration has long been linked to the display of reactive aggression (Berkowitz, 1993). Frustration occurs following the initiation of a behavior to achieve an expected reward and the subsequent absence of this reward. The suggestion has been made that ventromedial regions code expectancy of reinforcement and identify contingency changes whereas ventrolateral regions gate response choice following a detected change in contingency (Blair, 2004). In short, damage to these regions will profoundly alter the individual's ability to initiate behaviors appropriately to achieve expected rewards. In particular, if contingencies change, damage to these regions will give rise to a situation where the individual consistently initiates a behavior to achieve an expected reward, and this reward does not occur (the essence of a response reversal trial). The basic suggestion then is that if orbital/ventrolateral frontal cortex dysfunction disturbs the computational systems that allow rapid response reversal, the individual will be predisposed to frustration-based reactive aggression. Importantly, orbital/ventrolateral frontal cortex dysfunction is not

unique to psychopathy but common to a series of psychiatric conditions associated with an elevated risk of reactive aggression.

It is interesting to relate the above to the concept of "effortful control" introduced by Rothbart and her colleagues (Posner & Rothbart, 2000; Rothbart & Ahadi, 1994). They proposed this construct to denote a class of self-regulatory mechanisms that emerge around 6–12 months of age. They defined effortful control as the ability to suppress a dominant response to perform a subdominant response. Effortful control has been found to be positively associated with "fearfulness" and conscience development (Kochanska & Knaack, 2003). This is an interesting echo of work with individuals with psychopathy where there is impairment in both fearfulness (stimulus–reinforcement association due to amygdala dysfunction), conscience development and effortful control (the ability to perform response reversal reliant on orbital frontal cortex, ventrolateral frontal cortex).

The Development of Psychopathy

In the beginning of this paper, I made reference to those who have advocate that developmental disorders can only be understood following multiple levels of analysis (Cicchetti & Dawson, 2002; Kopnisky et al., 2002). In this section, I wish to sketch a specifically developmental model of psychopathy across these levels of analysis (gene, molecular, systems, cognitive and behavioral).

Growing evidence suggests a genetic contribution to psychopathy. It is important that the claim here is that there is a genetic contribution to the emotional disorder that is at the heart of psychopathy. The claim is not that there is a genetic contribution to antisocial behavior directly; given that we learn to use weapons/objects to commit antisocial behavior, it is unclear how there could be. Instead, the claim is that there is a genetic contribution to a specific form of emotional responsiveness that may put the individual at risk for learning antisocial behaviors.

Two recent studies have examined the heritability of psychopathy (Blonigen, Carlson, Krueger, & Patrick, 2003; Viding, Blair, Mof-

fitt, & Plomin, in press). Blonigen et al. (2003) collected data from 353 adult male twins using the self-report Psychopathic Personality Inventory (Lilienfeld & Andrews, 1996). This inventory forms a global index of psychopathy with eight subscales, most of which showed moderate heritability ($h^2 = 0.29$–0.56) and negligible shared environmental influence. Viding et al. (in press) examined the callous and unemotional component of psychopathic tendencies within almost 3,500 twin pairs at age 7 within the Twins Early Development Study. This study revealed a significant group heritability of $h_g^2 = 0.67$ and no shared environmental influence on the callous–unemotional component; that is, genetic factors account for two thirds of the difference between the callous–unemotional probands and the population. In short, both studies suggested a genetic contribution to psychopathy.

Genes have their impact at the molecular level. However, currently there is no molecular level account of psychopathy. In short, although it appears that there is a genetic contribution to the development of the disorder, how this contribution is achieved remains unknown. Suggestions can be offered with respect to specific neurotransmitter systems (Blair, 2003b). However, at present, such suggestions exist in a near absence of data.

What about environmental influences? The studies of Blonigen et al. (2003) and Viding et al. (in press) suggested little environmental impact. However, it is a common lay impression that psychopathy may be due to early stressors such as physical or sexual abuse. Moreover, there is a considerable scientific literature demonstrating the impact of early stressors on the development of neural systems involved in the basic response to threatening stimuli. Stressors in early life have profound and long-term effects on hypothalamic–pituitary–adrenal axis activity (Bremner & Vermetten, 2001; Charney, 2003). During infancy animals do not demonstrate hypothalamic–pituitary–adrenal axis responses to stress. However, infant animals exposed to stressors demonstrate increases in immediate early genes (e.g., c-*fos* and nerve growth factor inducible gene) in the paraventricular nucleus of the hypothalamus (Smith, Kim, van

Oers, & Levine, 1997). Chronic stress is also associated with potentiated release of norepinephrine following exposure to subsequent stressors (Nisenbaum, Zigmond, Sved, & Abercrombie, 1991) and a general lifelong increase in the sensitivity of the noradrenergic system (Francis, Caldji, Champagne, Plotsky, & Meaney, 1999). In short, environmental stressors can alter the development of systems that mediate the basic response to threat. However, their impact is to increase the responsiveness of the system. As yet, there are no data that environmental stressors can lead to the type of suppression of emotional responsiveness that is seen in individuals with psychopathy. In other words, currently there is no reason to believe that there could be a social cause to psychopathy (although this does not imply that environmental factors play no role in determining the nature of the pathology seen in any given individual; see below).

Although there is no molecular level account of psychopathy, the data presented above suggest that the genetic contribution is manifested as reduced amygdala responsiveness. Of course, any genetic contribution is unlikely to affect only one system, and while I did also present evidence of orbital frontal cortex/ventrolateral prefrontal cortex dysfunction, it is unknown whether the genetic contribution influences regions beyond these systems. On the basis of neuroimaging data, Kiehl (in press) has argued that there is dysfunction in individuals with psychopathy within paralimbic cortex (i.e., amygdala, anterior superior temporal gyrus, rostral and caudal anterior cingulate, posterior cingulate, ventromedial frontal cortex [orbital frontal cortex] and parahippocampal regions). However, neuroimaging data is notoriously unable to localize deficits; impairment in any region will lead to anomalous activity in any region reliant on the dysfunctional region for input. In addition, the anterior cingulate, at least, does not appear globally impaired in individuals with psychopathy. Damage to the anterior cingulate is known to disrupt performance on the Stroop task, for example (Swick & Jovanovic, 2002; Stuss et al., 2001). However, individuals with psychopathy show no indications of impairment on Stroop tasks (Hiatt

et al., 2004; Peschardt et al., in press; Smith, Arnett, & Newman, 1992).

I argued above that the functional significance of the amygdala dysfunction in individuals with psychopathy was that it interfered with the ability to form stimulus–reinforcement associations in individuals with psychopathy. I also argued above that the developmental consequence of this impairment was that it interfered with moral socialization; the ability of the individual with the disorder to learn the "badness" of moral transgressions is profoundly reduced.

It is worth considering environmental influences again now. As stated above, there are currently no reasons to believe that there might be an environmental cause to psychopathy. However, there is a considerable literature indicating a relationship between socioeconomic status (SES) and antisocial behavior (Raine, 1993). It would be surprising if social variables did not impact on the probability of antisocial behavior; SES, for example, is likely to constrain the possibility of alternative behavioral choices to antisocial behavior as well as increase the salience of the money contained in a potential victim's wallet. Indeed, in line with this, a relationship between SES and the antisocial behavior component of psychopathy has been reported (Hare, 2003). In other words, to develop the full disorder, the individual may have to be influenced by the genetic factors such that his/her emotional responsiveness (and specifically, ability to form stimulus–punishment associations) is reduced and also by environmental factors such that he/she regards antisocial behaviors as a potential way of achieving his/her goals.

Implications of Psychopathy Research for Understanding Typical Development

In the previous section, I briefly sketched a developmental model of psychopathy. In this section, I consider some of the implications that the findings obtained with this population can have for the understanding of typical development. Psychopathy represents an "experiment of nature." Cicchetti and others have argued for that the study of experiments of nature is important because they allow the

isolation of "the components of the integrated system" and that "investigation of these natural experiments sheds light on the normal structure of the system" (Cicchetti, 2003).

The genetic and environmental influence data discussed above suggest that emotional responsiveness (in particular, the ability to perform the forms of emotional learning necessary for moral socialization, i.e., the temperamental variable Kochanska [1993, 1997] terms "fearfulness") in typically developmentally children is likely to be under considerable genetic influence. Environmental influences, at least in respect to early exposure to stressors, are likely to contribute to increased emotional responsiveness/fearfulness in children, particularly children who are predisposed to be emotional responsive because of genetic factors.

In the work with individuals with psychopathy it is also suggested that the temperamental variable fearfulness is likely to be heavily reliant of the integrity of the amygdala. Following the arguments developed above concerning the amygdala's importance in instrumental learning reliant on stimulus–reinforcement learning but not in instrumental learning reliant on stimulus–response learning, we might also consider whether socialization strategies might be adapted, for those children showing reduced fearfulness, which would take advantage of their intact ability to form stimulus–response associations.

The research with individuals with psychopathy also suggests that the concept of "effortful control" might be usefully refined (Kochanska & Knaack, 2003; Posner & Rothbart, 2000; Rothbart & Ahadi, 1994). Currently, it is defined as the ability to suppress a dominant response to perform a subdominant response (Posner & Rothbart, 2000; Rothbart & Ahadi, 1994). A variety of tasks have been used to index "effortful control," but typically these have been response control tasks (i.e., tasks where the participant modulates behavior due to task demands) such as variants of Stroop and Go/No-Go tasks (Kochanska & Knaack, 2003). The work with psychopathy would suggest that more crucial measures should be based on the response reversal paradigm. Although ventrolateral prefrontal cortex is involved in mediating both response

control and response reversal paradigms, these two functions can be dissociated pharmacologically (Crean et al., 2002; Rogers, Blackshaw, et al., 1999; Rowley et al., 1997; Schmitt et al., 2000). It is the impairment in response reversal that is most marked in individuals with psychopathy (Mitchell et al., 2002).

As an experiment of nature, psychopathy is important for what it tells us about what is not developmentally reliant on the amygdala/the ability to form stimulus–reinforcement associations. Neurocognitive functions that are unimpaired in individuals with psychopathy can be considered developmentally independent of the integrity of the amygdala/the ability to form stimulus–reinforcement associations (at least, if we assume that there is no selective compensation occurring). Individuals with psychopathy, unlike other psychiatric disorders, are not impaired in a wide variety of cognitive domains. There are no indications of general executive dysfunction (LaPierre et al., 1995; Mitchell et al., 2002; Peschardt et al., in press), memory. or language impairment (Hart, Forth, & Hare, 1990), although there may be some specific difficulties with abstract words (Kiehl, Hare, McDonald, & Brink, 1999; Kiehl, Smith, Mendrek, Forster, Hare, & Liddle, 2004). Individuals with psychopathy may be impaired in specific functions of orbital/ventrolateral frontal cortex that may be reliant on input from the amygdala; for example, response reversal (LaPierre et al., 1995; Mitchell et al., 2002; Peschardt et al., in press). However, as stated above, it is unclear that they are impaired on all functions of ventrolateral prefrontal cortex; specifically, response control on the basis of task demands as indexed by Go/No-Go and Stop tasks.

Computational Modeling and Additional Deficits

In this final section of the paper, two additional points will briefly be considered: the potential future impact of computational modeling, and additional deficits that cannot be accounted for with respect to the IES model. Computational modeling has been one of the more recent planks of particularly the cognitive part of the cognitive neuroscience revolution, emerging most saliently in the seminal work of Rumelhart and McClelland (1986). Its advantage is that it forces precise specification of theoretical accounts, avoiding the vagueness that is almost inevitable with a verbal description.

Thus far there has been little formal modeling work with respect to psychopathy, and most accounts have remained verbal descriptions. Gray specified a functional architecture for the BIS using the visual "box and arrow" form of description (Gray, 1987). More recently, putative computational details have been provided with respect to the IES model (Blair, 2004) with at least some components of the IES model tentatively implemented (Blair et al., 2005). However, this formalization of the models of psychopathy remains in its infancy.

One of the important features of a cognitive neuroscience approach to psychopathy, or to any psychiatric disorder, is that the models should not only be models of the pathological state but, at least implicitly, they should be models of the healthy individual. Without a clear model of the functioning of the healthy individual it is difficult to consider how a patient's disorder might be pathological. The importance of this point with respect to computational modeling is that there is a new generation of computational cognitive neuroscience models of affect that are emerging. For example, Dayan et al.'s work on the temporal difference error has been used to understand and predict activity in striatal regions in both animals and humans (Montague, Dayan, & Sejnowski, 1996; O'Doherty et al., 2003). Siegle et al. have been modeling amygdala–cortical interactions in the context of depression and using these models to predict functional magnetic resonance imaging data in patients with this disorder (Siegle & Hasselmo, 2002; Siegle, Konecky, Thase, & Carter, 2002). The relevance of these and other models for understanding psychopathy has yet to be determined. However, their importance is that they provide highly specified predictions that can be tested through both behavioral and functional magnetic resonance imaging work.

The last, although connected, point concerns currently unexplained data at least from the perspective of the IES model. Two notable

points stand out. First, individuals with psychopathy show *reduced* interference on at least some Stroop type tasks (Hiatt et al., 2004; Newman et al., 1997). Second, recent work with the lexical decision task, where participants are presented with letter strings and must determine whether the string is a word or not, has indicated that although healthy individuals are faster to respond to emotional/high frequency words rather than neutral/low-frequency words, individuals with psychopathy are less so (Lorenz & Newman, 2002). Neither of these results can be explained by the IES model.

According to the RM hypothesis, the reduced interference on Stroop type measures and reduced frequency effects in lexical decision in individuals with psychopathy, is due to their inability to use the peripheral nontarget/frequency information because of their focus of attention on the dominant response set (responding to the target stimulus/deciding whether the stimulus was a word or not; Hiatt et al., 2004; Lorenz & Newman, 2002). Difficulties with the RM hypothesis were outlined above. However, in this context, it is worth considering that highly specified computational models of both Stroop performance and semantic processing exist. In Cohen et al.'s model of Stroop performance, the degree of interference is determined by the relative strength of the connections of the separate routes for the competing task demands (Cohen, Botvinick, & Carter, 2000; Cohen, Dunbar, & McClelland, 1990; Cohen & Servan–Schreiber, 1992). The strength of the connections is determined by the model's exposure to stimulus pairings mediated by the separate routes. In short, if the model has less experience with a form of stimulus pairing (i.e., less experience with reading), then reading will exert less interference in hue naming during the Stroop task. In recent computational models of semantic memory, frequency is considered as the model's experience with the stimulus pairing. Units representing more frequent items will be associated with greater weighted connections to other associated units than units representing less frequent items (Plaut & Booth, 2000; Rogers, Lambon Ralph, Gerrard, Bozeat, McClelland, Hodges, & Patterson, 2004).

Unlike the RM hypothesis, these models do not make reference to automatic switches of attention to nontarget or frequency information. However, these models do both imply that less experience with reading and verbal information more generally would give rise to the observed effects of reduced interference on Stroop tasks and reduced frequency effects in verbal fluency. Individuals with psychopathy, because of their limited educational exposure (Hare, 1991), might represent such a population and such an account has been offered of the above data sets (Peschardt, Newman, Mitchell, Richell, Leonard, Morton, & Blair, in press). It is important that this account makes clear predictions regarding individuals with subclinical psychopathy: individuals who present with the emotional dysfunction but who, because of other resources (intellectual/financial), do not present the full disorder. It suggests that these individuals should not present with these difficulties that may be a result of educational inexperience but should still present with impairment on the fundamental emotional paradigms. Future work will no doubt determine the validity of this prediction.

Conclusions

In this paper, four earlier models of psychopathy were reviewed from the perspective of cognitive neuroscience. These were the frontal lobe dysfunction, RM, fear dysfunction, and VIM hypotheses. Following this, an account inspired by recent findings in affective cognitive neuroscience as well as in relation to psychopathy, the IES model, was articulated.

The basis of the frontal lobe position is that frontal lobe dysfunction leads to aggression. The literature reviewed here suggests that this position needs greater specification. Specifically, it was suggested that the position should be qualified to orbital/ventrolateral frontal cortex dysfunction, if it disrupts the systems necessary for the rapid alteration of responding following contingency change, will be associated with an increased risk for a type of aggression, frustration-based reactive aggression (Blair, 2001, 2004). Importantly, this position was a cognitive neuroscience position; it spans

disorders. It does not only attempt to explain the behavior of neurological patients with acquired lesions of orbital/ventrolateral frontal cortex and individuals with psychopathy but also other disorders with comparable disruption; for example, childhood bipolar disorder and intermittent explosive disorder.

The RM hypothesis is an attention-based account. The basis of this position is that individuals with psychopathy are unable to use nontarget information because of their focus of attention on the dominant response set (responding to the target stimulus). It was suggested here that as an attention-based account it should either be compatible with contemporary cognitive neuroscience models of attention (e.g., Desimone & Duncan, 1995). It was suggested here that, at least as it is currently articulated, the RM hypothesis cannot be considered as compatible with dominant accounts of attention.

The fear dysfunction and VIM accounts were both emotion-based models. The basic position of both was that the behavior exhibited by individuals with psychopathy was due to a fundamental emotional deficit; either reduced fear or the malfunctioning of the VIM. It was argued here that the fear dysfunction accounts, as they are currently specified, have been effectively disproved. Fear is not mediated by a unitary system that functions to a greater or less degree in different individuals. Instead, fear-related behaviors are mediated by a collection of systems that may operate in an integrated fashion to achieve specific processing goals but are also considerably dissociable. Although the VIM account has not been disproved, its flaws are equally obvious; it can only explain a fraction of the pathology associated with psychopathy.

The proposed IES model can be considered an integration of the fear dysfunction and VIM positions. At the same time, a strong attempt has been made to be compatible with dominant cognitive neuroscience views on the nature of emotional processing. This position suggests a primary amygdala dysfunction in psychopathy. This deficit disrupts the ability of the individual with psychopathy to form stimulus–reinforcement associations (particularly stimulus–punishment associations though stimulus–reward associations are also affected, Peschardt et al., in press). This deficit interferes with socialization. The individual is less likely to learn to avoid the use of antisocial behavior to achieve their goals. Instead, the individual may learn to use antisocial behavior instrumentally to achieve their desires (they may receive the potential reward, e.g., financial gain, without the cost of the victim's distress).

This position also suggests orbital/ventrolateral frontal cortex dysfunction that disrupts the systems necessary for the rapid alteration of responding following contingency change, and is associated with an increased risk for frustration-based reactive aggression. Importantly, this second form of pathology, unlike the proposed amygdala dysfunction, is not unique to psychopathy but is also shared with other emotional disorders such as childhood bipolar disorder and intermittent explosive disorder.

One interesting question is why there are these two forms of pathology in psychopathy; amygdala and orbito/ventrolateral dysfunction. Several possibilities exist (Mitchell et al., 2002). First, the orbital/ventrolateral frontal cortex dysfunction might be due to the amygdala dysfunction; for example, because of reduced afferent input from the amygdala to orbital/ventrolateral frontal cortex. Second, the two types of pathology might be linked to a single pathology at a different level. For example, it is possible that both dysfunctions are due to disruption within a single neurotransmitter system (Peschardt, Leonard, et al., 2005). Third, the orbital/ventrolateral frontal cortex dysfunction might reflect the lifestyle of individuals with psychopathy. Part of the syndrome of psychopathy is an increased risk for drug use (Hare, 1991). Particular forms of drug use have been associated with orbital/ventrolateral frontal cortex dysfunction (Rogers & Robbins, 2001). Future work may determine which of these possibilities are correct.

In short, the cognitive neuroscience revolution has had an enormous impact on the general fields of psychology, neurology and psychiatry. It will be interesting to follow the impact of this revolution on the understanding of psychopathy.

References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*, 169–177.

Adolphs, R. (2003). Is the human amygdala specialized for processing social information? *Annals of the New York Academy of Sciences*, *985*, 326–340.

Adolphs, R., Baron–Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience*, *14*, 1264–1274.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, *393*, 470–474.

Amaral, D. G. (2001). *The amygdaloid complex and the neurobiology of social behaviour*. Paper presented at the Society for Research in Child Development, Minneapolis, MN.

Ambrogi Lorenzini, C. G., Baldi, E., Bucherelli, C., Sacchetti, B., & Tassoni, G. (1999). Neural topography and chronology of memory consolidation: A review of functional inactivation findings. *Neurobiology of Learning and Memory*, *71*, 1–18.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, I. M., Clark, L., Elliott, R., Kulkarni, B., Williams, S. R., & Deakin, J. F. W. (2002). 5-HT(2c) receptor activation by *m*-chlorophenylpiperazine detected in humans with fMRI. *Neuroreport*, *13*, 1547–1551.

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behaviour related to early damage in human prefrontal cortex. *Nature Neuroscience*, *2*, 1032–1037.

Angrilli, A., Mauri, A., Palomba, D., Flor, H., Birbaumer, N., Sartori, G., & di Paola, F. (1996). Startle reflex and emotion modulation impairment after a right amygdala lesion. *Brain*, *119*, 1991–2000.

Aniskiewicz, A. S. (1979). Autonomic components of vicarious conditioning and psychopathy. *Journal of Clinical Psychology*, *35*, 60–67.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*, 115–116.

Arsenio, W. F., & Fleiss, K. (1996). Typical and behaviourally disruptive children's understanding of the emotion consequences of socio-moral events. *British Journal of Developmental Psychology*, *14*, 173–186.

Asendorpf, J. B., & Nunner–Winkler, G. (1992). Children's moral motive strength and temperamental inhibition reduce their immoral behaviour in real moral conflicts. *Child Development*, *63*, 1223–1235.

Baddeley, A., & Della Sala, S. (1998). Working memory and executive control. In A. C. Roberts, T. W. Robbins, & L. Weiskrantz (Eds.), *The prefrontal cortex* (pp. 9–21). New York: Oxford University Press.

Bandura, A., & Rosenthal, T. L. (1966). Vicarious classical conditioning as a function of arousal level. *Journal of Personality and Social Psychology*, *3*, 54–62.

Barkley, R. A. (1999). Theories of attention-deficit/hyperactivity disorder. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 295–316). New York: Kluwer Academic/Plenum Press.

Baron–Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. (2000). The amygdala theory of autism. *Neuroscience Biobehavior Review*, *24*, 355–364.

Baron–Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience*, *11*, 1891–1898.

Barratt, E. S. (1994). Impulsiveness and aggression. In J. Monahan & H. Steadman (Eds.), *Violence and mental disorders: Developments in risk assessment* (pp. 61–79). Chicago: University of Chicago Press.

Barratt, E. S., Stanford, M. S., Dowdy, L., Liebman, M. J., & Kent, T. A. (1999). Impulsive and premeditated aggression: A factor analysis of self-reported acts. *Psychiatry Research*, *86*, 163–173.

Barratt, E. S., Stanford, M. S., Kent, T. A., & Felthous, A. (1997). Neuropsychological and cognitive psychophysiological substrates of impulsive aggression. *Biological Psychiatry*, *41*, 1045–1061.

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology Monographs*, *4*, 1–103.

Baumrind, D. (1983). Rejoinder to Lewis's interpretation of parental firm control effects: Are authoritative families really harmonious? *Psychological bulletin*, *94*, 132–142.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature Review Neuroscience*, *3*, 563–573.

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, *19*, 5473–5481.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science*, *269*, 1115–1118.

Berger, A., & Posner, M. I. (2000). Pathologies of brain attentional networks. *Neuroscience Biobehavioral Review*, *24*, 3–5.

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. Philadelphia, PA: Temple University Press.

Best, M., Williams, J. M., & Coccaro, E. F. (2002). Evidence for a dysfunctional prefrontal circuit in patients with an impulsive aggressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 8448–8453.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry*, *148*, 564–577.

Blackburn, R. (1988). On moral judgements and personality disorders: The myth of psychopathic personality revisited. *British Journal of Psychiatry*, *153*, 505–512.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, *57*, 1–29.

Blair, R. J. R. (1997). Moral reasoning in the child with psychopathic tendencies. *Personality and Individual Differences*, *22*, 731–739.

Blair, R. J. R. (1999). Responsiveness to distress cues in the child with psychopathic tendencies. *Personality and Individual Differences*, *27*, 135–145.

Blair, R. J. R. (2001). Neuro-cognitive models of aggres-

sion, the antisocial personality disorders and psychopathy. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*, 727–731.

Blair, R. J. R. (2003a). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 561–572.

Blair, R. J. R. (2003b). Neurobiological basis of psychopathy. *British Journal of Psychiatry*, *182*, 5–7.

Blair, R. J. R. (2003c). A neurocognitive model of the psychopathic individual. In M. A. Ron & T. W. Robbins (Eds.), *Disorders of brain and mind 2* (pp. 400–420). Cambridge: Cambridge University Press.

Blair, R. J. R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition*, *55*, 198–208.

Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal: A case of "acquired sociopathy." *Brain*, *123*, 1122–1141.

Blair, R. J., Colledge, E., & Mitchell, D. G. (2001). Somatic markers and response reversal: Is there orbitofrontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology*, *29*, 499–511.

Blair, R. J. R., Colledge, E., Murray, L., & Mitchell, D. G. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, *29*, 491–498.

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1995). Is the psychopath "morally insane"? *Personality and Individual Differences*, *19*, 741–752.

Blair, R. J. R., Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology*, *34*, 192–198.

Blair, R. J. R., Mitchell, D. G. V., & Blair, K. S. (2005). *The psychopath: Emotion and the brain*. Oxford: Blackwell.

Blair, R. J., Mitchell, D. G., Richell, R. A., Kelly, S., Leonard, A., Newman, C., & Scott, S. K. (2002). Turning a deaf ear to fear: Impaired recognition of vocal affect in psychopathic individuals. *Journal of Abnormal Psychology*, *111*, 682–686.

Blair, R. J. R., Monson, J., & Frederickson, N. (2001). Moral reasoning and conduct problems in children with emotional and behavioural difficulties. *Personality and Individual Differences*, *31*, 799–811.

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain*, *122*, 883–893.

Blair, R. J. R., Newman, C., Mitchell, D. G. V., Peschardt, K. S., & Leonard, A. (2005). *Passive avoidance learning in psychopathic individuals: Modulation by reward but not by punishment*. Manuscript submitted for publication.

Blonigen, D. M., Carlson, R. F., Krueger, R. F., & Patrick, C. J. (2003). A twin study of self-reported psychopathic personality traits. *Personality and Individual Differences*, *35*, 179–197.

Bremner, J. D., & Vermetten, E. (2001). Stress and development: Behavioral and biological consequences. *Development and Psychopathology*, *13*, 473–489.

Brody, G. H., & Shaffer, D. R. (1982). Contributions of parents and peers to children's moral socialisation. *Developmental Review*, *2*, 31–75.

Budhani, S., & Blair, R. J. R. (in press). Probabilistic response reversal in children with psychopathic tendencies. *Journal of Child Psychology and Psychiatry*.

Budhani, S., & Blair, R. J. R. (2005). *Probabilistic response reversal in children with psychopathic tendencies*. Manuscript submitted for publication.

Budhani, S., Richell, R. A., & Blair, R. J. R. (2005). *Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy*. Manuscript submitted for publication.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*, 263–272.

Burgess, P. W., & Wood, R. L. (1990). Neuropsychology of behaviour disorders following brain injury. In R. L. Wood (Ed.), *Neurobehavioural sequelae of traumatic brain injury* (pp. 110–133). London: Taylor & Francis.

Camras, L. A. (1977). Facial expressions used by children in a conflict situation. *Child Development*, *48*, 1431–1435.

Casey, B. J., Forman, S. D., Franzen, P., Berkowitz, A., Braver, T. S., Nystrom, L. E., Thomas, K. M., & Noll, D. C. (2001). Sensitivity of prefrontal cortex to changes in target probability: A functional MRI study. *Human Brain Mapping*, *13*, 26–33.

Chaplin, T. C., Rice, M. E., & Harris, G. T. (1995). Salient victim suffering and the sexual responses of child molesters. *Journal of Consulting and Clinical Psychology*, *63*, 249–255.

Charney, D. S. (2003). Neuroanatomical circuits modulating fear and anxiety behaviors. *Acta Psychiatrica Scandinavia*, *417*(Suppl.), 38–50.

Church, R. M. (1959). Emotional reactions of rats to the pain of others. *Journal of Comparative & Physiological Psychology*, *52*, 132–134.

Cicchetti, D. (2003). Editorial: Experiments of nature: Contributions to developmental theory. *Development and Psychopathology*, *15*, 833–835.

Cicchetti, D., & Dawson, G. (2002). Editorial: Multiple levels of analysis. *Development and Psychopathology*, *14*, 417–420.

Cleckley, H. M. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby.

Coccaro, E. F. (1998). Impulsive aggression: A behavior in search of clinical definition. *Harvard Review of Psychiatry*, *5*, 336–339.

Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, *3*, 421–423.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychological Review*, *97*, 332–361.

Cohen, J. D., & Servan–Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychological Review*, *99*, 45–77.

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *22*, 4563–4567.

Cornell, D. G., Warren, J., Hawk, G., Stafford, E., Oram, G., & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. *Journal of Consulting and Clinical Psychology*, *64*, 783–790.

Crean, J., Richards, J. B., & de Wit, H. (2002). Effect of tryptophan depletion on impulsive behavior in men

with or without a family history of alcoholism. *Behavioral Brain Research*, *136*, 349–357.

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development*, *67*, 993–1002.

Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York: Putnam Press (Grosset Books).

Davidson, R. J. (2003). Affective neuroscience and psychophysiology: Toward a synthesis. *Psychophysiology*, *40*, 655–665.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, *380*, 69–72.

Duncan, J. (1998). Converging levels of analysis in the cognitive neuroscience of visual attention. *Philosophical Transactions of the Royal Society B*, *353*, 1307–1317.

Elliot, F. A. (1978). Neurological aspects of antisocial behavior. In W. H. Reid (Ed.), *The psychopath*. New York: Bruner/Mazel.

Everitt, B. J., Cardinal, R. N., Parkinson, J. A., & Robbins, T. W. (2003). Appetitive behavior: Impact of amygdala-dependent mechanisms of emotional learning. *Annual New York Academy of Sciences*, *985*, 233–250.

Eysenck, H. J. (1964). *Crime and personality*. London: Routledge & Kegan Paul.

Eysenck, H. J., & Gudjonsson, G. H. (1989). *The causes and cures of criminality*. London: Plenum Press.

Fisher, L., & Blair, R. J. R. (1998). Cognitive impairment and its relationship to psychopathic tendencies in children with emotional and behavioural difficulties. *Journal of Abnormal Child Psychology*, *26*, 511–519.

Flor, H., Birbaumer, N., Hermann, C., Ziegler, S., & Patrick, C. J. (2002). Aversive Pavlovian conditioning in psychopaths: Peripheral and central correlates. *Psychophysiology*, *39*, 505–518.

Forth, A. E., Kosson, D. S., & Hare, R. D. (in press). *The Psychopathy Checklist: Youth Version*. Toronto: Multi-Health Systems.

Fowles, D. C. (1988). Psychophysiology and psychopathy: A motivational approach. *Psychophysiology*, *25*, 373–391.

Francis, D. D., Caldji, C., Champagne, F., Plotsky, P. M., & Meaney, M. J. (1999). The role of corticotropin-releasing factor–norepinephrine systems in mediating the effects of early experience on the development of behavioral and endocrine responses to stress. *Biological Psychiatry*, *46*, 1153–1166.

Frick, P. J., & Hare, R. D. (2001). *The antisocial process screening device*. Toronto: Multi-Health Systems.

Frick, P. J., O'Brien, B. S., Wootton, J. M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology*, *103*, 700–707.

Fuster, J. M. (1980). *The prefrontal cortex*. New York: Raven Press.

Giedd, J. N., Blumenthal, J., Molloy, E., & Castellanos, F. X. (2001). Brain imaging of attention deficit/hyperactivity disorder. *Annals of the New York Academy of Science*, *931*, 33–49.

Gorenstein, E. E. (1982). Frontal lobe functions in psychopaths. *Journal of Abnormal Psychology*, *91*, 368–379.

Gorrindo, T., Blair, R. J. R., Budhani, S., Pine, D. S., & Leibenluft, E. (2005). *Probabilistic response reversal deficits in pediatric bipolar disorder*. Manuscript submitted for publication.

Goyer, P. F., Andreason, P. J., Semple, W. E., Clayton, A. H., King, A. C., Compton–Toth, B. A., Schulz, S. C., & Cohen, R. M. (1994). Positron-emission tomography and personality disorders. *Neuropsychopharmacology*, *10*, 21–28.

Grafman, J., Schwab, K., Warden, D., Pridgen, B. S., & Brown, H. R. (1996). Frontal lobe injuries, violence, and aggression: A report of the vietnam head injury study. *Neurology*, *46*, 1231–1238.

Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge: Cambridge University Press.

Gregg, T. R., & Siegel, A. (2001). Brain structures and neurotransmitters regulating aggression in cats: Implications for human aggression. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *25*, 91–140.

Hare, R. D. (1982). Psychopathy and physiological activity during anticipation of an aversive stimulus in a distraction paradigm. *Psychophysiology*, *19*, 266–271.

Hare, R. D. (1991). *The Hare Psychopathy Checklist—Revised*. Toronto: Multi-Health Systems.

Hare, R. D. (2003). *The Hare Psychopathy Checklist—Revised* (2nd ed.). Toronto: Multi-Health Systems.

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. *Journal of Abnormal Psychology*, *103*, 604–609.

Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *1*, 6–17.

Hart, S. D., Forth, A. E., & Hare, R. D. (1990). Performance of criminal psychopaths on selected neuropsychological tests. *Journal of Abnormal Psychology*, *99*, 374–379.

Hiatt, K. D., Schmitt, W. A., & Newman, J. P. (2004). Stroop tasks reveal abnormal selective attention among psychopathic offenders. *Neuropsychology*, *18*, 50–59.

Hinshaw, S. P. (1987). On the distinction between attentional deficits/hyperactivity and conduct problems/aggression in child psychopathology. *Psychological Bulletin*, *101*, 443–463.

Hoffman, M. L. (1984). Empathy, its limitations, and its role in a comprehensive moral theory. In J. Gewirtz & W. Kurtines (Eds.), *Morality, moral development, and moral behavior* (pp. 283–302). New York: Wiley.

Hoffman, M. L. (1988). Moral development. In M. Bornstein & M. Lamb (Eds.), *Developmental psychology: An advanced textbook* (pp. 497–548). Hillsdale, NJ: Erlbaum.

Hoffman, M. L. (1994). Discipline and internalisation. *Developmental Psychology*, *30*, 26–28.

Hoffman, M. L., & Saltzstein, H. D. (1967). Parent discipline and the child's moral development. *Journal of Personality and Social Psychology*, *5*, 45–57.

House, T. H., & Milligan, W. L. (1976). Autonomic responses to modeled distress in prison psychopaths. *Journal of Personality and Social Psychology*, *34*, 556–560.

Kandel, E., & Freed, D. (1989). Frontal lobe dysfunction and antisocial behavior: A review. *Journal of Clinical Psychology*, *45*, 404–413.

Kiehl, K A. (in press). A paralimbic dysfunction hypothesis of psychopathy. In D. Barch (Ed.), *Cognitive and affective neuroscience of psychopathology*. London: Oxford University Press.

Kiehl, K. A., Hare, R. D., McDonald, J. J., & Brink, J. (1999). Semantic and affective processing in psycho-

paths: An event-related potential (ERP) study. *Psychophysiology*, *36*, 765–774.

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, *50*, 677–684.

Kiehl, K. A., Smith, A. M., Mendrek, A., Forster, B. B., Hare, R. D., & Liddle, P. F. (2004). Temporal lobe abnormalities in semantic processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Psychiatry Research*, *130*, 27–42.

Killcross, S., Robbins, T. W., & Everitt, B. J. (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature*, *388*, 377–380.

Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, *64*, 325–347.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments: From toddlerhood to age 5. *Developmental Psychology*, *33*, 228–240.

Kochanska, G., De Vet, K., Goldman, M., Murray, K., & Putman, P. (1994). Maternal reports of conscience development and temperament in young children. *Child Development*, *65*, 852–868.

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates and consequences. *Journal of Personality*, *71*, 1087–1112.

Kopnisky, K. L., Cowan, W. M., & Hyman, S. E. (2002). Levels of analysis in psychiatric research. *Development and Psychopathology*, *14*, 437–461.

Kosson, D. S., Cyterski, T. D., Steuerwald, B. L., Neumann, C. S., & Walker–Matthews, S. (2002). The reliability and validity of the psychopathy checklist: Youth version (PCL:YV) in nonincarcerated adolescent males. *Psychological Assessment*, *14*, 97–109.

Krakowski, M., Czobor, P., Carpenter, M. D., Libiger, J., Kunz, M., Papezova, H., Parker, B. B., Schmader, L., & Abad, T. (1997). Community violence and inpatient assaults: Neurobiological deficits. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *9*, 549–555.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *Journal of Neuroscience*, *15*, 6846–6855.

LaPierre, D., Braun, C. M. J., & Hodgins, S. (1995). Ventral frontal deficits in psychopathy: Neuropsychological test findings. *Neuropsychologia*, *33*, 139–151.

LeDoux, J. (1998). *The emotional brain*. New York: Weidenfeld & Nicolson.

Leibenluft, E., Blair, R. J., Charney, D. S., & Pine, D. S. (2003). Irritability in pediatric mania and other childhood psychopathology. *Annual New York Academy of Sciences*, *1008*, 201–218.

Levenston, G. K., Patrick, C. J., Bradley, M. M., & Lang, P. J. (1996). *Psychopathy and startle modulation during affective picture processing: A replication and extension*. Paper presented at the Annual Meeting of the Society for Psychophysiological Research, Vancouver, Canada.

Levenston, G. K., Patrick, C. J., Bradley, M. M., & Lang, P. J. (2000). The psychopath as observer: Emotion and attention in picture processing. *Journal of Abnormal Psychology*, *109*, 373–386.

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, *66*, 488–524.

Linnoila, M., Virkkunen, M., Scheinin, M., Nuutila, A., Rimon, R., & Goodwin, F. K. (1983). Low cerebrospinal fluid 5-hydroxy indoleacetic acid concentration differentiates impulsive from nonimpulsive violent behavior. *Life Sciences*, *33*, 2609–2614.

Lorenz, A. R., & Newman, J. P. (2002). Deficient response modulation and emotion processing in low-anxious Caucasian psychopathic offenders: Results from a lexical decision task. *Emotion*, *2*, 91–104.

Luria, A. (1966). *Higher cortical functions in man*. New York: Basic Books.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology*, *55*, 6–10.

Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale, NJ: Erlbaum.

Masserman, J. H., Wechkin, S., & Terris, W. (1964). "Altruistic" behavior in rhesus monkeys. *American Journal of Psychiatry*, *121*.

McNaughton, N., & Gray, J. A. (2000). Anxiolytic action on the behavioural inhibition system implies multiple types of arousal contribute to anxiety. *Journal of Affective Disorders*, *61*, 161–176.

Mitchell, D. G. V., Colledge, E., Leonard, A., & Blair, R. J. R. (2002). Risky decisions and response reversal: Is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? *Neuropsychologia*, *40*, 2013–2022.

Moffitt, T. E. (1993a). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Moffitt, T. E. (1993b). The neuropsychology of conduct disorder. *Development and Psychopathology*, *5*, 135–152.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, *16*, 1936–1947.

Morgan, A. B., & Lilienfield, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*, 113–136.

Newman, J. P. (1998). Psychopathic behaviour: An information processing perspective. In D. J. Cooke, A. E. Forth, & R. D. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 81–105). Dordrecht: Kluwer Academic.

Newman, J. P., & Kosson, D. S. (1986). Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*, *95*, 252–256.

Newman, J. P., Patterson, C. M., & Kosson, D. S. (1987). Response perseveration in psychopaths. *Journal of Abnormal Psychology*, *96*, 145–148.

Newman, J. P., Schmitt, W. A., & Voss, W. D. (1997). The impact of motivationally neutral cues on psychopathic individuals: Assessing the generality of the response modulation hypothesis. *Journal of Abnormal Psychology*, *106*, 563–575.

Nisenbaum, L. K., Zigmond, M. J., Sved, A. F., & Abercrombie, E. D. (1991). Prior exposure to chronic stress results in enhanced synthesis and release of hippocampal norepinephrine in response to a novel stressor. *Journal of Neuroscience*, *11*, 1478–1484.

Nucci, L. P., & Herman, S. (1982). Behavioral disordered children's conceptions of moral, conventional, and personal issues. *Journal of Abnormal Child Psychology*, *10*, 411–425.

O'Brien, B. S., & Frick, P. J. (1996). Reward dominance: Associations with anxiety, conduct problems, and psychopathy in children. *Journal of Abnormal Child Psychology*, *24*, 223–240.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, *304*, 452–454.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*, 329–337.

O'Keefe, J. (1991). The hippocampal cognitive map and navigational strategies. In J. Paillard (Ed.), *Brain and space* (pp. 273–295). Oxford: Oxford University Press.

Ogloff, J. R., & Wong, S. (1990). Electrodermal and cardiovascular evidence of a coping response in psychopaths. *Criminal Justice and Behaviour*, *17*, 231–245.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Pastor, M. C., Molto, J., Vila, J., & Lang, P. J. (2003). Startle reflex modulation, affective ratings and autonomic reactivity in incarcerated Spanish psychopaths. *Psychophysiology*, *40*, 934–938.

Patrick, C. J. (1994). Emotion and psychopathy: Startling new insights. *Psychophysiology*, *31*, 319–330.

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the criminal psychopath: Startle reflex modulation. *Journal of Abnormal Psychology*, *102*, 82–92.

Patterson, C. M., & Newman, J. P. (1993). Reflectivity and learning from aversive events: Toward a psychological mechanism for the syndromes of disinhibition. *Psychological Review*, *100*, 716–734.

Pennington, B. F., & Bennetto, L. (1993). Main effects or transaction in the neuropsychology of conduct disorder? Commentary on "the neuropsychology of conduct disorder." *Development and Psychopathology*, *5*, 153–164.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*, 51–87.

Perry, D. G., & Perry, L. C. (1974). Denial of suffering in the victim as a stimulus to violence in aggressive boys. *Child Development*, *45*, 55–62.

Peschardt, K. S., Leonard, A., Morton, J., & Blair, R. J. R. (2005). *Differential stimulus–reward and stimulus–punishment learning in individuals with psychopathy*. Manuscript submitted for publication.

Peschardt, K. S., Morton, J., & Blair, R. J. R. (2005). *They know the words but don't feel the music: Reduced affective priming in psychopathic individuals*. Manuscript submitted for publication.

Peschardt, K. S., Newman, C., Mitchell, D. G. V., Richell, R. A., Leonard, A., Morton, J., & Blair, R. J. R. (in press). Differentiating among prefrontal substrates in psychopathy: Neuropsychological test findings. *Neuropsychology*.

Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., & Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience*, *23*, 11078–11094.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*, 271–275.

Raine, A. (1993). *The psychopathology of crime: Criminal behavior as a clinical disorder*. San Diego, CA: Academic Press.

Raine, A. (1997). *The psychopathology of crime*. New York: Academic Press.

Raine, A. (2002). Annotation: The role of prefrontal deficits, low autonomic arousal, and early health factors in the development of antisocial and aggressive behavior in children. *Journal of Child Psychology and Psychiatry*, *43*, 417–434.

Raine, A., Buchsbaum, M. S., & LaCasse, L. (1997). Brain abnormalities in murderers indicated by positron emission tomography. *Biological Psychiatry*, *42*, 495–508.

Raine, A., Buchsbaum, M. S., Stanley, J., Lottenberg, S., Abel, L., & Stoddard, J. (1994). Selective reductions in prefrontal glucose metabolism in murderers. *Biological Psychiatry*, *15*, 365–373.

Raine, A., Meloy, J. R., Birhle, S., Stoddard, J., LaCasse, L., & Buchsbaum, M. S. (1998). Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers. *Behaviour Science and Law*, *16*, 319–332.

Rice, G. E. (1965). *Aiding responses in rats: Not in guinea pigs*. Paper presented at the Annual Convention of the American Psychological Association.

Rice, G. E., & Gainer, P. (1962). "altruism" in the albino rat. *Journal of Comparative & Physiological Psychology*, *55*, 123–125.

Richell, R. A., Mitchell, D. G., Newman, C., Leonard, A., Baron–Cohen, S., & Blair, R. J. (2003). Theory of mind and psychopathy: Can psychopathic individuals read the "language of the eyes"? *Neuropsychologia*, *41*, 523–526.

Richell, R. A., Mitchell, D. G. V., Peschardt, K. S., Winston, J. S., Leonard, A., Dolan, R. J., & Blair, R. J. R. (2005). *Trust and distrust: The perception of trustworthiness in faces in psychopathic and non-psychopathic offenders*. Manuscript submitted for publication.

Roberts, A. C., Robbins, T. W., & Weiskrantz, L. (1998). *The prefrontal cortex: Executive and cognitive functions*. Oxford: Oxford University Press.

Rogers, R. D., Blackshaw, A. J., Middleton, H. C., Matthews, K., Hawtin, K., Crowley, C., Hopwood, A., Wallace, C., Deakin, J. F., Sahakian, B. J., & Robbins, T. W. (1999). Tryptophan depletion impairs stimulus–reward learning while methylphenidate disrupts attentional control in healthy young adults: Implications for the monoaminergic basis of impulsive behaviour. *Psychopharmacology (Berlin)*, *146*, 482–491.

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., & Robbins, T. W. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *Journal of Neuroscience*, *19*, 9029–9038.

Rogers, R. D., & Robbins, T. W. (2001). Investigating the neurocognitive deficits associated with chronic

drug misuse. *Current Opinion in Neurobiology*, *11*, 250–257.

Rogers, T. T., Lambon Ralph, M. A., Gerrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*, 205–235.

Rolls, E. T. (1997). The orbitofrontal cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *351*, 1433–1443.

Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*, 1518–1524.

Rothbart, M., Ahadi, S., & Hershey, K. L. (1994). Temperament an social behaviour in children. *Merrill–Palmer Quarterly*, *40*, 21–39.

Rothbart, M. K., & Ahadi, S. (1994). Temperament and the development of personality. *Journal of Abnormal Psychology*, *103*, 55–66.

Rowley, B., Van, F., Mortimore, C., & Connell, J. (1997). Effects of acute tryptophan depletion on tests of frontal and temporal lobe function. *Journal of Psychopharmacology*, *11*, A60.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Schmitt, J. A., Jorissen, B. L., Sobczak, S., van Boxtel, M. P., Hogervorst, E., Deutz, N. E., & Riedel, W. J. (2000). Tryptophan depletion impairs memory consolidation but improves focussed attention in healthy young volunteers. *Journal of Psychopharmacology*, *14*, 21–29.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.

Siegle, G. J., & Hasselmo, M. E. (2002). Using connectionist models to guide assessment of psychological disorder. *Psychological Assessment*, *14*, 263–278.

Siegle, G. J., Konecky, R. O., Thase, M. E., & Carter, C. S. (2002). Can't shake that feeling: Event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biological Psychiatry*, *51*, 693–707.

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, *52*, 1333–1336.

Smetana, J. G. (1985). Preschool children's conceptions of transgressions: Effects of varying moral and conventional domain-related attributes. *Developmental Psychology*, *21*, 18–29.

Smetana, J. G. (1993). Understanding of social rules. In M. Bennett (Ed.), *The child as psychologist: An introduction to the development of social cognition* (pp. 111–141). New York: Harvester Wheatsheaf.

Smith, M. A., Kim, S. Y., van Oers, H. J., & Levine, S. (1997). Maternal deprivation and stress induce immediate early genes in the infant rat brain. *Endocrinology*, *138*, 4622–4628.

Smith, S. S., Arnett, P. A., & Newman, J. P. (1992). Neuropsychological differentiation of psychopathic and non-psychopathic criminal offenders. *Personality and Individual Differences*, *13*, 1233–1243.

Stevens, D., Charman, T., & Blair, R. J. R. (2001). Recognition of emotion in facial expressions and vocal tones in children with psychopathic tendencies. *Journal of Genetic Psychology*, *162*, 201–211.

Stone, V. E., Baron–Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, *41*, 209–220.

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia*, *39*, 771–786.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*, 135–170.

Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: Neuropsychological evidence for topographic specificity. *Neuropsychologia*, *40*, 1240–1253.

Taylor, E. A., Schachar, R., Thorley, G., & Wieselberg, M. (1986). Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *British Journal of Psychiatry*, *149*, 760–767.

Trasler, G. (1978). Relations between psychopathy and persistent criminality—Methodological and theoretical issues. In R. D. Hare & D. Schalling (Eds.), *Psychopathic behaviour: Approaches to research*. Chichester: Wiley.

Veit, R., Flor, H., Erb, M., Hermann, C., Lotze, M., Grodd, W., & Birbaumer, N. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters*, *328*, 233–236.

Verona, E., Patrick, C. J., Curtin, J. J., Bradley, M. M., & Lang, P. J. (2004). Psychopathy and physiological response to emotionally evocative sounds. *Journal of Abnormal Psychology*, *113*, 99–108.

Viding, E., Blair, R. J. R., Moffitt, T. E., & Plomin, R. (in press). Psychopathic syndrome indexes strong genetic risk for antisocial behaviour in 7-year-olds. *Journal of Child Psychology and Psychiatry*.

Volkow, N. D., & Tancredi, L. (1987). Neural substrates of violent behaviour. A preliminary study with positron emission tomography. *British Journal of Psychiatry*, *151*, 668–673.

Volkow, N. D., Tancredi, L. R., Grant, C., Gillespie, H., Valentine, A., Mullani, N., Wang, G. J., & Hollister, L. (1995). Brain glucose metabolism in violent psychiatric patients: A preliminary study. *Psychiatry Research*, *61*, 243–253.

Williamson, S., Hare, R. D., & Wong, S. (1987). Violence: Criminal psychopaths and their victims. *Canadian Journal of Behavioral Science*, *19*, 454–462.

Williamson, S., Harpur, T. J., & Hare, R. D. (1991). Abnormal processing of affective words by psychopaths. *Psychophysiology*, *28*, 260–273.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, *5*, 277–283.