Paranoid Thinking, Suspicion, and Risk for Aggression: A Neurodevelopmental Perspective

Erin B. Tone  
Georgia State University, etone@gsu.edu

Jennifer S. Davis

Follow this and additional works at: http://scholarworks.gsu.edu/psych_facpub

Part of the Psychology Commons

Recommended Citation
Beliefs that “others are likely to mistreat me” occur from childhood throughout the life span; some authors argue that such beliefs are becoming increasingly common (Freeman & Freeman, 2009). It is thus not surprising that the term paranoia has migrated from the clinical literature into the popular lexicon (Freeman & Freeman, 2008). Although paranoia originally signified psychopathology or madness (Lewis, 1970), the word now more broadly denotes “a tendency on the part of an individual or group toward excessive or irrational suspiciousness and distrustfulness of others” (Merriam-Webster, 2012). For many individuals, beliefs that appear paranoid stem from real life experiences and may constitute rational responses to threatening environments. Discrimination and microaggression (Rippy & Newman, 2006; Sue, Capodilupo, & Holder, 2008), economic inequity (Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997), peer rejection and bullying (Campbell & Morrison, 2007; Schreier et al., 2009), and childhood abuse and neglect (De Loore et al., 2007; Lataster et al., 2006; Natsuaki, Cicchetti, & Rogosch, 2009; Price & Glad, 2003) all predict biases or proclivities to anticipate physical or psychological aggression from others. As each of the studies cited above also illustrates, however, the correspondence between negative life experiences and paranoid thinking is far from perfect. Many individuals who experience mistreatment never develop paranoid ideas or conceptualize ostensibly benign others as sources of danger; conversely, not all individuals who hold such beliefs show clear histories of having been aggression targets.

Further, although some evidence links paranoid/suspicious attitudes or related types of beliefs and biases to concurrent or later aggressive behavior (e.g., Bjørkly, 2002; De Castro et al., 2002; Dodge, 1980; Haggård-Grann, Hallqvist, Långström, & Möller, 2006), associations vary according to the ways in which suspicious or paranoid perceptions and attributions are operationally defined (e.g., cognitive biases vs. impairing clinical symptoms), the types of aggression under study (e.g., reactive, physical, relational, instrumental), the affective contexts in which suspicious beliefs manifest (e.g., anxious, angry, depressed), and individual demographic characteristics such as age (Bailey & Ostrov, 2008; Bjørkly, 2002; De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002). Thus, the likely complex ways in which aggression and paranoid patterns of thought might fuel or inhibit each other’s development remain unclear.

The results from a number of studies provide evidence that it may be useful to examine the associations with aggressive behavior separately for clinically significant delusional paranoia and normal-range suspiciousness. The relationship between delusions that one is at risk of harm from others or delusions that override typical self-control (threat/control override delusions) and violence has received substantive attention in the literature; the findings have yielded mixed results. Data regarding the associations between a broadly suspicious stance and violence are sparser; however, the findings from at least one study suggest that this relationship merits closer attention (Appelbaum, Robbins, & Monahan, 2000).

A number of studies since 1990 have yielded evidence of a strong relationship between threat/control override delusions...
and violence in adult psychiatric patients (Hodgins, Hiscoe, & Freese, 2003; Link, Monahan, Stueve, & Cullen, 1999; Link & Stueve, 1994; Swanson, Borum, Swartz, & Monahan, 1996; Swanson et al., 1997). In contrast to these retrospectively designed studies, however, the MacArthur Study of Violence Risk (Monahan, 2002; Monahan et al., 2001), a large-scale, prospective longitudinal study of violence in the community among adults discharged from psychiatric hospitals, found a nonsignificant association between delusional beliefs that others intended harm or controlled one’s thoughts or actions and violent behavior in the subsequent year (Appelbaum et al., 2000). Some more recent studies suggest that accounting for moderating variables, such as severity of violent offenses (Stompe, Ortwein-Swoboda, & Schanda, 2004) or gender (Teasdale, Silver, & Monahan, 2006), may help explain discrepancies among research results.

Although it has yielded limited evidence of associations between delusions and violence in psychiatric patients, the MacArthur Study of Violence Risk has produced support for the idea that a generally suspicious attitude toward others, marked by anger and impulsivity, significantly predicts later violent behavior, including physical and sexual assault (Appelbaum et al., 2000). Other lines of research in nonclinical samples also suggest links between aggression and both focal individual and diffuse societal suspicion. Suspicions of infidelity, for instance, have been tied to risk for intimate partner violence among men (Kaighobadi & Shackelford, 2009; Kaighobadi, Starratt, Shackelford, & Popp, 2008), and a suspicious cognitive schema has been linked similarly to male sexual aggression (Malamuth & Brown, 1994). In addition, qualitative research findings indicate that among male youths in conflict-ridden regions such as Northern Ireland, suspicious attitudes partly fuel a subculture of violence (Harland, 2011). More research into the associations between normal-range suspicion and aggression is clearly needed; a focus on subclinical paranoia may be of particular value in studies of children and young adolescents, who rarely exhibit clinically significant paranoid thoughts but are likely to show broad variability in the emergence of normative suspicions.

A Developmental Psychopathology Perspective on Paranoia and Aggression

Several tenets drawn from the developmental psychopathology perspective provide a coherent and useful framework for examining the relationship between paranoid or suspicious types of thinking and aggression, as well as putative mechanisms underlying this relationship as they emerge and change across the life span (Cicchetti, 2006; Cicchetti & Cohen, 1995). This perspective asserts that multiple factors interact in a dynamic, transactional manner to influence the emergence of both adaptive and maladaptive patterns of thought, affect, and behavior. The effects of a given etiological factor can vary dramatically, depending on the contexts in which the factor manifests (principle of multifinality); conversely, a given outcome can emerge from the influences of a wide range of factors (principle of equifinality). Thus, the appearance of paranoia or suspicion at any particular point in development depends on the ever-changing interplay among vulnerabilities and risk or protective factors for a given individual.

The developmental psychopathology perspective also emphasizes the value inherent in concurrently examining both normal and atypical patterns of development (Cicchetti, 2006; Cicchetti & Cohen, 1995). This emphasis is particularly important in light of ongoing debates about the relative merits of categorical and dimensional models of function and dysfunction (David, 2010; Kessler, 2002); unless we understand normal or adaptive developmental paths, what constitutes significant deviation from these paths, and both when and how individuals might come to veer from them, we are unlikely to generate accurate models of dysfunction. The ability to detect or anticipate threats in the social environment, for instance, is a critically important interpersonal skill from childhood onward; only under some circumstances and in certain individuals does this capacity become problematic or dysfunctional (Green & Phillips, 2004). A clear understanding of normative experiences of paranoia or suspiciousness, contextualized appropriately according to an individual’s current developmental stage, is thus a necessary component of research on their pathological manifestations.

Several rich literatures shed light both on ways in which paranoid beliefs or interpersonal suspiciousness might develop and on factors, particularly biologically based characteristics, that might strengthen or weaken their links to aggressive behavior. Ample clinical research has examined delusional paranoid or persecutory beliefs in the context of psychotic disorders such as schizophrenia (Bentall, 1994; Freeman, 2010); more recently, this line of work has expanded to consider mechanisms and correlates of “normal-range” paranoid beliefs, primarily among adults, in the broader community (Freeman et al., 2005; Green et al., 2011). Research in these veins has yielded important insight into how people process threatening information and how individuals might come to veer from them, we are unlikely to generate accurate models of dysfunction. The ability to detect or anticipate threats in the social environment, for instance, is a critically important interpersonal skill from childhood onward; only under some circumstances and in certain individuals does this capacity become problematic or dysfunctional (Green & Phillips, 2004). A clear understanding of normative experiences of paranoia or suspiciousness, contextualized appropriately according to an individual’s current developmental stage, is thus a necessary component of research on their pathological manifestations.

The literature that explicitly examines paranoia in both clinical and normal ranges, however, is notably limited by its heavy focus on adults. Attention to a set of distinct but related bodies of research that more tightly integrate developmental, cognitive, and neuroscience studies may better help illuminate how this style of thought begins and evolves in early life at psychological and neural levels. These literatures trace the emergence, maintenance, and potential mechanisms of several cognitive biases and anomalies that careful empirical work has identified as probable components of paranoid thought. Given that the construct of paranoia encompasses both cognitive and affective components (Freeman, 2007), it may also be helpful to examine how these biases and atyp-
Paranoid thinking and aggression risk

Such an examination of constructs that map onto portions of the conceptual space that paranoia occupies has advantages. In particular, it facilitates the integration of the distinct insights that multiple literatures offer into putative foundations and developmental courses of excessive or irrational suspiciousness and distrustfulness of others. By reviewing them in an integrated fashion, we aim to advance the understanding of the ways in which individual biological characteristics, such as patterns of structure and activity in the brain, interact dynamically over time with environmental contexts to yield paranoid types of cognitive and emotional experience and, in some cases, elevated risk for aggression.

We first define paranoia or paranoid thought and examine its prevalence across the life span in both clinical and community contexts. We then summarize theoretical models that postulate core components of paranoia that span both cognitive and affective domains. Finally, we review the literature on neural correlates of both paranoia and several critical component processes and examine whether and how convergent findings overlap with data regarding neural risk factors for aggressive behavior.

Paranoia

Definitions and prevalence

The term paranoia was initially used in ancient Greece to describe those who were “mad” or “out of their minds”; in subsequent centuries, it evolved in meaning across both time and language (Lewis, 1970). By the early 20th century, however, the American and European medical and psychological literatures were notably consistent in defining paranoia as a clinical state marked by delusions or false beliefs that are firmly maintained, even in the face of indisputable evidence that they are false (Merriam-Webster, 2012). Although delusions vary widely in content, in the context of paranoia they tend to manifest as delusions of reference, with a tendency toward either “ideas of injury” (e.g., persecutory delusions) or “ideas of exaltation” (Kraepelin, 1921, p. 220). Such traditional definitions classify paranoid thinking as a psychotic process marked by disconnection from reality, such that “numerous impressions and occurrences are not accepted in their sober, everyday character, but they enter into some or other relation to the patient’s own fortunes and misfortunes” (Kraepelin, 1921, p. 217).

However, a parallel definition that encompasses paranoid attitudes, beliefs, and behaviors that are not associated with psychosis also emerged in the early 20th century (e.g., Kant, 1927). Paranoia has since become widely used in ordinary, nonclinical discourse and the popular media to denote exaggerated or, in some cases, baseless fears or beliefs that others intend you harm or distress (Keltner & Davidson, 2009). Theorists have argued for decades that processes such as paranoid thought, which have historically been considered psychotic, might be better conceptualized as “points on continua of function” (Strauss, 1969, p. 581) or as quasi-continuous variables with both dichotomous and continuous characteristics (van Os, Linscott, Myin-Germeys, Delespaul, & Krabbendam, 2009; van Os & Verdoux, 2003).

Such continua could be defined in terms of an individual’s certainty that atypical experiences are objectively real, the ways in which cultural and environmental factors support such beliefs, the degree of preoccupation with atypical experiences, and the extent to which others find the experiences plausible (Strauss, 1969). Individuals who do not meet the criteria for active psychosis and who function adequately in their typical environments could thus still occupy points toward the paranoid ends of these continua, but not at the extremes. This set of ideas has attracted renewed attention in recent years (Linscott & van Os, 2010), provoking considerable controversy (e.g., Braithwaite, 2008; Freeman, 2008). Active debate continues about proposals that symptoms such as paranoia, when they manifest persistently at distressing but still subclinical levels, be assigned to diagnostic categories such as “attenuated psychosis syndrome” (Carpenter, van Os, & van Os, 2011) that signify elevated risk for progression to active psychosis.

Epidemiological data lend support to the idea that paranoid thought might be best conceptualized as distributed along continua such as those that Strauss (1969) proposed, with relatively few individuals clustered at the extremes. In particular, much evidence indicates that full-blown, clinically significant delusions of persecution are rare at any stage of development. For example, although nearly 10% of 6,455 12-year-old participants in the Avon Longitudinal Study of Pregnancy and Childhood responded “yes” or “maybe” when asked if they believed others were persecuting them, follow-up interviews confirmed clinically significant paranoid delusions of persecution in only 0.3% (Horwood et al., 2008). Similarly, in a careful epidemiological study of adults’ responses to screening questions about nonaffective psychotic symptoms, only 0.8% of 2,232 participants endorsed such beliefs; further probing suggested that roughly a fifth of those individuals who reported delusions of persecution were reporting culturally appropriate, realistic, or odd but nonpsychotic thoughts (Kessler et al., 2005). Comparable rates were evident in a study of 894 nondemented older adults (aged 70–82 years) in Sweden, with paranoid symptoms evident in less than 2% of the sample, regardless of whether the participant or a close relative or friend served as the informant (Sigström et al., 2009).

“Normal-range” paranoid thinking or suspiciousness, in contrast, appears much more pervasive (for a review, see Freeman, 2007). Horwood and colleagues’ (2008) findings suggest that a sizable proportion of healthy youths are vulnerable to at least mild or transient beliefs that others intend them harm, and studies of community and college-student samples of adults across varied cultural contexts yield similar results. Research examining data from different waves (2000 and 2008) of the Adult Psychiatric Morbidity Survey in England,
for example, has shown that even when individuals with probable psychosis were excluded, 18.6% to 21.2% of the general adult population endorsed broad beliefs that people were “against them” and 1.5% to 1.8% reported that there had been plots to cause them serious harm (Freeman et al., 2011; Johns et al., 2004). Chan and colleagues (2011) found comparable, if slightly lower, rates of normal-range paranoid ideation in a sample of nearly 5,000 Chinese undergraduates.

In one of the few adult epidemiological studies to distinguish between plausible and implausible suspicious beliefs, van Os, Hanssen, Bijl, and Ravelli (2000) found that 8.7% of 18- to 64-year-old participants in the Netherlands Mental Health Survey and Incidence Study reported persecutory or other delusional experiences that clinicians deemed “not clinically relevant” and a further 3.8% reported similar beliefs that the researchers identified as plausible. Notably, such subclinical experiences proved persistent for a proportion of the sample; at 2-year follow-up, 8% of individuals who had previously reported subclinical symptoms such as delusions reported that their symptoms had recurred (Hanssen, Bak, Bijl, Vollebergh, & van Os, 2005).

**Theoretical models of paranoid thought**

With remarkable prescience about the now mounting evidence that paranoid patterns of thought occur regularly outside of the context of psychosis, Cameron (1943) published a cogent review in which he conceptualized paranoid thought as a “disorder of interpretation” (p. 220) that is not limited to individuals who are actively psychotic. He noted that “the possibilities of misinterpretation are being continuously realized in every normal person from a very early age” (p. 220), such that “even normal persons can build up fairly comprehensive systems of misrepresentation in a very little time” (p. 221). In Cameron’s view, it is critical that considerations of paranoia take into account the roles not only of biological diatheses and psychological processes but also of social and environmental contexts that may foster the emergence of paranoid thought. In its clear integration of processes at multiple levels of analysis as contributing factors to adaptive and potentially disabling paranoid ideation in adulthood, it makes sense with regard to the label of paranoid. However, if we are to prevent or circumvent the emergence of distressing paranoid thinking in childhood and adolescence may also stem partly from concerns about the risks of prematurely or inaccurately attaching potentially damaging labels to individual children; such risks have been the topic of ample debate over the past several decades (e.g., Leigh, 1983; MacCulloch, 2010). Given the high rates of childhood exposure to bullying, abuse, and other experiences that may lead youths to rationally anticipate negative treatment from others, these concerns make sense with regard to the label of paranoid.

Fortunately, decades of careful study of persecutory delusions and related patterns of thought among adults has yielded evidence that multiple cognitive processes whose developmental trajectories are well mapped may contribute to the formation of persecutory delusions not only among adults with schizophrenia (Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001; Blackwood et al., 2004; Blackwood, Howard, Bentall, & Murray, 2001) but also in patient samples that cross a range of diagnostic categories (Bentall et al., 2009). A related and rapidly expanding line of research suggests further that the same processes may be relevant to the emergence of paranoid thinking in nonclinical populations as well (Freeman, 2007; Freeman et al., 2005; Green et al., 2011). These perceptual and inferential processes, some of which are more consistently documented than others, include attention and recall biases for threat cues (Bentall, Kaney, & Bowen-Jones, 1995; Kaney, Wolfenden, Dewey, & Bentall, 1992), a tendency to jump to conclusions when engaging in probabilistic reasoning (So et al., 2011), and exaggerated and labile causal and intent attribution biases (Bentall & Kaney, 2005; Combs et al., 2009; Kinderman & Bentall, 1997; Randall, Corcoran, Day, & Bentall, 2003).

Several additional cognitive processes, which overlap to varying degrees with those noted above, have also been implicated in paranoid thinking; these include deficits in theory of mind (Mehl et al., 2010; Randall et al., 2003), intolerance of ambiguity
(Bentall & Swarbrick, 2003), biased use of the availability heuristic (Bennett & Corcoran, 2010), and elevated self-focused attention (Bodner & Mikulincer, 1998; Fenigstein, 1995).

The following sections first briefly summarize current ideas regarding the neural correlates of adaptive, appropriately trusting interpersonal perception, social cognition, and behavior across development. We then examine several component cognitive processes thought to contribute to paranoid thinking. Highlighting both evidence regarding their neural mechanisms in both normative and atypical development and, where possible, data regarding their independent links to aggressive behavior. Each of the processes listed in the preceding paragraph is worthy of detailed attention; however, in the interest of brevity and clarity, we elected to focus on only three. Each of these three is relatively narrow in scope, yet incorporates multiple aspects of social cognition; has been examined to some degree across development; and has some documented evidence regarding potentially relevant neural mechanisms: (a) attention bias for threat, (b) jumping to conclusions biases, and (c) hostile intent attribution biases.

Neural Mechanisms of Adaptive Interpersonal Perception and Behavior

Early social development comprises an array of interrelated tasks that include learning to regulate one’s own behavior, thoughts, and emotions in the context of rapidly changing interpersonal cues that vary in their personal relevance for each individual and that signal, with varying degrees of clarity, whether another’s intentions are hostile or benign (Bradley, 2000; Shonkoff & Phillips, 2000). Mastering these skills is of critical importance for effective social functioning. Individuals who overlook or misread salient interpersonal cues are at heightened risk for inaccurately assuming or anticipating that others mean them harm; individuals who struggle to regulate their emotional, cognitive, and behavioral reactions to social cues, particularly those that they have misread, are vulnerable to respond inappropriately, potentially with aggression (Blackwood et al., 2001; Crick & Dodge, 1994). Mastery of these social cognitive tasks and consequent diminished risk of interpersonal problems is most likely to occur in children who inhabit supportive environments (Landry & Smith, 2010) and who also experience healthy maturation of the social information processing network (SIPN), a core set of functionally linked neural structures, and a diverse group of additional regions with which the SIPN component structures share connections (Blakemore, 2008; Burnett, Sebastian, Cohen, & Blakemore, 2011; Nelson, Leibenluft, McClure, & Pine, 2005; Paterson, Heim, Friedman, Choudhury, & Benasich, 2006).

Although this loosely defined network functions as a unified whole, Nelson and colleagues (2005) parse it descriptively into three dynamically interactive “nodes,” each of which plays distinct roles in processing and responding to social cues. Processing initiates in structures within the detection node, which support the identification and decoding of salient cues in the social environment. Structures within this node, some of which show sophisticated patterns of function as early as the first 3 months of life, include the superior temporal sulcus, fusiform face area, and inferior temporal and occipital cortices (Johnson et al., 2005; Tzourio-Mazoyer et al., 2002).

Once social cues have been perceived, structures within the affective node, including, but not limited to, the amygdala, ventral striatum, septum, and ventral tegmental area, evaluate their reward or punishment value and support the generation of relevant and appropriate responses (Adolphs, Tranel, & Damasio, 1998; McGinty et al., 2011; O’Doherty, 2011). Of note, the amygdala, septum, and ventral tegmental area have also been implicated in the generation and maintenance of trusting behavior in adults (Koscik & Tranel, 2011; Krueger et al., 2007). The early developmental course of human structures in this node is still poorly understood; however, evidence from studies of nonhuman species suggests that their functional properties, like those of structures in the detection node, may come on line in the first year of life (Schumann, Bauman, & Amaral, 2011; Stopczynski, Poloskey, & Haber, 2008; Tau & Peterson, 2010).

These brain regions appear to grow steadily, undergoing changes in structure and function at the onset of puberty (Brenhouse & Andersen, 2011; Burnett et al., 2011; Giedd, Castellanos, Rajapakse, Vaituzis, & Rapoport, 1997; Giedd et al., 1999). Notably, as Ernst and Fudge’s (2009) triadic model of motivated behavior suggests, structures within the affective node appear to follow distinct developmental trajectories during adolescence that lead to imbalances in function. Specifically, striatally mediated reward responsiveness takes precedence over amygdala-mediated punishment or threat avoidance, as a function of differential patterns across regions of both anatomical development and regulatory input from prefrontal cortical structures.

More complex integrative functions, such as attributing mental states to others, inhibiting prepotent responses, and generating goal-directed behavior, appear to engage structures within the cognitive-regulatory node, which is distributed primarily across regions of the prefrontal cortex (PFC; Nelson et al., 2005). Considerable evidence indicates that prefrontal areas within the cognitive-regulatory node, as well as their reciprocal white matter projections to the affective/detection nodes and other salient regions, continue to develop throughout childhood and into early adulthood, paralleling the emergence of increasingly sophisticated cognitive skills (Asato, Terwilliger, Woo, & Luna, 2010; Gogtay et al., 2004). Given the complexity of the functions that this node mediates, it is likely that additional brain regions also support their implementation.

Recent reviews, for instance, highlight evidence that areas within the insula and anterior cingulate cortex (ACC) may operate as a “saliency network” that triggers prefrontal attentional, control, and response processes as a function of select input from sensory components of the detection node (Med-
Evidence is mixed regarding whether and how attention bias for threatening or other emotional cues changes across the life span, and the conclusions that can be drawn are limited by the relatively small number of studies of children and older adults (Bar-Haim et al., 2007). Although some have proposed that a threat bias is normative early in development but decreases as children mature (Kindt, Bierman, & Broschot, 1997; Kindt, van den Hout, de Jong, & Hoekzema, 2000), studies vary in the degree to which their findings conform to this proposal (Monk et al., 2006; Morren, Kindt, van den Hout, & van der Kasten, 2003; Waters, Lipp, & Spence, 2004). In adults, attention biases for threat have been found most consistently in the context of varied types of psychopathology (e.g., Bar-Haim et al., 2007; Gotlib et al., 2004; Mogg & Bradley, 1998); conversely, attention biases for positive cues appear to strengthen with increasing age in healthy samples (Lindstrom et al., 2009; Mather & Carstensen, 2005).

Neural mechanisms of attention bias for threat have received some research attention, particularly in anxious adults and adolescents. Bishop (2008) proposed an elegant and comprehensive model that describes putative anxiety-associated neural mechanisms of a selective bias for attention to threat cues. According to this biased competition model, “lower-order” brain regions such as the amygdala, which detect and evaluate stimulus salience, and “higher-order” cortical regions, including the lateral PFC and anterior cingulate, which exert control over perceptual attention and cognitive processing, compete for resources in the presence of perceived threats. Anxiety modulates the function of this system at different levels that span the detection, affective, and cognitive-regulatory nodes that implement healthy social thought and behavior. First, state anxiety amplifies the amygdala response to a detected threat. Second, trait anxiety disrupts attentional, inhibitory, and executive functions mediated by the ACC and lateral PFC. Bishop’s (2008) model focuses explicitly on ways in which anxiety may contribute to biased attentional processing for threat cues; given evidence that individuals vulnerable to paranoid patterns of thought also commonly demonstrate high levels of anxiety or other negative affect (Freeman, Brugha, Meltzer, Jenkins, Stahl, & Bebbington, 2010; Tone et al., 2011), it seems plausible that a comparable, if not identical, biased competition model operates in threat attention biases associated with paranoia.

Neuroimaging research findings to date have yielded a range of findings that lend mixed support to Bishop’s (2008) model (Monk et al., 2004, 2006; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2006; Telzer et al., 2008). Interpretation of this small literature, however, is complicated by methodological and sampling variations across studies. In healthy adults, Monk and colleagues (2004) found evidence of a progressively increasing bias away from angry faces that was accompanied by increases in occipitotemporal activation. In contrast, using a different task paradigm that incorporated fearful rather than angry faces as stimuli, Pourtois and colleagues (2006) found preferential activation in a net-
### Table 1. Summary of relevant neuroimaging findings

<table>
<thead>
<tr>
<th>Study (Grouped by Domain)</th>
<th>Behavioral Task</th>
<th>Neuroimaging Results</th>
<th>Population Sampled</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attention Bias for Threat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monk et al. (2004)</td>
<td>Attention bias for angry/happy faces</td>
<td>Increases in bias away from angry faces correlated positively with occipitotemporal activation</td>
<td>Healthy adult</td>
<td>$N = 12$</td>
</tr>
<tr>
<td>Pourtois et al. (2006)</td>
<td>Covert orienting toward fearful/ happy faces</td>
<td>Increased activation in intraparietal and orbitofrontal cortex for targets following invalid fearful faces; increased activation in right lateral occipital cortex for targets following valid fearful faces</td>
<td>Healthy adult</td>
<td>$N = 15$</td>
</tr>
<tr>
<td>Telzer et al. (2008)</td>
<td>Attention bias for angry/happy faces</td>
<td>Trait anxiety associated with increased right DLPFC activation on trials reflecting attentional bias for angry faces</td>
<td>Healthy adolescent</td>
<td>$N = 20$</td>
</tr>
<tr>
<td>Monk et al. (2006)</td>
<td>Attention bias for angry/happy faces</td>
<td>Anxiety disorder associated with both increased ventral IPFC activation to angry face trials and bias away from angry faces</td>
<td>Adolescent (clinical and healthy)</td>
<td>$N = 18$ (anxiety disorder); $n = 15$ (control)</td>
</tr>
<tr>
<td>Lindstrom et al. (2009)</td>
<td>Attention bias for angry/happy faces</td>
<td>Nonsignificant correlations between age and angry-face attention bias scores and brain activation to angry faces; age correlated with left cuneus and left caudate activation in association with bias toward happy faces</td>
<td>Healthy adult/adolescent</td>
<td>$N = 37$ (aged 9–40 years)</td>
</tr>
</tbody>
</table>

| **Jumping to Conclusions Bias** | | | | |
| Moore & Sellen (2006)      | Beads task | Simulation results consistent with increased activity in the mesolimbic dopamine system (ventral striatum, ventral tegmental area of the midbrain) | Adult (simulated models of neural activity based on data from five published studies) | NA |

| **Attributional Bias** | | | | |
| Blackwood et al. (2004)   | Determine self-relevance of ambiguous or unambiguous neutral and threatening statements | Patients showed increased posterior cingulate gyrus activation and decreased rostral–ventral ACC activation during self-relevance evaluations relative to controls. | Adult (clinical and healthy) | $N = 8$ (delusional patients with schizophrenia); $N = 8$ (control) |
| Lombardo et al. (2009)    | Make “mentalizing” or “physical” judgments about self or a familiar nonclose other (the British Queen) | Preferential activation during mentalizing (vs. physical) judgments in ventral mPFC, PCC, TPJ, anterior temporal lobe, and primary sensorimotor cortex | Healthy adult (male only) | $N = 33$ |
| Mitchell et al. (2005)    | Evaluate “how pleased” or “how symmetrical” faces appeared | Preferential activation of dorsal mPFC during mentalizing (“how pleased”) trials | Healthy adult | $N = 18$ |
| Wolf et al. (2010)        | View a 15-min film and make inferences about the mental states or physical conditions of movie characters | Increased left TPJ and left precuneus activation during mental state inferences | Healthy adult | $N = 19$ |
| Young et al. (2011)       | Determine whether to blame or praise others based on conjunctions of harmful, helpful, and neutral intentions and outcomes | TPJ (especially right hemisphere) and mPFC activation strongest when participants blamed or withheld praise on the basis of an actor’s reported intentions | Healthy adult | $N = 17$ |

*Note:* DLpFC, dorsolateral prefrontal cortex; IPFC, lateral prefrontal cortex; ACC, anterior cingulate cortex; mPFC, medial prefrontal cortex; PCC, posterior cingulate cortex; TPJ, temporoparietal junction.
work of parietal, occipital, and orbital prefrontal regions in response to trials that incorporated fearful, but not happy, face stimuli.

Monk et al. (2006) and Telzer et al. (2008) each focused on adolescents; however, whereas Monk and colleagues (2006) compared youths with generalized anxiety disorder and controls, Telzer et al. (2008) focused on associations between normal-range trait anxiety and attention bias for angry faces in a healthy sample. Each study yielded suggestions that varied PFC perturbations relate separately to anxiety and to biased attention toward or away from threat. In the one study to directly examine correlations among age, neural activity patterns, and behaviorally observed attention biases in both healthy adolescents and healthy adults, Lindstrom and colleagues (2009) found no evidence of an association between age and either behaviorally observed threat biases or neural activity associated with a bias to attend to threat.

Taken together, these findings from a small pool of neuroimaging studies suggest broadly that perturbations in different prefrontal regions, as well as a variety of areas across the temporal, parietal, and occipital lobes, may support the emergence of biased attentional processing of threat cues. No clear developmental patterns of change in either attention bias itself or in putative neural mechanisms have emerged yet from research, but Lindstrom and colleagues’ (2009) study comparing adults and adolescents provided an important first step toward clarifying the neurodevelopmental course of this cognitive characteristic. More study is needed, especially with a specific focus on how these biases relate to paranoid or persecutory ideation in addition to or instead of anxiety.

It also remains unclear whether and how attention bias for threat per se might relate to elevated risk for aggression among individuals who experience paranoid types of thought. The few studies to date to examine attention bias and aggression have focused primarily on biases for aggression-related words in adults who show other risk factors for violence, including alcohol intoxication (Gallagher & Parrott, 2011) or Type A personality characteristics (Faunce, Mapledoram, & Job, 2004). Research examining the links between heightened attention to cues of threat to self and risk for aggression, particularly in the context of paranoid thinking, would help advance this literature and lay the foundation, if associations emerged, for neuroimaging research targeted toward identifying their neural correlates.

**Jumping-to-conclusions (JTC) bias**

Garety and colleagues (e.g., Garety & Freeman, 1999; Garety & Hemsley, 1994) coined the phrase “jumping-to-conclusions bias” to describe their findings that adults with delusions show an elevated tendency to rapidly accept hypotheses on the basis of limited evidence and, to a lesser degree, a tendency to reject hypotheses more quickly than controls. Most research on this bias has used iterations of the experimental “beads task,” which requires individuals to guess which of two jars (each holding beads of two colors in different ratios) a series of beads is drawn from. In a widely used variant, they are permitted to request as many beads as they would like before making their guess (Garety & Freeman, 1999). Findings commonly (but not uniformly; see, e.g., Colbert, Peters, & Garety, 2010) show a bias to jump to conclusions or to request a small number of beads before making a decision among roughly half of adults who experience delusions, compared to only 10% to 20% of nondelusional adults (for reviews, see Fine, Gardner, Craigie, & Gold, 2007; Freeman, 2007; Garety, Bebington, Fowler, Freeman, & Kuipers, 2007). At least one study demonstrates that the JTC bias is more prevalent among adults who are specifically prone to persecutory delusions than among nonclinical controls (Startup, Freeman, & Garety, 2008).

In the limited number of studies to date that examine JTC biases in nonclinical populations, consistent associations with paranoid ideation emerge. Freeman, Pugh, and Garety (2008), for example, found a significant association between a JTC bias and vulnerability to paranoid thoughts in a sample of 200 adults drawn from the community. Similar findings appeared in a study of 92 adults, with participants who endorsed high levels of paranoid ideation more quickly generating decisions than members of moderate- and low-paranoia groups during a computerized social reasoning measure modeled on the beads task (Lincoln, Salzmann, Ziegler, & Westermann, 2011).

Psychological mechanisms underlying this bias remains unclear; theorists have postulated variously that it stems from a need for closure or avoidance of ambiguity (Bentall et al., 2001), from a desire for rapid confirmation of threats in the environment (Dudley & Over, 2003), from a tendency to overattribute salience to currently experienced stimuli (Monk, Pomarol-Clotet, McKenna, & McCarthy, 2006), or from an impaired ability to process and use sequential information (Young & Bentall, 1995). In their thoughtful review of the literature, Fine and colleagues (2007) point out weaknesses and unclear aspects of each of these hypotheses and call for greater theoretical precision in considerations of the JTC bias. Behavioral correlates of the JTC bias also have yet to be elucidated; it is thus not clear whether and how a JTC bias might relate to risk for aggressive behavior.

Very little empirical information is available regarding the neural mechanisms of the JTC bias. Moore and Sellen (2006) put forth a neural network model that implicates enhanced or overactive mesolimbic dopamine (DA) system functioning in the JTC bias in delusional patients and that may have applicability to healthy adults as well. The stress-sensitive mesolimbic DA system comprises the ventral striatum and the ventral tegmental area of the midbrain; the ventral tegmentum provides the ventral striatum, a core structure within the affective node of the SIPN, with dopaminergic input (Trainor, 2011). Dysfunction in this system or its inputs has been implicated in an array of psychological conditions, including schizophrenia (Stahl, 2007), substance abuse and addiction (Kauer & Malenka, 2007), psychopathy (Buckholtz et al., 2010), and depression (Nestler & Carlezon, 2006).
Simulations of behavioral differences between delusional individuals and controls during performance of the bead task using Moore and Sellen’s (2006) model suggest that in the context of delusions the elevated mesolimbic DA alters representations of stimulus salience, which influences their subsequent processing by regulatory and attentional systems based in the PFC. It is notable that the results of simulations were inconsistent with dysfunction in higher order attentional processes; instead, the authors suggest that the altered salience representations generated via the striatum may differentially influence the allocation of attention.

This model is loosely consistent with the hypothesis that the JTC bias stems from a tendency to overattribute salience to environmental cues (e.g., Menon et al., 2006); empirical research comparing neural activation between actively paranoid and control adults during paradigms modeled on the bead task thus might be hypothesized to yield evidence of group differences in patterns of ventral striatal activation during the decision-making process. Ernst and Fudge’s (2009) triadic model of the neural substrates of motivated decision making might further suggest that these differences would be less pronounced during adolescence, when striatal responsiveness is generally enhanced and prefrontal cortical attentional and control processes are still maturing. Earlier in development, it is unclear whether the JTC bias would manifest differentially among youths at either behavioral or neural levels. Research findings, however, show evidence of motivated decision making as early as 24 months of age (Kenward, Folke, Holmberg, Johansson, & Gredebäck, 2009), which suggests that it may be possible to examine normative and biased decision-making patterns well before youths reach adolescence.

Whether and how the JTC bias might relate to elevated risk for aggression remains unstudied. However, given the findings that link more broadly defined cognitive and behavioral impulsivity and aggressive behavior (Elliott & Mirsky, 2002), it is plausible that this cognitive pattern might contribute to an individual’s vulnerability to aggress. Evidence that anomalies in the ventral striatum structure and function may also relate to both impaired processing of threat cues in humans and violent behavior in some species also make this set of associations worthy of further examination (Calder, Keane, Lawrence, & Manes, 2004; Ferrari, van Erp, Tornatzky, & Miczek, 2003).

Attributional bias

In the mid-20th century, personality theorists focused considerable attention on the idea that the appraisals and expectations that an individual attaches to a stimulus are as important as or more important than actual stimulus characteristics in determining the responses that the stimulus elicits (e.g., Bowers, 1973; Harvey, Hunt, & Schroder, 1961; Mischel, 1973). Nashby, Hayden, and DePaulo (1980) conducted two studies testing this theory in adolescent boys hospitalized for emotional problems; specifically, they examined the associations between aggressive behavior and a tendency to misattribute hostility to nonverbal interpersonal cues. Their findings, which were consistent with the presence of what they termed “a generalized and marked attributional bias to infer hostility” (p. 465), converged with those of Dodge (1980), who found that aggressive, but not nonaggressive, school-aged boys were likely to respond in a hostile manner to provocative peer behavior that was ambiguous in its intent and to assume that peer hostility drove such ambiguous behavior.

These landmark papers initiated a cascade of research on the hostile attributional style, which shares notable similarities with paranoid patterns of thought. Over the subsequent three decades, dozens of studies have explored the construct, its developmental antecedents, and its associations with behavior, particularly aggression (for extensive and thorough reviews, see De Castro et al., 2002; and Dodge, 2006). Dodge and colleagues have built a particularly impressive body of longitudinal research that elucidates the patterns of association between hostile attribution biases measured just prior to kindergarten and aggressive behavior at numerous points throughout childhood and adolescence (Dodge, Bates, & Pettit, 1990; Dodge, Pettit, Bates, & Valente, 1995; Lansford, Malone, Dodge, et al., 2006; Pettit, Bates, & Dodge, 1997).

A few recent studies also show evidence that hostile intent attribution biases constitute a core element of paranoid thought, both in healthy samples (Combs, Penn, Wichert, & Waldheter, 2007; Fornells-Ambrojo & Garety, 2009) and in adults with persecutory delusions (Combs et al., 2009). The findings from one study in patients with either current or past persecutory delusions suggest that negative attribution biases such as these may be particularly pronounced in the context of active paranoid thought (Lincoln, Mehl, Exner, Lindenmeyer, & Rief, 2010). However, how such findings might map onto normal-range paranoia is unclear.

Note that hostile intent attribution biases can function in both adaptive and maladaptive ways. For some individuals in chronically punitive environments, for example, intent attribution biases may foster resilience by diminishing risk for developing internalizing problems such as anxiety and depression that might emerge more commonly among those prone to self-blame (Lansford, Malone, Stevens, et al., 2006). Prolonged anticipatory hostile behavior from others, however, is also associated with elevated risk for engaging in aggression, particularly reactive physical aggression, which constitutes retaliatory acts in response to perceived environmental threats (Crick & Dodge, 1994; DeCastro et al., 2002; Dodge, 1980, 2003). In addition, although findings are less consistent than for physical aggression, some researchers have found relationships between hostile intent attribution biases and relational aggression toward others (Mathieson et al., 2011; Yeung & Leadbeater, 2007; but see also Crain, Finch, & Foster, 2005; Nelson, Mitchell, & Yang, 2008).

Hostile intent attribution biases appear to constitute one link between early negative life experiences that vary from peer rejection to physical abuse and later aggression (DeWall, Twenge, Gitter, & Baumeister, 2009; Dodge et al., 1995; Re-
Blackwood and colleagues (2004) found different patterns of peak activation between groups in both frontal and posterior regions of the brain. Whereas both groups showed significant activation during both neutral and threatening statements in the temporal poles, only controls showed elevated activation in the anterior paracingulate cortex and both medial and superior frontal cortical areas, and only patients showed elevated activity in the posterior cingulate. However, the only direct contrasts of patients and controls to yield evidence of significantly stronger neural responses found greater anterior cingulate activation in controls and greater posterior cingulate activity in patients.

Thus, although the literature is still small, findings are beginning to suggest roles for a range of frontal, cingulate, and temporoparietal regions in the evaluation of and generation of responses to others’ intentions, consistent with involvement of the cognitive-regulatory and affective nodes. Developmentally oriented neuroimaging research is much needed in this area; most work to date focuses on adults, although a few researchers have begun to compare patterns of activation during mentalizing tasks between healthy adolescents and adults (e.g., Burnett & Blakemore, 2009). In addition, more work directed specifically at examining the neural correlates of explicitly hostile intent attribution biases, rather than simply attributions of intent, would help clarify the emergence of atypical and typical patterns of social cognition and their links with aggressive behavior.

**Discussion and Conclusions**

Research on paranoid thinking and its potential association with aggression that adopts a developmental psychopathology perspective is only beginning to emerge; research on the ways in which biological substrates figure in the development of a paranoid style of cognition, particularly outside of the context of psychosis, is even more limited. By examining the data regarding the behavioral and neural correlates of varied cognitive processes that are likely components of a paranoid thinking style, it may be possible to advance both theoretical and empirical research in this domain. The present article constitutes an effort to summarize a select portion of the work on paranoid thought and a few of its cognitive components that have been articulated more or less thoroughly across development. A few key points emerge to guide future research.

First, although the research base is small, convergent evidence suggests that dysfunction in the SIPN, particularly the affective and cognitive-regulatory nodes and related structures, may be associated with the emergence of paranoid thinking across the spectrum from normal range to clinically significant, at least in adults. Prefrontal involvement is the most common finding across component cognitive processes and paradigms under study; however, structures that participate in the evaluation of stimulus valence, salience, and threat value also appear worthy of further attention. It may be of value to combine novel paradigms that elicit paranoid types
of thought, such as a virtual reality task that Freeman, Pugh, Vorontsova, Antley, and Slater (2010) developed to simulate real life neutral and hostile interactions, with neuroimaging tools to better characterize what occurs in the brain when paranoid thoughts are active.

Second, it will be critically important for researchers to clarify when paranoid patterns of thought, both clinically significant and normal range, first emerge in development and what factors produce and maintain them. The hostile intent attribution bias literature provides much relevant data; however, because the conceptual overlap between hostile intent attribution biases and paranoia is incomplete, many questions remain about how early and under what circumstances an individual might begin to think consistently in a paranoid way. Similar developmental extension of the neuroimaging literature to younger samples would be of value; researchers have had impressive success in completing functional scans of children as young as preschoolers (e.g., Gaffrey et al., 2011; James & Maouene, 2009), and replication of adult studies across the early life span would provide useful data regarding patterns of change and consistency.

Third, although links between paranoid thought and aggression are unclear and appear to be mediated by an array of variables, there is some evidence of consistency between patterns of brain function and dysfunction associated with the two. Regions of the PFC, for instance, which have been implicated in the emergence of reactive aggression in some samples (Blair, 2010), might show similarly atypical activity in the context of active paranoid or suspicious thought. A number of neural circuits that show links to impulsive aggression also overlap with those that appear relevant to paranoid types of thought (Coccaro, Sripada, Yanowitch, & Phan, 2011). Adaptation of interactive neuroimaging paradigms, such as the rejection-simulating cyberball task (Eisenberger, Lieberman, & Williams, 2003), that could both elicit paranoid ideation and provide opportunities for aggressive, neutral, or prosocial responses, might help clarify areas of convergence and divergence in the neural substrates of paranoid thinking and aggression.

Fourth, given that findings relevant to the understanding of paranoid thought and its neural mechanisms are emerging from studies of both clinical/high-risk (e.g., Kumari et al., 2011; Tost et al., 2010; Whalley et al., 2007; Williams et al., 2007) and healthy samples (e.g., Prévost et al., 2011), cross-talk between researchers in both realms will facilitate much more rapid advancement of the field. Similar collaborations among researchers focused on different developmental periods, with attention to finding common conceptual ground in the absence of consistent labels, would be of value. For example, bringing together scholars who study the normative development of trust in children (e.g., Sharp, Ha, & Fonyagy, 2011) and those who study the correlates of extreme mistrust (i.e., persecutory delusions) may lead to novel and illuminating hypotheses and findings.

References


