Serotonergic Function, Two-Mode Models of Self-Regulation, and Vulnerability to Depression: What Depression Has in Common With Impulsive Aggression

Charles S. Carver, Sheri L. Johnson, and Jutta Joormann
University of Miami

Evidence from diverse literatures supports the viewpoint that two modes of self-regulation exist, a lower-order system that responds quickly to associative cues of the moment and a higher-order system that responds more reflectively and planfully; that low serotonergic function is linked to relative dominance of the lower-order system; that how dominance of the lower-order system is manifested depends on additional variables; and that low serotonergic function therefore can promote behavioral patterns as divergent as impulsive aggression and lethargic depression. Literatures reviewed include work on two-mode models; studies of brain function supporting the biological plausibility of the two-mode view and the involvement of serotonergic pathways in functions pertaining to it; and studies relating low serotonergic function to impulsiveness, aggression (including extreme violence), aspects of personality, and depression vulnerability. Substantial differences between depression and other phenomena reviewed are interpreted by proposing that depression reflects both low serotonergic function and low reward sensitivity. The article closes with brief consideration of the idea that low serotonergic function relates to even more diverse phenomena, whose natures depend in part on sensitivities of other systems.

Keywords: serotonin, depression, two-mode models, dual-process models, self-regulation

Depression is a disorder that afflicts countless people worldwide. A substantial literature has examined the role of neurotransmitters in depression; research has also turned to molecular genetics as a tool to gain leverage on neurobiological processes underlying depression. This body of neurobiological work has yielded valuable evidence. An important question remains, however, about how to interpret the evidence in psychological terms. A confluence of ideas from other areas of work suggests an answer to that question. This article reviews some of those literatures (almost exclusively human research) and advances a psychological viewpoint that may help in interpreting the neurobiological evidence.

A central idea examined here is that low serotonergic function is a marker of a deficit in executive control processes, processes that override or inhibit lower-level influences on behavior. A result is that persons with low serotonergic function (and thus diminished executive control) are especially responsive to associative and affective cues of the moment. Lack of executive control ultimately has divergent effects on behavior, however, because what follows depends on other aspects of the person’s makeup. In the cases that are most familiar, a deficit in executive control yields impulsive action and is associated with externalizing disorders. The position advanced here is that in depressed persons, the same executive control deficit often (though not always) leads to automatic (and in that sense “impulsive”) action, plus absorption in negative emotions.

It should be clear from this statement that the focus on serotonergic function in this article is not intended to imply that serotonin function is the entire story behind depression (or externalizing disorders either, for that matter). Dopamine function also appears to play a role in depression (Naranjo, Tremblay, & Busto, 2001). Indeed, the very idea that low serotonergic function and deficits in executive control affect different people in quite different ways clearly implies that there is an interaction among causal influences, leading some but not others to experience depression in the context of serotonergic deficits.

The view that is advanced here relies on several quite different sources of information, which are described in separate sections. One source is a set of theories from cognitive, social, personality, motivational, and developmental psychology, which converge on the idea that there exist two modes of processing information and regulating action, which operate simultaneously and in competition with each other. This set of theories forms the psychological perspective from which the other literatures are considered. A second body of evidence, considered next, describes functions of certain brain regions and the interplay among them, which appear consistent with the two-mode models, and thus supports the biological plausibility of the two-mode view.
The review then turns to cognitive and behavioral effects related to differences in serotonergic function, with the goal of mapping these effects onto the two-mode models. First described are studies of cognitive processing of emotional stimuli, impulsiveness in behavior, impulse-related disorders, and self-reports of several aspects of personality, all of which are then shown to reflect principles in the two-mode models. Next is evidence linking low serotonergic function to depression vulnerability. These findings represent a striking contrast to much of what was reviewed just previously, because the behavioral qualities now being related to low serotonergic function are quite different. This divergence between literatures is interpreted by incorporating another variable: sensitivity of approach motivation. The result is an interactive model, with two classes of potentially problematic outcomes (poor impulse control vs. depression).

The review continues by addressing additional neurobiological and behavioral findings concerning depression, which offer further support for the two-mode interpretation. Finally a set of remaining issues and further directions is considered.

Two-Mode Models of Functioning

Several theoretical models have emerged in diverse areas of psychology over the past two decades which posit that that people experience the world through two simultaneous but somewhat distinct modes of processing (Barrett, Tugade, & Engle, 2004; Epstein, 1994; J. St. B. T. Evans, 2008; Kahneman, 2003; Stanselich, 1999; Toates, 2006). These two processing modes may use different aspects of available information (Rudman, Phelan, & Heppen, 2007). There is also evidence that the two modes learn in different ways, and that the two patterns of learning create parallel and competing paths to action, which require continuous arbitration (Daw, Niv, & Dayan, 2005).

In cognitive psychology, this idea helped deal with a clash between two views of cognition. One view treats cognition as sequential symbol processing; another view assumes simultaneous parallel processing (e.g., Bechtel & Abrahamsen, 1991; Dawson, 2005; J. L. McClelland, 1999). Many cognitive psychologists now believe that both views are partly correct, that cognition (broadly conceived) uses two kinds of processes. One process—effortful, top-down, symbolic, and reflective—is used for planning and strategic behavior. The other—automatic, reflexive, bottom-up, and associationist—is used for acts that are heuristic, skilled, or urgent (Norman, 1986; Shastri & Aijanagadde, 1993; Sloman, 1996; Smolensky, 1988; Smolensky & Legendre, 2006).

Two-mode models (also labeled dual-process models) exist in the literature of personality psychology as well. Indeed, Epstein’s (1973, 1985, 1990, 1994) cognitive–experiential self theory may have been the first explicit two-mode model in contemporary psychology. It was based on the premise that humans experience reality via two systems. One is a symbolic processor—the rational mind. The other is associative and intuitive, uses short cuts and heuristics, and functions automatically and quickly. Epstein argued that both systems are always at work and that they jointly determine behavior.

Metcalfe and Mischel (1999), drawing on several decades’ work on delay of gratification, proposed a similar model. Delay of gratification research poses a choice between a smaller, less desired, but immediate reward versus a larger, more desired reward later (Mischel, 1974). Metcalfe and Mischel (1999) proposed that two systems determine the ability to restrain in this and many other contexts: a “hot” system (emotional, impulsive, reflexive, and connectionist) and a “cool” system (strategic, flexible, slower, and unemotional). How a person responds to a difficult situation depends on which system presently dominates.

A conceptually similar line of argument about two modes of functioning also exists in social psychology. An early step in this direction was the argument that there are two paths to persuasion, one involving careful and thoughtful processing and the other involving heuristics and associations (Chaiken & Trope, 1999; E. R. Smith & DeCoster, 2000; Wilson, Lindsey, & Schooler, 2000). More recently, an analysis of automaticity in attribution posited two processing modes that are quite similar to those in the cognitive theories (Lieberman, Gaunt, Gilbert, & Trope, 2002). Strack and Deutsch (2004) proposed a farther-reaching model of overt social behavior, in which action is a joint output of two modes of functioning—reflective and impulsive—that occur simultaneously and may be either mutually supportive or in conflict with each other.

The two-mode idea has made a number of other appearances in social psychology as well. There is a rapidly expanding literature on implicit attitudes, implicit self-concepts, and the like (Custers & Aarts, 2005; Greenwald et al., 2002; McConnell, Rydell, Strain, & Mackie, 2008; Rydell & McConnell, 2006). Indeed, the idea that implicit and explicit mental processes play distinct roles in behavior is being applied to a steadily expanding range of topics, including (for example) moral judgments (Haidt, 2001), reactions to cues of stigma (Pryor, Reeder, Yeadon, & Hesson-McNittis, 2004), temporal discounting ( McClure, Laibson, Loewenstein, & Cohen, 2004), and addiction (Wiers & Stacy, 2006). Of course, the idea that both implicit and self-attributed self-aspects influence behavior in different ways also traces to classic motivation theory (D. C. McClelland, 1984, 1985; D. C. McClelland, Koestner, & Weinberger, 1989; see also Schultheiss & Pang, 2007).

The two-mode idea also has an important presence in developmental psychology. Rothbart and her colleagues (e.g., Derryberry & Rothbart, 1997; Rothbart, Ahadi, & Evans, 2000; Rothbart, Ahadi, Hershey, & Fisher, 2001; Rothbart & Bates, 1998; Rothbart, Ellis, Rueda, & Posner, 2003; Rothbart & Posner, 1985) have argued for the existence of basic temperament systems for approach and avoidance and a third temperament generally termed effortful control (see also Kochanska & Knaack, 2003; Nigg, 2000, 2003, 2006). Effortful control concerns attentional management (both in terms of sensitivity to new stimuli and in terms of the focusing or persistence of attention during long-lasting tasks) along with inhibitory control (the ability to suppress an approach behavior when doing so is situationally appropriate).

This model postulates that effortful control is superordinate to both approach and avoidance temperaments (e.g., Ahadi & Rothbart, 1994; see also Clark, 2005). The label “effortful” conveys the sense that this is an executive, planful activity, entailing the use of cognitive resources to deter the tendency to react impulsively. Rothbart (e.g., Rothbart & Bates, 1998), Eisenberg (e.g., Eisenberg et al., 2004), and Kochanska (e.g., Kochanska & Knaack, 2003) have seen effortful control as dependent on certain cortical functions (see also Nigg, 2001, 2003, regarding what he calls executive inhibition), and a variety of evidence from neuroimaging studies of both adults and
children has supported that argument (e.g., Durston, Thomas, Worden, Yang, & Casey, 2002; Durston, Thomas, Yang, et al., 2002).

**Commonalities Among Theories**

Two-mode theories emerged in various areas for various reasons. Specific theories stress different issues, but they are similar in many respects (for broad views, see Barrett et al., 2004; J. St. B. T. Evans, 2008; Toates, 2006). Across theories, the two modes could be depicted fairly accurately as follows: A reflexive (or reactive or implicit) system, which is nonverbal and associationist, spontaneously creates action when its schemas or production systems are sufficiently activated. Being association based, it is said by many to be intuitive (e.g., Lieberman, 2000). It is often characterized as “emotional.” This label sometimes seems to imply that emotional experience is subjectively salient in this mode and sometimes seems to imply that this mode is very responsive to emotions triggered by situational cues. This system dominates when speed is needed (as when a situation is emotionally charged) and also when processing resources are diminished. That is, it requires relatively little capacity and thus can function under suboptimal conditions (see also De Neys, 2006). It tends to respond to short-term contingencies without consideration for the future or for broader consequences of the action. Thus, although it can be highly adaptive with respect to short-term contingencies, it may be less adaptive when a longer view is beneficial (cf. Daw et al., 2005; Kurzhan & Aktipsis, 2007; LeDoux, 2002).

The reflective (or rational or explicit) system, in contrast, operates mostly consciously, uses logical rules, is verbal and deliberative, and thus is comparatively slow. This rational system provides a more cautious, analytic, planful way of proceeding. It anticipates future conditions, makes decisions based on those anticipations, and forms top-down intentions. It is more wide ranging in its search for relevant information. Because the reflective system requires substantial cognitive capacity, it loses efficiency under high mental load or other conditions that limit cognitive capacity.

Although two-mode theorists tend to characterize the modes partly in terms of differences in emotionality, that point requires clarification. The theorists do not say that planful behavior is devoid of emotion. It is clear that planning ahead (as in delay of gratification) implies evaluating the affective consequences of various courses of action. However, an urgent and intense emotion appears likely to compete with, and to potentially short circuit, the tendency to be planful. The two-mode theories typically do not specify in detail the mechanism by which emotional salience or intensity promotes ascendance of the reactive system. It is generally implied, however, that greater subjective intensity means that the emotion cue is strongly activating an associative network within the reflexive system.

**Behavioral Implications of Two-Mode Models**

As a group, these theories suggest that the two modes, given their different operating characteristics, have differing influences on behavior. Behavior that is managed by what is typically characterized as the associationist, automatic, implicit, or reflexive mode tends to be impulsive. More precisely, this system responds to readily available associative cues of the moment. Behavior that is managed by what is typically characterized as the rational, reflective, sequential, or deliberative mode tends to be planful. It responds to two kinds of parameters—social and temporal—that go beyond immediate cues of incentive or threat.

This difference in operating characteristics tends to translate, for the most part, to a dimension of impulsive responsivity versus planful constraint (Carver, 2005; Nigg, 2001, 2003; Spoont, 1992; Strack & Deutsch, 2004). This characterization requires an important qualification, however: Specifically, in principle, each mode of functioning can promote either action or inaction.

The developmental two-mode model is perhaps most explicit on this point. As described earlier, Rothbart and others posit three temperaments. Approach and avoidance temperaments together form the basis of what is termed reactive control. Reactive control is relatively reflexive and involuntary. The reactions in the term reactive control are automatic tendencies to approach incentives and avoid threats. These tendencies compete (Figure 1), and the resultant overt action is determined by which reactive tendency is stronger.

Reactive approach and avoidance tendencies can be relatively balanced, or one can dominate over the other. What is called

![Figure 1. Three temperamental influences on behavior. A: A reactive system for approaching rewards and a reactive system for avoiding threats or punishment compete for ascendance; in the absence of effortful control, the resultant of that competition is expressed in behavior. B: The engagement of an effortful control system permits the resultant arising from the competition of the reactive systems to be overridden, thus dampening the role of the reactive systems in determining behavior. Adapted from various statements by Rothbart, Eisenberg, and others.](image-url)
reactive undercontrol (Eisenberg et al., 2004) refers to cases in which the automatic system for approach dominates the automatic system for inhibition or withdrawal; what is called reactive overcontrol refers to cases in which the latter dominates (Nigg, 2001, 2003, termed this motivational inhibition).

Thus, in the absence of executive override (effortful control), a person with a sensitive reactive approach system and an insensitive reactive avoidance system will display impulsive pursuit of incentives. A person with a sensitive reactive avoidance system and an insensitive reactive approach system will display reflexive freezing or avoidance. Reactive undercontrol relates to impulsiveness and externalizing disorders; reactive overcontrol relates to behavioral inhibition and internalizing disorders (Rothbart, Ellis, & Posner, 2004; Valiente et al., 2003).

The development of the capacity for effortful control (which depends on emergent executive functions) changes this situation, because effortful control can counter whatever is the resultant reactive tendency (effortful control is superordinate to reactive control). If enough effortful control capacity is available, the impulsive grabbing of incentives that arises from a sensitive approach system can be restrained (Kochanska & Knaack, 2003; Murray & Kochanska, 2002). This child (or adult) can delay gratification. More generally, impulsivity is dampened.

However, effortful control is, in principle, not entirely a matter of restraining approach (Figure 2). Effortful control sometimes means forcing the production of an action that one does not want to take (overriding a reflexive tendency toward inaction). For example, effortful control can lead a sedentary adult to exercise, or can lead a person to stay engaged in a boring task (Sansone, Wiebe, & Morgan, 1999). Effortful control can lead children to look happy when they receive gifts they do not like (Kieras, Tobin, Graziano, & Rothbart, 2005). Effortful control may be required for a spendthrift to inhibit the tendency to buy things, but it may also be required for a tightwad to overcome the tendency not to buy things (Rick, Cryder, & Loewenstein, 2008).

Thus, effortful control can have diverging effects, depending on the context. Effortful control can overcome relatively reflexive tendencies triggered by reactive systems. However, exerting effortful control can move a person either toward restraint or toward action, depending on what reflexive response is being overcome.

Two-Mode Models and Brain Functions

Dual-process models in the literatures just discussed are based on behavioral research. However, arguments have also been made on largely neurobiological grounds that effortful and automatic modes of functioning are managed by different brain areas (Bechara, 2005; Rolls, 2005; see also Corbetta & Shulman, 2002; Daw et al., 2005). The automatic or reflexive system is said to include the ventromedial prefrontal cortex (PFC), the lateral temporal cortex, and subcortical regions such as the amygdala and the basal ganglia (Casey, Tottenham, & Fossella, 2002; Posner & DiGrolamo, 2000; Satpute & Lieberman, 2006). Controlled, effortful processing recruits regions that include rostral and dorsal areas of the anterior cingulate cortex (ACC) and the lateral and dorsal PFC.

It should be noted that assignment of brain areas to reflexive and reflexive systems is not fully consistent across studies and models, and the picture continues to evolve as more data come in. Thus, the linking of specific brain regions to the two modes is by no means complete. It should also be noted that subcomponents of larger structures likely serve different systems. For example, the orbitofrontal cortex (OFC) has been found to be activated both in (reflective) responses to stimuli evoking a sad mood and during (effortful) attempts to repair that sad mood (Cooney, Joormann, Atlas, Eugene, & Gotlib, 2007). Thus, some regions cannot be meaningfully assigned to one mode or the other. Nonetheless, there is evidence that does tend to support the two-mode idea. The sections that follow review this evidence; later comes discussion of how this view of the function of these regions helps interpret literature on serotonin and depression.

Automatic, Reflexive Mode

The reflexive mode includes the basal ganglia, which play a key role in automatic aspects of affect, cognition, and behavior. The putamen and caudate nuclei, for example, are involved in implicit learning and motor skill acquisition, responses that are learned slowly but executed quickly. The ventral striatum (a circuit consisting of the nucleus accumbens, parts of the putamen, globus pallidus, and caudate nucleus) is important in reward processing, contributing to the motivational salience of stimuli. Activity in this subcortical area increases in anticipation of reward (Ernst et al., 2004; Knutson, Fong, Adams, Varner, & Hommer, 2001). Craving and compulsive drug seeking are tied to activity in this area (Hyman & Malenka, 2001), and reactivity in this area relates to preference for immediate over delayed rewards (Hariri, Brown, et al., 2006).

The amygdala appears to play a key role in assigning emotional value to stimuli, both aversive and appetitive (Cunningham, Van Bavel, & Johnsen, 2008), although most available evidence concerns aversive stimuli (Davis & Whalen, 2001). For example,

---

1 In years past, it would have been common to refer globally to the reflexive mode of functioning as depending on cortical or prefrontal functions and the reflexive mode as depending more on subcortical functions (indeed, we have sometimes done so). It has become increasingly apparent, however, that that dichotomy is not tenable.
viewing faces expressing fear activates the amygdala (Whalen et al., 1998). The amygdala also is activated by novelty: for instance, an unfamiliar face (with no emotion) versus a familiar one (Schwartz, Wright, Shin, Kagan, Whalen, et al., 2003). The amygdala is also reactive to positive stimuli and stimuli that draw interest for reasons other than emotional content (Davis & Whalen, 2001). Reactivity of the amygdala appears to be a fairly stable trait; children characterized as inhibited at 2 years of age show hyper-reactivity of the amygdala to novel stimuli as adults (Schwartz, Wright, Shin, Kagan, & Rauch, 2003).

In light of such evidence, the amygdala is widely viewed as flagging stimuli quickly for emotional relevance (Aggleton, 2000; Anderson & Phelps, 2001; Cunningham et al., 2008; see also Kagan, 2007), thereby rendering the stimulus salient (Spoont, 1992). In this way, the amygdala plays an important role both in conditioning emotional reactions and in early stages of producing emotional reactions (Cahill, Babinsky, Markowitsch, & McGaugh, 1995; Cheng, Knight, Smith, & Helmstetter, 2006; LeDoux, 2000).

Besides emotion processing, subcortical areas are also involved in impulsive action. An example is impulsive aggression (Blair, 2003; Siever et al., 1999). For example, early positron emission tomography (PET) studies (Raine et al., 1994) showed high activity in the amygdala and other subcortical areas in persons who had committed impulsive murders.

Despite the fact that the PFC is linked in many people’s minds to higher-order processing and executive control, activation in some PFC regions has been consistently linked to automatic and affective processing (Ochsner & Gross, 2005). For example, some ventral and medial areas of the PFC with strong direct connections to subcortical structures including the amygdala and the basal ganglia are activated in response to emotion-eliciting stimuli. The OFC in particular, with strong reciprocal connections to the amygdala and other subcortical structures, has been linked to learning through stimulus-reinforcement associations and to responses to changing reinforcers (Kringelbach & Rolls, 2004). These PFC regions represent information pertaining to reinforcers and are used to form intuitions about stimulus–outcome probabilities prior to conscious awareness (Bechara, Damasio, & Damasio, 2000).

**Deliberative, Reflective Mode**

With respect to the effortful mode, regions of the dorsal PFC are activated when people engage in reflective top-down processing, as in reappraising emotional stimuli in order to regulate emotional responses to them (Ochsner, Bunge, Gross, & Gabrieli, 2002). Similarly, effortfully suppressing emotion activates the dorsolateral PFC. As an example, when men viewed erotic pictures, the amygdala and related areas were activated; when they were asked to inhibit their emotional reactions, however, dorsolateral PFC and ACC regions were activated (Beauregard, Lévesque, & Bourgouin, 2001), along with diminished activation of the amygdala. Low activation in the dorsal PFC has also been linked to prolonged activation of subcortical structures in response to an emotional challenge (Siegle, Steinhauer, Thase, Stenger, & Carter, 2002).

It is important to keep in mind that executive control does not always imply reduction of emotion. For example, one study had some participants dampen emotions and others increase them (Ochsner et al., 2004). As would be expected, what happened in the amygdala depended on the task. Efforts to reduce emotion reduced amygdala activity; efforts to increase emotion increased amygdala activity. Efforts to increase also uniquely activated left medial PFC; efforts to reduce uniquely activated regions of right lateral PFC and OFC. However, there were also areas in which both tasks yielded the same effects. Both efforts to increase and efforts to decrease emotion led to increased activity in left lateral PFC and ACC, and dorsal medial PFC.

It was noted earlier that some ventral PFC areas are involved in reflexive processes; there is also evidence, however, that some ventral PFC areas (parts of the OFC and ventral medial PFC) play a role in regulating reflexive processes (Drevets, 2000; Pizzagalli et al., 2004)—inhibiting the amygdala (Beauregard et al., 2001; Hariri et al., 2002, 2005) and ventral striatum (Nestler & Carlezon, 2006). For example, activation of some ventral PFC areas has been found during suppression or reappraisal of emotional stimuli (Lévesque et al., 2003; Ochsner et al., 2004) and when suppressing influences of emotional stimuli on behavior (Beer, Knight, & D’Esposito, 2006). Impulsivity has also been related to low activation in ventral PFC (Brown, Manuck, Flory, & Hariri, 2006). Thus, evidence links ventral PFC and OFC to both postulated systems (Cooney et al., 2007; Lévesque et al., 2003; Ochsner et al., 2002). Perhaps functional differentiation in the OFC links it to both modes (Zald & Kim, 1996); perhaps this region’s high degree of interconnectivity with other regions makes it a bridge between systems (Wallis, 2007).

The focus of this section thus far has been on control of emotions and impulses, but dorsal and lateral parts of the PFC actually are involved primarily in control of cognition (Horn, Dolan, Elliott, Deakin, & Woodruff, 2003; Jonides, Smith, Marshuetz, Koepp, & Reuter-Lorenz, 1998), inhibiting attention to distracting information (Chao & Knight, 1998) and inhibiting response tendencies when a task changes (Jonides et al., 1998). Thus, lateral and dorsal PFC regions are involved in cognitive processes that seem involved in the effortful regulation of emotion. However, these PFC areas are not themselves directly connected to subcortical areas (Clark, Cools, & Robbins, 2004; Lieberman, 2007; Schaefer et al., 2003), though they do connect closely to ventral and medial areas of the PFC, which have strong reciprocal connections to subcortical areas. It is possible that dorsal and lateral areas work through ventral and medial areas to affect amygdala activity (Lieberman, 2007; S. F. Taylor, Phan, Decker, & Liberzon, 2003).

Another cortical area that plays a key role in the regulation of subcortical regions is the ACC. A core function of the ACC is detection of conflict between responses cued by different aspects of available information (Bush, Luu, & Posner, 2000; Carter et al., 1998; Mitchell, 2006), but there is also functional differentiation within the ACC (e.g., Milham & Banich, 2005; Mohanty et al., 2007; van Veen & Carter, 2006). The dorsal ACC seems especially responsive to cognitive aspects of conflict, calling on dorsolateral PFC areas to exert top-down attentional control, to maintain focus on the task at hand by inhibiting processing of task-irrelevant information (Paus, 2001; van Veen & Carter, 2006). The dorsal ACC may be further differentiated, as well, with anterior areas involved in early-stage response evaluation and posterior areas involved in later-stage processes involving the response itself (Milham & Banich, 2005). This leads some to see this area as being involved in cognitive–motor mechanisms (Mohanty et al., 2007).
2007). Overall, however, the function of these areas is to minimize distraction and keep the intended behavior on track.

In contrast to these cognitive influences, rostral and ventral areas of the ACC represent its “affective” subdivision. They connect extensively with the OFC and with subcortical systems including amygdala and nucleus accumbens. Pizzagalli et al. (2001) argued that a main function of this subdivision of the ACC is assessing conflict between what the organism is currently doing and incoming information about potentially rewarding and emotional stimuli. Indeed, several studies have shown activation of this area when effortful emotion regulation is required (Bush et al., 2000; Ochsner & Barrett, 2001).

The affective subdivision may integrate salient affective and cognitive information and may change attentional processes in the cognitive subdivision (e.g., Pizzagalli et al., 2001). Indeed, pathways related to attentional processing and amygdala pathways converge within the rostral ACC. It thus has been argued that the rostral ACC may be especially involved in cognitive control in the context of emotional stimuli (Mohanty et al., 2007).

**Interplay Between Reflexive and Reflective Systems**

Dual-process models hold that both systems are at work and potentially competing with each other. Certainly some of that competition was implied in the preceding section. However, the interplay between the systems has been related more explicitly to a variety of outcomes, including impulsivity and emotional processing. With regard to impulsivity, impulsivity on tasks such as the go/no-go task has been linked to greater activation in subcortical areas relevant to emotion identification (e.g., amygdala, right insula), coupled with lower activation in ventral PFC areas (Brown et al., 2006; Horn et al., 2003). Thus, when areas of the reflective system are not inhibiting areas of the reflexive system that generate prepotent responses to cues, the person is more impulsive.

There is also evidence linking this impulsivity to emotional processing per se. Brown et al. (2006) found that higher scores on the Barratt Impulsivity Scale predicted more activity in the amygdala and less activity in ventral PFC during viewing of emotional faces (a task that engages amygdala activity). This is consistent with the idea that impulsivity relates more broadly to a general vigilance for immediate emotional relevance, coupled with lower effortful control.

Another study showing the interplay of these systems in regulating emotion and impulsive behavior (Hariri, Brown, et al., 2006) determined behavioral preferences for delayed versus immediate reward in one session, then related preference to activation of brain areas in another session. Initial preference for immediate reward (suggesting relative dominance of the reflexive system) predicted greater ventral striatum responses to both success and failure feedback on the later task (compared with no feedback) and greater activation of medial PFC (after success), indicating relative lack of executive control. In contrast, initial preference for delayed reward (suggesting relative dominance of the deliberative system) predicted greater activation of lateral OFC and dorsolateral PFC after success feedback, indicating relative presence of executive control. Thus, activity in prefrontal areas highly connected with the ventral striatum—the medial PFC and the OFC—related to preference for immediate rewards; activity in dorsolateral PFC (which may be involved in constraining the other areas) related to preference for larger, delayed rewards.

These findings are also consistent with those from a study by McClure et al. (2004) who reported that ventral striatum activity was significantly higher as participants chose smaller, immediate rewards compared with larger, delayed rewards. Activation of medial PFC and OFC also correlated positively with the selection of immediate rewards. In contrast, activation of dorsolateral PFC correlated with selection of larger, delayed rewards.

Consistent with this central role of certain prefrontal areas in delaying reward and controlling impulses, adults with PFC damage have been characterized as resembling children in terms of emotional reactivity and difficulties in regulating emotion (Fox, Henderson, Rubin, Calkins, & Schmidt, 2001). Galvan et al. (2006) recently reported that adolescents show greater activation in ventral striatum and more diffuse recruitment of the OFC in a reward task. They suggested that the fact that subcortical areas develop earlier than dorsal areas of PFC might account for impulsive and risk taking behavior in this age group (see also Steinberg, 2007).

As another example of the interplay between systems, functional connectivity (by within-subject correlations) has been seen between amygdala and dorsolateral, dorsal medial, ACC, and OFC during affect regulation by cognitive reappraisal (Banks, Eddy, Angstadt, Nathan, & Phan, 2007). Indeed, the degree of within-subject coupling between the amygdala and the OFC and dorsal medial cortical areas predicted the degree of affect reduction that occurred during the task.

**Section Summary**

The overall pattern of effects described in this section appears to be generally consistent with a two-mode picture of self-regulation. It fits the idea that certain brain areas (associated with the reflexive system) quickly identify emotional relevance and initiate emotional reactions, whereas specific areas of the prefrontal cortex (associated with the deliberative system) sometimes act to inhibit those reactions.

**Methods of Human Research on Serotonergic Function**

The two-mode viewpoint outlined in the preceding two sections is the perspective from which the rest of the article proceeds. Specifically, the sections that follow briefly review other literatures and interpret them from the two-mode framework. Those literatures also share one more feature: Each of them examines the involvement of serotonergic function in some type of behavior. All those sections raise the possibility that serotonergic function influences the balance between the two processing modes. Before turning to those sections, however, some comment is in order about research methods bearing on serotonergic function in humans (for greater detail, see Manuck, Kaplan, & Lotrich, 2006).

The processes by which serotonin operates in the nervous system are complex and not fully understood (Hensler, 2006; Lesch & Canli, 2006). It is clear, however, that more is involved than levels of available serotonin (e.g., Neumeister et al., 2006), including such factors as sensitivity and density of receptors, efficiency of metabolizing neurotransmitters released into the synaptic cleft, efficiency of re-uptake of neurotransmitter from the cleft, dietary
intake of amino acids, and recent history of the cell’s firing (thereby depleting available stores).

Thus it can be misleading to think of the serotonergic system in terms of level of serotonin per se, though there are contexts in which that view is appropriate. That is, temporary perturbations can change how much serotonin is available and thus have transient effects on behavior. In any long-term view of the serotonergic system, however (such as a person’s chronic experience over an extended period of time, or the tuning of the system by the prolonged use of selective serotonin reuptake inhibitors [SSRIs]), the issue really is the overall functioning of the system, rather than the levels of serotonin per se.

In light of such complexities, researchers have turned to several indirect ways to examine the role of serotonergic function in a given behavior or experience. Some of these methods focus on temporary perturbations of functioning, some on longer-standing patterns.

**Acute Tryptophan Depletion or Enhancement**

One relatively common and direct procedure temporarily lowers levels of serotonin experimentally, by artificially depleting the person of tryptophan, a precursor to serotonin. Tryptophan can be depleted with a drink (or capsules) containing high levels of 15 amino acids but no tryptophan. A control condition is a drink that has no effect on the balance of tryptophan to other amino acids. The manipulation results in an episode of acute tryptophan depletion with its lowest point several hours later, which in turn temporarily reduces available serotonin. The question of interest in this paradigm is the nature and extent of the response (behavioral, affective, or cognitive) to the lowering of available serotonin. Less frequently, manipulations have been used that work in the opposite direction, enhancing tryptophan levels and thus serotonin availability.

**Drug Challenge**

Another approach is to assess serotonergic function indirectly, by means of a drug challenge to the serotonergic system (Coccaro & Kavoussi, 1994; Power & Cowen, 1992). A challenge test involves administering precursors, agonists, or releasers of serotonin and assessing subsequent change in a hormone that is sensitive to change in serotonin. In challenge studies of serotonin and aggression, the most commonly used drug in the existing literature is fenfluramine, a serotonin releaser (Lopez Ibor, Saiz Ruiz, & Iglesias, 1988), and the most commonly measured outcome is prolactin response 3–4 hr later (Manuck et al., 2006). Serotonergic function is inferred from the prolactin response (Cowen, 2002; Flory, Mann, Manuck, & Muldoon, 1998; Hennig, Toll, Schonlau, Rohrmann, & Netter, 2000). Blunted prolactin responses are taken to indicate lower serotonergic functioning. The level of prolactin response then is related to the target behavior or psychological characteristic.2

One further issue is important in this particular literature (Manuck et al., 2006): Women’s prolactin response to serotonergic agonists can vary as much as threefold as a function of phase of the menstrual cycle (O’Keane, O’Hanlon, Webb, & Dman, 1991). Unfortunately, this is often not considered in studies using this method. This is a probable reason for inconsistent results among women when researchers use this methodology.

**Variation in Genotype**

Another methodological strategy is to relate the behavior of interest to genetic polymorphisms that have independently been associated with serotonergic function (Manuck et al., 2006). Though a large number of genes relate to serotonergic function (Martinowich & Lu, 2007), most of the genetic research discussed here examined the gene that codes the serotonin transporter. The transporter plays a key role in serotonergic transmission by facilitating reuptake of serotonin from the synaptic cleft (Heils et al., 1996). Indeed, under normal circumstances, this is the principal mechanism for actively clearing serotonin from the synaptic cleft.

Transcriptional activity of the transporter gene is believed to be influenced by (or at least associated with) a repetitive sequence in a polymorphic region called 5-HTTLPR, which has a short version and a long version (i.e., which has more repetitions). A variety of indirect evidence links this polymorphism to serotonergic function, though it is important to note that the evidence is not completely uniform (disconfirming findings have been reported, e.g., by Mann et al., 2000, and van Dyck et al., 2004).

Supportive evidence includes the following: The long version is associated in vitro with higher transcription efficiency than that of the short one. In human lymphoblast cells, for example, the long–long combination was linked to serotonin uptake 1.9 to 2.2 times that of the short–short or long–short combination, which had comparable serotonin uptake (Lesch et al., 1996). The alleles have also shown similar ratios in binding levels, in imaging measures of radioligand binding to 5-HT in vivo (Heinz et al., 2000) and in postmortem calculations of 5-HTT density (Little et al., 1998). Neuroimaging studies in humans and nonhuman primates have found an inverse relation between 5-HTT availability and cerebrospinal fluid concentrations of a 5-HT metabolite (Heinz et al., 2002), suggesting that the 5-HTTLPR is functional and influences serotonergic neurotransmission. One recent study found carriers of the short allele had twice the turnover of brain serotonin as those without the short allele (Barton et al., 2008). Of particular importance, in some of the literatures to be discussed, the 5-HTTLPR polymorphism has been associated with outcomes that resemble those obtained through direct manipulation of the serotonergic system. Thus, although there remain questions about causal mechanisms, the short allele is widely viewed as a marker of low serotonergic function (e.g., Canli & Lesch, 2007).

Two variants of the long allele have been identified relatively recently (Hu, Zhu, Lipsky, & Goldman, 2004). The expression of one variant is comparable with that of the short allele, and the two long variants behave differently in response to pharmacological challenge (Hu et al., 2005; Neumeister et al., 2006). Such findings encourage closer attention to variability in this polymorphism (this polymorphic region may have as many as 10 variants in humans; **2**

---

1 Fenfluramine is not the only possible drug challenge, and because of its potential toxicity with long-term or repeated usage, newer studies are turning to other challenges. Other drugs used include m-CPP, buspirone, and ipsapirone. Unfortunately, m-CPP and buspirone influence dopamine function as well as serotonin, and ipsapirone and m-CPP both bind to adrenergic receptors (Yatham & Steiner, 1993). For these reasons, even though findings from these challenges are generally comparable with those from fenfluramine (Manuck et al., 2006), consideration of challenge studies here is limited almost entirely to fenfluramine studies.
Nakamura, Ueno, Sano, & Tanabe, 2000). Further, polymorphisms at other sites on this gene also appear to have functional effects on transporter expression (Kraft et al., 2007). More generally, it is important to acknowledge that the genetic association studies we describe here evaluate correlations between a phenotype and the single marker that is under study, rather than a causal link.

Nonetheless, most published work on the serotonin transporter is limited to comparisons on 5-HTTLPR. Most published work is further limited to comparisons that involve the long and short allele without further differentiation. Thus, in this article, mention of polymorphisms in the serotonin transporter gene always refers specifically to 5-HTTLPR and primarily to short- versus long-allele carriers. A few studies that do go beyond this are noted later on.

**Serotonergic Function and the Brain**

This section addresses the idea that serotonergic function plays an important role in differentiating the two modes of functioning introduced in the article’s first major section. This idea is advanced by relating serotonergic function to brain areas that were discussed in the second major section as subserving those two modes.

The central serotonergic system projects from the brainstem extensively throughout cortical and subcortical structures. Serotonergic projections originate primarily in the dorsal and median raphe nuclei. Neurons within these nuclei have different but overlapping projection patterns. Serotonergic synapses are densely concentrated in subcortical regions, including the amygdala, ventral striatum, and hypothalamus. Raphe neuron firing is regulated mostly by projections from brainstem nuclei, with some input from thalamus, hypothalamus, and limbic areas, but there is evidence that the ventral OFC has a direct inhibitory effect on raphe activity (Hensler, 2006; Ressler & Nemeroff, 2000).

The focus here is on areas of serotonergic innervation that are most directly relevant to the two-mode idea. Ascending serotonergic projections from the raphe nuclei have a widespread cortical distribution, with particularly dense serotonergic projections to the medial, ventral, and orbital regions of the PFC (Way, Láca, Fairbanks, & Melega, 2007). Tryptophan depletion has been found to decrease 5-HT2 receptor binding in the dorsolateral PFC (Yatham et al., 2001), suggesting that receptors in this region are particularly sensitive to serotonergic variation. Tryptophan depletion has also been found to decrease activation in the dorsolateral/medial PFC during a verbal working memory task (P. P. Allen et al., 2006). Tryptophan depletion has been shown to affect ACC activity (K. A. Smith, Morris, Friston, Cowen, & Dolan, 1999), particularly among persons with the short allele of the serotonin transporter gene (Neumeister et al., 2006). Thus, serotonin seems important in the functioning of cortical structures identified in the earlier section on brain function and two-model modes.

A previous section also noted that cortical and subcortical structures are interconnected, forming complex circuits that regulate emotion and behavior. Corticolimbic neural circuits that mediate emotion are densely innervated by serotonergic neurons and exhibit rich expression of serotonin receptors (Parsey et al., 2006). Indeed, serotonin appears critical for the development of this emotional circuitry, and even transient alterations during early development modify neural connections (Hariri & Holmes, 2006).

Within these circuits, serotonergic function is thought to be involved particularly in constraining excitatory influences on amygdala activity (Hariri & Holmes, 2006; Hariri & Weinberger, 2003). It has been argued that increased serotonin levels may thereby decrease the sensitivity of the amygdala to external stimuli (particularly aversive stimuli; Hariri et al., 2002). Support for this general idea has emerged from studies of the 5-HTTLPR polymorphism (Hariri et al., 2005), drug challenge studies, and neuroimaging studies (Clark et al., 2004).

In one of the first projects to combine genetic analyses and brain imaging, Hariri et al. (2002, 2005) found that carriers of the short allele of 5-HTTLPR had greater amygdala reactivity to emotional faces than did those with the long allele. Considerable additional evidence that short-allele carriers display amygdala hyperreactivity has accumulated since then (for a meta-analysis, see Munafò, Brown, & Hariri, 2008, who report that 5-HTTLPR may account for up to 10% of phenotypic variance in amygdala activation to a broad range of salient environmental stimuli). There is also evidence that the short allele of 5-HTTLPR relates to lower resting activity in the ventromedial PFC, an area that constrains amygdala activity (Rao et al., 2007).

Other data relate the 5-HTTLPR polymorphism to both structural and functional variations in this circuit as a whole. Pezawas et al. (2005) found smaller gray matter volume in a specific region of the PFC (the rostral ACC) and in the amygdala in carriers of the short allele compared with those without it. These authors also used functional connectivity analysis, a measure of correlation in activity between regions, derived from fMRI data. Amygdala and ACC activity were correlated overall; the subgenual rostral part of the ACC (BA 32/25/24) was positively correlated with amygdala activity, and the supragenual more caudal part of the ACC (BA 32) was negatively correlated with amygdala activity. These two regions of the ACC also showed strong positive connectivity with each other, suggesting that they may form a feedback loop with the amygdala.

These analyses revealed less coupling of these structures when viewing angry and fearful faces among carriers of the short allele.

---

3 It may be desirable to distinguish here between inhibition and constraint as concepts. Spoons (1992) argued that constraint means regulation of information flow so as to prevent being overwhelmed by it, rather than suppression of information flow. Her view of serotonergic function was that it constrains—regulates—communication rather than inhibiting it (see also Depue & Spoons, 1986). By limiting the flow of information, constraint serves an important homeostatic function, preventing overshoot of other elements within the overall system (e.g., excessive stress reactivity).

4 It was initially assumed that low serotonergic function relates to stronger amygdala reactions to emotion-relevant stimuli. Other evidence, however, casts a somewhat different light on the findings. Canli et al. (2005) found that the short allele predicted greater amygdala activity to negative stimuli than to neutral stimuli (replicating previous findings) but also predicted more amygdala activity to a fixation point than to neutral stimuli (replicated by Heinz et al., 2007). Canli et al. (2005, 2006) suggested that the short allele relates to greater tonic activation; that is, at rest the amygdala of this person is active, and that a nonemotional cognitive event (perceiving a neutral stimulus) reduces the activity (Canli & Lesch, 2007; Shulman et al., 1997), creating a difference between neutral stimuli and fixation. An accumulation of other evidence (Brocke et al., 2006; Canli et al., 2006) suggests that the amygdala in a person with the short allele does not just respond more intensely to emotional stimuli but is more vigilant—more active whenever there is uncertainty or novelty (cf. Davis & Whalen, 2001).
Carriers of the short allele therefore appear to show less inhibitory regulation of the amygdala than do those with the long allele (see also Heinz et al., 2005). It is noteworthy that the connectivity difference between groups was most noticeable in a PFC area that has the highest density of serotonin transporter terminals in the human cortex and is the target of dense projections from the amygdala.

The observations of amygdala hyperreactivity in carriers of the short allele in other studies may reflect the decoupling of this circuit. If activity of this prefrontal region inhibits activity in the amygdala, less connectivity in the carriers of the short allele would permit greater excitability of the amygdala and deficits in affect regulation (Hariri, Drabant, & Weinberger, 2006).

Pezawas et al. (2005) also asked their participants to complete a self-report measure of harm avoidance, as an indicator of emotionality. Measures of prefrontal activity and amygdala activity per se did not predict individual differences in harm avoidance. However, indices of the connectivity between these areas accounted for almost 30% of the variance in this trait. This suggests a link between decoupling of the amygdala–rostral-ACC feedback circuitry and elevation in dispositional emotionality. This circuit as a whole, then, appears to be affected by serotonergic function.5

Section Summary

As described earlier, specific brain regions and their interconnections appear to function in ways that resemble the functions of the two-mode behavioral and cognitive models. Evidence reviewed in this section appears to be consistent with the view that serotonergic function plays an important role in these brain areas and their interaction. The data suggest that low serotonin function (as manipulated by acute tryptophan depletion and assessed by serotonin transporter polymorphism) results in less constraint of amygdala activity.

Cognitive and Behavioral Effects of Serotonergic Function

The previous section examined how serotonergic function is involved in brain regions that were identified earlier with the two-mode model. What evidence links serotonergic function to psychological manifestations of the two processing modes? If serotonergic function is related to constraint of areas associated with the reflexive system (e.g., the amygdala), a number of other things should follow. Most simply, high serotonergic function should facilitate cognitive and behavioral inhibition of emotional states and impulses. This does appear to be so.

Serotonergic Function and Emotion-Related Processing

Much of the human research relevant to this idea has examined behavioral effects of serotonergic function experimentally, using acute tryptophan depletion to temporarily reduce serotonin. Tryptophan depletion has been shown to impair performance on cognitive tasks of executive processes such as reversal learning, memory, and attention tasks (Park et al., 1994; Riedel, Klaassen, Deutz, van Someren, & van Praag, 1999). Tryptophan depletion affects performance on behavioral and cognitive tasks that have emotional components. As an example, tryptophan depletion impairs the ability to inhibit responses to previously rewarded cues (Cools, Blackwell, et al., 2005; Park et al., 1994; Rogers et al., 2003). It is noteworthy that the strongest effect of tryptophan depletion found by Cools, Blackwell, et al. (2005) was among people who reported being highly impulsive. This pattern is consistent with the two-mode model.

A particular focus has been on how cognitive tasks that involve emotional stimuli are affected by tryptophan depletion. For example, in one study acute tryptophan depletion led female participants to show slower processing of happy words (but not sad ones) in an affective go/no-go task (F. C. Murphy, Smith, Cowen, Robbins, & Sahakian, 2002). This study also found impairment on the first round of a reversal task involving reward and punishment, but no impairment on a task without affective stimuli.

Tryptophan depletion generally does not affect mood per se among persons with no personal or family history of depression (Benkelfat, Ellenbogen, Dean, & Palmour, 1994; Park et al., 1994). However, the impairment in suppressing emotion-relevant stimuli suggests that tryptophan depletion might exaggerate mood change in response to a stressor. Consistent with that possibility, tryptophan depletion has led persons to report more sadness during exposure to uncontrollable stress (aversive noise), whereas the effect was only minor when the noise was controllable (Richell, Deakin, & Anderson, 2005).

Evidence for the role of serotonin in inhibition of responses to emotion-relevant stimuli has also emerged in one study of tryptophan supplements (which enhance serotonin availability). Women given tryptophan supplements for 14 days showed decreased recognition of disgusted faces and a diminished startle response to negative pictures, as well as an increase in the ability to recognize subtle facial expressions of happiness, compared with a control group (S. E. Murphy, Longhitano, Ayres, Cowen, & Harmer, 2006). The effect did not emerge for men.

Behavioral Effects of SSRIs in Nonclinical Samples

Researchers have also manipulated the serotonergic system by administration of SSRIs. The vast majority of such studies are of patients under treatment. However, in clinical samples an important question arises as to whether any effect observed reflects the change in serotonergic function or relates instead to other changes that follow from improved psychiatric status. Given such issues, effects of SSRIs in nonclinical samples are of particular interest. Several such studies have been published.

Knutson et al. (1998) administered 4 weeks of SSRI to a group of 26 nonclinical participants and compared them with a control group of 25. SSRI led to less negative affect (from the Positive and Negative Affect Schedule, PANAS; Watson, Clark, & Tellegen, 1988). There was no change in positive affect, and no tests were raised but not answered at this point is whether the effect of serotonergic function in this circuitry is better characterized as activating the deliberative system, which then constrains the reflexive system, or as deactivating the reflexive system, which then permits the deliberative system to play a larger role (cf. Gotlib & Hamilton, 2008). For present purposes, the end result is the same in either case.

5 This discussion of serotonergic function was framed in terms of circuitry that seems relevant to two-mode models. A question that can be raised but not answered at this point is whether the effect of serotonergic function in this circuitry is better characterized as activating the deliberative system, which then constrains the reflexive system, or as deactivating the reflexive system, which then permits the deliberative system to play a larger role (cf. Gotlib & Hamilton, 2008). For present purposes, the end result is the same in either case.
reported for specific affects. Those with SSRI also displayed more cooperative behavior in a group task than did control participants.

In another small study of nonclinical participants, Tse and Bond (2002) had roommates make ratings of the participants and measured several qualities of behavior during a mixed-motive game with another person. Those receiving SSRI ($n = 10$) were rated by roommates as being less submissive than control participants ($n = 10$), and they showed more dominant eye-contact behavior during the mixed-motive game than did control participants. However, they also displayed more cooperative, affiliative behavior during the game.

**Impulsiveness**

The idea that serotonin function plays a role in constraining emotion also has implications for behaviors involving impulsive responses to emotions. Evidence of behavioral impulsiveness comes from a number of sources (e.g., Crockett, Clark, Tabibnia, Lieberman, & Robbins, 2008; Walderhaug et al., 2002), but of particular importance is work bearing on impulsive responses to anger. That literature includes both studies of low-level aggression in the laboratory and studies of violence in naturalistic settings. Some of the former research is noted here; the latter studies are taken up in a later section.

An important conceptual point is made in a study of aggressiveness by Cleare and Bond (1995). Participants were pre-assessed as being either high or low in aggressive tendencies. Those high in aggressive tendencies became more aggressive, hostile, and quarrelsome after tryptophan depletion, and less so after tryptophan enhancement, but there was no effect for those low in aggressive tendencies. Similar results were reported by Finn, Young, Pihl, and Ervin (1998). This pattern suggests that effects of low serotonergic function on aggression are less about aggression per se and more about the release of existing habitual tendencies to be aggressive (see also Manuck et al., 2006; Spoont, 1992). This is consistent with the idea that serotonergic function promotes effortful control over underlying vulnerabilities.

A later study (Bjork, Dougherty, Moeller, & Swann, 2000) reinforced this point: Tryptophan depletion in this case led to greater aggressive response to provocation among men high in aggressiveness but had an opposite effect among those low in aggressiveness. Thus, the effect of serotonergic function on aggression appears to depend at least partly on dominant preexisting tendencies.

**Section Summary**

The studies described in this section appear to be relatively consistent in suggesting that experimentally increasing serotonergic function reduces responsiveness to negative emotional stimuli, decreases aggression, and increases cooperativeness; there also were increases in what might be thought of as social potency (see also Spoont, 1992; Young & Leyton, 2002). Reducing serotonergic function impaired performances on tasks with emotional elements; it also increased expression of aggression and deterioration of cooperativeness, though these effects often occurred only among persons who were relatively high in aggressive tendencies by disposition.

Serotonergic Function, Impulse-Related Clinical Disorders, and Violence

Evidence linking serotonergic function to lack of executive control, impulsiveness, and aggression in nonclinical samples suggests that low serotonergic function should also relate to disorders that are characterized by these qualities. There are in fact intriguing parallels between phenomena described in the preceding section and certain phenomena that are more clinically relevant. This section considers some of those parallels, though it is not an attempt to review externalizing disorders in general (for broader treatment see, e.g., Krueger et al., 2002). Specifically, evidence is briefly considered here regarding serotonergic involvement in attention-deficit/hyperactivity disorder (ADHD) and externalizing behaviors such as conduct disorder, disruptive disorders, and antisocial personality disorder.6

**ADHD**

There is a substantial basis for believing that ADHD is actually a disorder of inability to control behavioral impulses (Barkley, 1997; Nigg, 2001, 2003; Quay, 1988). Symptoms include distractibility, difficulty following instructions, disorganization, and talking out of turn. Evidence of actual deficits in attention has generally been lacking (Barkley, 1997). Barkley (1997) argued that this disorder rests on deficits in executive functions and consequent deficits in motor control (for evidence that problems labeled attention problems reflect primarily difficulties with response inhibition, even in unselected children, see Friedman et al., 2007).

A more recent review led Nigg (2001, 2003) to a similar conclusion. Indeed, Nigg explicitly utilized a two-mode approach in analyzing ADHD. Inhibition by anxiety, uncertainty, or fear he termed motivational inhibition. What he termed executive inhibition resembles the concept of effortful control (though he applied it only to inhibitory tendencies). His review led him to conclude that ADHD reflects executive deficits. Consistent with this, an association has been found between the 5-HTTLPR polymorphism and ADHD (Beitchman et al., 2003).

**Conduct Disorder, Antisocial Personality Disorder, Violence**

A lack of executive control has also been linked to conduct disorder and other displays of antisocial tendencies. Dennis and Brotman (2003) found that a measure of effortful control predicted lower levels of (mother-reported) aggression in a high risk sample (younger siblings of youths adjudicated through the family court system). Other evidence ties weak effortful control to externalizing disorders as well (Rothbart et al., 2004; Valiente et al., 2003).

There is also evidence linking this sort of behavioral problem to serotonergic function. Within an ADHD group, one study related aggression to low serotonergic function assessed by fenfluramine challenge (Halperin et al., 1994). A more recent study of children 6 Apart from space considerations, studies of serotonergic function in substance-related disorders are not addressed because of the clear evidence that exposure to alcohol and drugs itself influences the function of neurotransmitter systems and the complexity of evaluating pre- and postdiagnostic data within humans (Koob & Le Moal, 2008).
with ADHD found a prospective link from low serotonergic function (again by fenfluramine challenge) to antisocial personality disorder 9 years later (Flory, Newcorn, Miller, Harty, & Halperin, 2007). Another developmental study tested boys with disruptive behavior disorders (ages 7–11 years) with fenfluramine challenge and reevaluated the boys’ clinical status nearly 7 years later (Halperin et al., 2006). Adjusting for baseline aggression, low serotonin function prospectively predicted higher adolescent aggression and antisocial behavior. Regardless of level of aggression in childhood, no child who had displayed high serotonin function in childhood was particularly aggressive in adolescence, suggesting that high serotonin function may have a protective role.

A good deal of research has also examined serotonergic function assessed by drug challenge in adults with clinical conditions reflecting impulsive aggression (for a more extensive review, see Manuck et al., 2006). O’Keane et al. (1992) found that convicted murderers with antisocial personality disorder had a blunted response to fenfluramine, compared with that of control participants. Lower serotonergic function has long been linked to history of fighting and assault (Coccaro, Kavoussi, Cooper, & Hauger, 1997), domestic violence (George et al., 2001), and impulsive aggression more generally, particularly among men (Cleare & Bond, 1997; Coccaro, Kavoussi, Hauger, Cooper, & Ferris, 1998). Consistent with the idea that impulsivity is a key element in these disorders, a serotonin releaser was found in one study to cause lower impulsivity on a delayed reward task among persons with conduct disorders (Cherek & Lane, 2000).

Genetic evidence also connects serotonergic function to violent and antisocial behavior. One recent study of young adults, some with personality disorders and others without, related the short allele of the 5-HTTLPR polymorphism to traits pertaining to borderline personality disorder and antisocial personality disorder (Lyons-Ruth et al., 2007), though another study (Sakai et al., 2007) found no link from this polymorphism to adolescent conduct problems. A sample of men in China convicted of extremely violent crimes was more likely to include short-allele carriers of the 5-HTTLPR polymorphism than was a control group (Liao, Hong, Shih, & Tsai, 2004), though the polymorphism did not relate to antisocial personality disorder within the criminal group. Another study (men only) found that the combination of the 5-HTTLPR polymorphism short allele and adverse childhood environment predicted violent adult behavior (Reif et al., 2007).

Other studies also make more differentiated points regarding serotonin and impulsive aggression. M. C. Dolan, Anderson, and Deakin (2001) linked low serotonergic function to higher impulsivity and higher aggression in male aggressive offenders. Interestingly, both impulsivity and aggression also related to higher anxiety in this sample, arguing against a path in which impulsive aggression is a product of low fear. This pattern thus fits Nigg’s (2001) description of lack of executive inhibition rather than lack of motivational inhibition.

M. C. Dolan, Deakin, Roberts, and Anderson (2002) examined men with aggressive personality disorders (from a maximum security psychiatric hospital) and a control sample (from hospital staff) on some of the same outcomes and some different ones. Men diagnosed with psychopathy (the majority of the aggressive group) had poorer executive function and greater self-report impulsivity than did the others. The psychopathic subset of aggressors had lower serotonergic response than did the nonpsychopathic subset, but it did not differ from that of the control group (the authors noted, however, that the control group was likely not representative of the broader community). Impulsivity, but not aggression, correlated with serotonergic function.

M. C. Dolan and Anderson (2003) examined criminal offenders with diagnoses of personality disorders, relating response to fenfluramine challenge with scores on the Psychopathy Checklist. Lower serotonergic function did not predict overall psychopathy scores but did relate to scores on the impulsive–antisocial component, which combines ratings of irresponsibility, adolescent antisocial behavior, adult antisocial behavior, impulsiveness, and poor behavioral control.

There is also research linking impulsive aggression to brain activity that occurs in response to serotonin. New et al. (2002) tested control participants and persons with a history of impulsive aggression by PET scan under a serotonin releaser and under placebo. Scans revealed differences between groups in drug response. Patients showed less activation (vs. placebo) than did control participants in left medial posterior OFC, leading New et al. (2002) to conclude that impulsive aggression by the patients reflects lack of activity in areas involved in planning and long-term regulation of behavior. Consistent with these findings, the OFC has also been implicated in the control of aggressive impulses in other research (Dutton, 2002).

Raine and Yang (2006) reviewed a variety of imaging research relating to antisocial behavior. They concluded that the evidence indicates impairments of a number of brain areas in antisocial populations, including both dorsolateral and orbitofrontal regions of the PFC, amygdala, and anterior cingulate. This pattern suggests problems in the circuitry constraining emotion, consistent with a two-mode view.

Suicidality

Suicidal behavior has long been linked to serotonin, via serotonin metabolites in cerebrospinal fluid (e.g., Ågren, 1980). Over 1,000 studies have been conducted on personality traits and suicidality (Brezo, Paris, & Turecki, 2006). Although negative affective traits such as neuroticism and hopelessness are involved in development of suicidal ideation, a large body of research suggests that suicidal acts and completed suicides are reliably tied to impulsivity, particularly impulsive aggression (for a detailed review, see Brezo et al., 2006). Mann, Waternaux, Haas, and Malone (1999) suggested that impulsiveness in particular is a risk factor for suicide, rendering a person who has suicidal feelings more likely to act on them.

Cross-sectional studies suggest that lower cerebral-spinal fluid 5-HIAA (a serotonin metabolite) is related to greater lethality of suicide attempts (Placidi et al., 2001). Although prospective research is limited, several studies of high risk samples have now found that lower cerebrospinal levels of serotonin among persons hospitalized for an initial suicide attempt predict prospectively a three- to fourfold increase in risk of subsequent death by suicide (for a review, see Mann & Currier, 2007).

7 Similar conclusions come from studies of borderline personality disorder, where serotonergic function has been related to impulsiveness and emotional reactivity (see Carver & Miller, 2006; Looper & Paris, 2000).
Section Summary

A good deal of evidence links low serotonergic function, assessed in various ways, to disorders and other kinds of problem behaviors that are characterized by deficient impulse control. Those problems often result in disruptive behavior, sometimes extending to impulsive violence, and sometimes extending to suicide. This pattern is consistent with the two-mode model, and with the idea that effortful control is necessary to restrain behaviors of this sort.

Personality Traits and Serotonergic Function

Much research has also examined relationships of serotonergic function to personality as measured by self-report, using several different procedures to assess serotonergic function. Some of this work has focused on qualities of personality pertaining to aggression and impulsiveness; other studies have examined a broader spectrum of personality qualities.

Trait Hostility and Impulsiveness

As noted in the preceding section, an extensive literature links serotonergic function to impulsive aggression. Hostility as a trait has also been related to low serotonergic function in nonclinical samples (Cleare & Bond, 1997; Depue, 1995; Netter, Hennig, & Rohrmann, 1999). Depue (1995) reported a more extensive array of associations, relating low serotonergic function (assessed by fenfluramine challenge) as well to the Control–Impulsivity facet scale from the Constraint factor of the Multidimensional Personality Questionnaire (Tellegen, 1985), the Aggression facet of the Multidimensional Personality Questionnaire’s Negative Emotionality factor (but not other facets), two sensation seeking subscales, and several indices of impulsiveness. Depue (1995) reported that low serotonergic function related most strongly to hostility subscales that he characterized as reflecting impulsive, action-oriented aggression (irritability and assertiveness), as opposed to more passive, cognitive forms (resentment and suspiciousness).

A more recent study (Hennig, Reuter, Netter, Burk, & Landl, 2005) yielded results that resemble these conceptually, by using both a drug challenge and a genetic method. The drug challenge was a single dose of SSRI, with the outcome being change in cortisol as an indirect indicator of serotonin function. High responders to the SSRI scored high on what Hennig et al. (2005) called aggressive hostility (marked by assertiveness and negativity). What they called neurotic hostility (marked by irritability and resentment, but also guilt) did not relate to cortisol response. Hennig et al. (2005) also assessed the tryptophan hydroxylase gene, which expresses an enzyme that helps synthesize serotonin. Variations in that gene related to aggressive but not neurotic hostility. Hennig et al. (2005) noted that aggressive hostility was not related to a separate measure of neuroticism except for the neuroticism facet concerning aggression. They concluded that aggressive hostility resembles psychoticism (defined by impulsivity and lack of social connection), not neuroticism.8

Personality and Serotonergic Function Assessed by Drug Challenge

Studies have also related serotonergic function as reflected in responses to drug challenge to personality in community samples, assessing personality more broadly. Manuck et al. (1998) assessed prolactin responses to fenfluramine and related them to the Neuroticism–Extroversion–Openness Personality Inventory (NEO–PI), plus additional measures. Among men, blunted prolactin response (suggesting low serotonergic function) related to higher neuroticism (assessed by NEO–PI) and the Neuroticism facet Angry Hostility (other facets were not tested); blunted prolactin response also related to low conscientiousness (assessed by NEO–PI) and to separate measures of impulsiveness and life history of aggression. No effects emerged in the full sample of women (perhaps because of variation in menstrual cycle), but analyses limited to postmenopausal women yielded effects for conscientiousness and impulsiveness. The association with neuroticism was replicated in a larger sample by the same research group (Flory, Manuck, Matthews, & Muldoon, 2004), but in this case no tests were reported for conscientiousness or any facet of neuroticism.9

Several studies have related serotonergic function assessed by drug challenge to the Tridimensional Personality Questionnaire (TPQ; Cloninger, 1987), with mixed results. Two studies found links between low levels of harm avoidance and lower serotonergic function (Gerra et al., 2000; Hanssen & Ansseau, 1999). One found the opposite (Hennig et al., 2000). J. Evans, Platts, Lightman, and Nutt (2000) found no link from serotonergic function to any TPQ scale. Apart from this inconsistency in results, there are measurement issues concerning the TPQ that urge caution in interpreting findings relating to it (Carver & Miller, 2006; B. P. O’Connor, 2002).

SSRIs and Personality

A previous section described behavioral results from a slim literature on SSRI effects in nonclinical samples. Two studies from that literature also examined effects on personality self-reports. Knutson et al. (1998) found that 4 weeks of SSRI led to lower self-report scores on the Assaultiveness and Irritability scales of the Buss–Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957). It is of interest that these are the same two subscales as Depue (1995) had earlier found to be related to low serotonergic function as assessed by drug challenge.

8 A good deal of the evidence links low serotonergic function to traits related to negative emotions. There is far less evidence concerning serotonergic function and positive affects, and the evidence that does exist is mixed. Zald and Depue (2001) found an inverse relation between prolactin response to fenfluramine and ambient positive (and negative) feelings; Flory et al. (2004) found a positive relation between prolactin response to fenfluramine and positive feelings.

9 Just as neuroticism has facets, so does conscientiousness, and there is reason to believe that some facets relate to serotonergic function but not others. This would compromise research on conscientiousness that did not test facets. Roberts, Chernyshenko, Stark, and Goldberg (2005) found six factors in multiple measures of conscientiousness, only moderately intercorrelated, which differed substantially in their relations to criterion measures. Indeed, in other research (Moon, 2001) conscientiousness was unrelated to a study outcome, but two facets had substantial and opposite relationships with the outcome. Just as there is merit in examining how facets of neuroticism relate to markers of serotonergic function, it also seems important to test facets of conscientiousness separately.
Another small study of nonclinical participants (Tse & Bond, 2001) examined effects of SSRI on an abbreviated version of the Temperament and Character Inventory (Cloninger, Svrakic, & Pryzbeck, 1993). SSRI led to an increase on a measure of self-direction (reflecting such qualities as responsibility, purposefulness, and resourcefulness) but produced no other difference.

**Personality and the Serotonin Transporter Polymorphism**

There is also a substantial literature on the 5-HTTLPR polymorphism and personality as assessed by broad-ranging self-report inventories, which permitted investigation of diverse possible associations. This work began with several large-scale studies in which the data were examined thoroughly. Lesch et al. (1996) studied a sample of 505 normal participants, mostly men. They found that the short allele (identified with low serotonergic function) related positively to neuroticism and inversely to agreeableness (both assessed by NEO–PI–Revised [NEO–PI–R]). Not all facets of neuroticism related to the polymorphism, however. In analysis of Neuroticism facet scales, the polymorphism related most strongly to Angry Hostility, Depression, and Impulsiveness scales; a link with the Anxiety scale was barely significant and there was no link to the Vulnerability scale.

The Impulsiveness facet in neuroticism requires further clarification. Its inclusion in that trait domain reflects the fact that negative feelings, themselves the hallmark of neuroticism, sometimes lead to impulsive actions such as overindulgence (which is well represented in the items). There is some question as to whether impulsivity as assessed by this scale is a good representative of impulsivity as a construct.

A second major study examined 397 participants, mostly women (Greenberg et al., 2000). Again, the 5-HTTLPR polymorphism related to both neuroticism and agreeableness, with an additional weak but significant association for conscientiousness. Analysis of Neuroticism facets revealed the strongest relations again for Angry Hostility and Depression scales. Associations with Anxiety, Impulsiveness, and Vulnerability scales were all nonsignificant.

A third large-scale study was reported by Sen, Villafuerte, et al. (2004). This sample of 419 participants yielded an association of the polymorphism with NEO–PI neuroticism and a weaker but significant association with agreeableness. As in the other studies, the facet of neuroticism most strongly related to the polymorphism was the Depression scale. The short allele related significantly to Anxiety and Self-Consciousness scores in this sample, to the Angry Hostility scores marginally (p = .052), and not to Impulsiveness or Vulnerability scores.

Many smaller-scale studies have been conducted, with varying outcomes. There have by now been several meta-analyses of the relationship of variation in 5-HTTLPR to neuroticism. In one meta-analysis (Sen, Burmeister, & Ghosh, 2004), with all studies combined there was no association. When the studies were separated by personality measure, 5-HTTLPR related to NEO–PI–R Neuroticism, but not Temperament and Character Inventory (TCI) or TPQ Harm Avoidance. Schinka, Busch, and Robichaux-Kenne (2004) reached a similar conclusion from their meta-analysis. Munafò, Clark, and Flint (2005a) found the opposite: a significant relation with TPQ/TCI Harm Avoidance, but not with NEO–PI Neuroticism. However, reanalysis of their data using the procedures used in the other meta-analyses reversed the pattern, with an association now emerging for NEO–PI Neuroticism (Munafò, Clark, & Flint, 2005b).

Unfortunately, these meta-analyses have two serious problems. First, none examined neuroticism facets. Because facet analyses were very informative in the earlier large-scale studies, this appears a serious omission. Second, traits other than neuroticism were disregarded. Thus, there was no test of a link from genetic difference to agreeableness, which had been found in all three large-scale studies (Greenberg et al., 2000; Lesch et al., 1996; Sen, Villafuerte, et al., 2004). The pattern presented by these three large-scale studies as a group is a picture relating low serotonergic function to disagreeable, dysphoric surliness.

Schmitz, Hennig, Kuepper, and Reuter (2007) took a more fine-grained approach to examining neuroticism in genetic research. They collected data on three measures of neuroticism and closely related traits (though again other traits were not considered), and tested the measures item by item against the polymorphism. Twelve items significantly related to genotype; item content focused on depression and stress reactivity. Consistent with that finding, the 5-HTTLPR polymorphism has been linked to differences in cortisol response to a standard laboratory stressor (Gotlib, Joormann, Minor, & Hallmayer, 2008).

**Section Summary**

As a group, the bulk of the findings reviewed in this section appear consistent with the idea that serotonergic function relates to a group of traits with agreeableness and constraint at one end and impulsivity, along with aggressive and depressive tendencies, at the other end (Spoont, 1992). The pattern appears consistent with the view that serotonergic pathways are involved in impulse control (Depue, 1995; Depue & Collins, 1999; Depue & Spoont, 1986; Manuck, Flory, Muldoon, & Ferrell, 2003; Soubrie, 1986; Spoont, 1992; Zuckerman, 2005), particularly impulses reflecting strong negative emotions. High serotonergic function appears to relate to consideration of the future consequences of one’s behavior (conscientiousness) and to positive social connection (agreeableness, cooperation, affiliation). This includes social potency (even dominance), but a positive social engagement rather than dominance by force (cf. Manuck et al., 2006). The literature reviewed here is quite different from that reviewed by Spoont (1992), but the conclusions from the two reviews are in many ways quite similar.10

**Interpretation in Terms of Two-Mode Models**

The preceding sections related serotonergic function to cognitive processing of emotional stimuli, overt behaviors such as impulsiveness and aggression, disorders involving poor impulse control, and a

---

10 An early view of serotonergic function related it specifically to the experience of anxiety (Handley, 1995). Indeed, it remains common to refer to “anxiety-related traits” in discussing this literature (e.g., George et al., 2001; Hennig et al., 2000; Lesch & Canli, 2006; Munafò et al., 2005a; Schinka, 2005; Schinka et al., 2004). The tendency to invoke anxiety doubtlessly reflects observed associations of serotonergic function with global measures of neuroticism. However, it seems clear from this review that anxiety is not the only emotion that must be considered here (see also Panksepp & Cox, 1986; Soubrie, 1986).
range of personality traits. This evidence appears amenable to interpretation in terms of two-mode models, as follows.

Relations between low serotonergic function and hostile and depressive trait tendencies can be seen as reflecting frequent intrusion of strong emotions; as noted earlier, this has been argued to characterize reflexive-mode processing. Frequent strong emotions would also be consistent with brain mechanisms reviewed earlier that yield a relative lack of executive control over emotional reactions. Relations between low serotonergic function and impulsive aggression (or other problematic impulsive actions) appear to reflect a lack of effortful control by the deliberative system over spontaneous action (see also Wilkowski & Robinson, 2008, regarding the role of effortful control in aggression).

Relations between serotonergic function and agreeableness (found frequently) and conscientious and related qualities (found less frequently) both concern qualities of personality that are addressed by two-mode models. Both sets of associations imply that high serotonergic functioning permits a relatively broader perspective on behavioral decisions—broader than the impulsive grabbing of incentives or fleeing from threats. Conscientiousness entails a broadened temporal perspective: taking future contingencies into account in one’s actions. Agreeableness entails a broadened social perspective: taking the needs of others into account in one’s actions. Indeed, consistent with the view that these traits reflect operation of the reflective mode, some argue that both of these traits have their temperamental origins in effortful control (Ahadi & Rothbart, 1994; Caspi & Shiner, 2006; Jensen-Campbell et al., 2002).

In further support of the idea that taking future contingencies into account entails effortful control, trait self-control (a closely related concept) has been linked to higher grades (Tangney, Baumeister, & Boone, 2004; Wolfe & Johnson, 1995), less procrastination (Steel, 2007), and less criminality (Krueger, 2002). In further support of the idea that taking into account the needs of others entails effortful control, trait self-control has been linked to accommodative behaviors in intimate relationships, which helps to sustain those relationships (Finkel & Campbell, 2001), an effect that held even when controlling for level of commitment to the relationship. Similarly, trait self-control has been linked to relationship satisfaction and relationship longevity among both romantic partners and best friends (Vohs, Lasalata, & Fennis, in press).

As noted earlier, Henning et al. (2005) suggested that the low serotonergic pattern resembles psychotimism, a trait characterized by impulsive social disconnection. Impulsive social disconnection seems an apt characterization of the combination of low agreeableness and low conscientiousness, both of which have been linked to psychopathy (Derefinko & Lynam, 2006).

Serotonin and Depression Vulnerability

Previous sections of the article relate low serotonin function to behavioral qualities such as reactivity to emotional information, poor impulse control, and uncooperativeness. A section on personality related serotonin function to traits that are conceptually similar to those behaviors, traits such as aggressiveness, impulsiveness, and disagreeableness. Correlates of low serotonin function from these various sources appear to share the property of poor control over (usually emotion-driven) responding. In contrast, this section turns away from impulsive behavior to what appears to be a quite different quality: depression.

Some information on depression has already been reviewed: The section on traits included evidence linking low serotonergic function to depressive tendencies. That evidence came from correlations with broad personality inventories, however, and is more suggestive than definitive. This section reviews research that makes a case for a link between low serotonergic function and the more explicit assessment of elevated depressive symptoms and diagnosis of depression.

Depression and Response to Drug Challenge and Manipulated Serotonin Availability

Several studies have found blunted prolactin response to fenfluramine among depressed persons (Flory et al., 1998; Lichtenberg et al., 1992; O’Keane & Dinan, 1991; Siever, Murphy, Slater, de la Vega, & Lipper, 1984). There is also evidence of a blunted response to serotonin challenge (including citalopram) among persons with a history of depression who are not currently depressed and are not medicated (Bhagwagar, Whale, & Cowen, 2002; Flory et al., 1998). This evidence is not entirely consistent, however. Occasional studies have found elevated rather than blunted responses among depressed persons, though generally in small samples (e.g., Strickland, et al., 2002).

Evidence of the involvement of serotonin in depression also comes from studies of acute tryptophan depletion (temporarily reducing available serotonin) among persons who were formerly depressed (see Moore et al., 2000, for a review). This procedure causes a temporary increase in depressive symptoms among some formerly depressed persons (Booij & Van der Does, 2007; Booij et al., 2002; Booij, Van der Does, et al., 2006; Delgado, Charney, Price, & Aghajanian, 1990; Delgado et al., 1999, 1999; Lam et al., 1996; Moreno et al., 1999; Munafò, Hayward, & Harmer, 2006; Neumeister et al., 2006; Neumeister et al., 2004; K. A. Smith, Clifford, Hockney, Clark, & Cowen, 1997; but see Leyton et al., 1997, for a nonreplication). In one case, tryptophan depletion induced cognitive deficits on the Mini-Mental State Examination (Porter et al., 2005). In other studies, tryptophan depletion produced memory impairments (Sambeth et al., 2007). In another study, tryptophan depletion increased cerebral glucose metabolism in the OFC cortex, anterior and posterior cingulate cortices, and ventral striatum in remitted participants but not in control participants (Neumeister et al., 2004). The return of depression symptoms has been found among remitted persons who are medicated free (Neumeister et al., 2006) as well as among those taking SSRIs (Delgado et al., 1999).

Persons whose first-degree relatives have a history of major depression also appear to be at elevated risk of negative affect after tryptophan depletion, compared with those with no family history of major depression (Benkelfat et al., 1994; Klaassen et al., 1999; Neumeister et al., 2002). Depression family history has been shown to increase risk of vegetative symptoms of depression after tryptophan depletion (Klaassen et al., 1999).

Serotonin Transporter Polymorphism and Depression

Early studies of the serotonin transporter gene looked for a direct (i.e., unmoderated) link between variations in 5-HTTLPR...
and clinical depression. Several studies found such relations (Collier et al., 1996; Hoefgen et al., 2005), but others did not. A meta-analysis of earlier studies (Anguelova, Benkelfat, & Turecki, 2003) reached a negative conclusion. Hoefgen et al. (2005) noted that that meta-analysis involved samples of great diversity, in contrast to their homogenous sample. They suggested that obtaining small effects (such as they found) may require substantial samples and also less variability attributable to sources other than the polymorphism.

A number of other studies have investigated the possibility of an interaction between this polymorphism and adverse experience (a diathesis–stress approach). For example, Caspi et al. (2003) reported that this polymorphism interacted with stressful life events to predict depressive symptoms and depression diagnosis: Negative life events had an adverse effect on those carrying at least one short allele, but not among those with two long alleles. This suggests that the short allele marks a genetic vulnerability: A factor present before onset, which interacts with triggers to increase the risk of onset (Monroe & Simmons, 1991). Without adversity, this genetic makeup did not yield depression, but this genetic makeup rendered people reactive to adversity.

A number of other studies have yielded similar results. These studies have examined children (Kaufman et al., 2004, 2006), adolescents (Eley et al., 2004), schoolteachers (Wilhelm et al., 2006), young adults (Kendler, Kuhn, Prescott, & Riley, 2005), and adults exposed to hurricanes (Kilpatrick et al., 2007). In all of those cases, the short allele of the serotonin transporter polymorphism was associated with elevated risk of depression, but only in combination with recent stressful life events or histories of maltreatment (and low social support, in the case of Kilpatrick et al., 2007).

Another recent study of Korean elders (Kim et al., 2007) found an interaction between serotonin transporter gene and stressful life events, such that carriers of the short allele who had had higher levels of adversity over the preceding year were more likely to be depressed. At 2.5 year follow-up of those not depressed at baseline, those with two short alleles who had experienced more negative events in the preceding year were more likely to be depressed.

Other studies have yielded partial replications of this pattern, but with qualifications. One study examined a large general population sample (Grabe et al., 2005), finding effects among women, but not men, and in interaction with other variables. In this case, the short allele was related to reports of both mental and physical distress among women who were unemployed, and (independently) among those who had from one to three chronic diseases. Another study of adolescents found that the short allele predicted depressive symptoms in interaction with environmental stress among female participants but had the opposite effect among male participants (Sjoberg et al., 2006).

S. E. Taylor et al. (2006) found an interaction on self-reported depressive symptoms: Among persons reporting adverse early family environment, those homozygous for the short allele (but not heterozygous) reported elevated depressive symptoms. Among persons reporting particularly benign early environment, those homozygous for the short allele reported the lowest depressive symptoms. S. E. Taylor et al. (2006) concluded that the short–short genotype is not so much a risk factor for depression as it is a magnifier of the impact of the environment on the person’s well-being. If the person is exposed to great adversity, the result may be enhanced symptoms of depression. If the person is instead exposed to a very supportive environment, the result may be a positive emotional life.

Jacobs et al. (2006) examined a large sample of women over a 15-month span, using self-reports of depressive symptoms as the outcome. They found an interaction of genotype with neuroticism, such that the presence of the short allele appeared to enhance the relation of neuroticism with depressive symptoms. This effect was independent of adverse life events, though there also was an interaction of neuroticism with adverse life events. The authors concluded that the polymorphism moderates the way in which persons respond to their daily life experiences. That is, neuroticism is associated with more emotional responses to adversity, but the impact of neuroticism on that emotional response is greater if the person carries the short allele of the serotonin transporter gene. Importantly, this suggests that low serotonergic function enhances stress reactivity even within a given level of neuroticism.

Zalsman et al. (2006) examined 191 persons with mood disorder and 125 healthy control participants, in a triallelic design (grouping the less expressive long allele with the short allele). They found that genotype did not distinguish patients from control participants, but among patients it did distinguish those with more severe depressive symptoms. The latter finding emerged both as a main effect of genotype and an interaction of genotype with life stresses.

Not all attempts to replicate the diathesis–stress effect have been successful. For example, in a sample of 1,206 twins, Gillespie, Whitfield, Williams, Heath, and Martin (2005) found that stress predicted major depressive disorder, but genotype played no role. Similar results were obtained by Surtees et al. (2005). Gillespie et al. pointed out that earlier demonstrations of the interaction had generally used young adults and that only 20% of their sample was below age 30. They speculated that genetic differences may contribute to vulnerability to different degrees at different ages. However, that interpretation does not fit the findings of Grabe et al. (2005) or Kim et al. (2007) described earlier.

Although most genetic research pertaining to serotonergic function and depression has examined the 5-HTTLPR polymorphism, there have also been a more limited number of studies of other polymorphisms. An example is a study of the tryptophan hydroxylase 1 gene (Jokela, Rääkkönen, Lehtimäki, Rontu, & Keltikangas-Järvinen, 2007). As noted earlier, tryptophan is involved in the synthesis of serotonin, and genes influencing tryptophan would ultimately influence serotonin. This study found that depressive symptoms were more common, both cross-sectionally and prospectively, in persons with a combination of low social support and the A allele of the gene. Another example is a study of the serotonin receptor 2A gene (Jokela, Keltikangas-Järvinen, et al., 2007). In this study, the lowest symptoms of depression were among people with the combination of the T genotype of the T102C polymorphism and high maternal nurturance.

A methodologically sophisticated study that combined diagnostic status, the serotonin transporter polymorphism, and tryptophan depletion was conducted by Neumeister et al. (2006). They gathered PET data from persons with remitted recurrent major depressive disorder who had not taken antidepressants in at least 3 months (plus never-depressed control participants). A significant interaction emerged among depression status, genotype, and tryp-
tophan depletion predicting neural responses. Persons with no depression history generally did not have strong responses to tryptophan depletion. Among persons with remitted depression, however, those with the short allele displayed greater activation of the hippocampus, the amygdala, and ACC in response to tryptophan depletion.

**Summary of Association Research**

The results outlined in the previous sections can be summarized as follows. Despite some nonreplications and qualifications, research with diverse methodologies links low serotonergic function to elevated symptoms of depression and to clinical depression. Depressed persons have displayed a blunted response to drug challenge (compared with other persons), and manipulations that reduce available serotonin have induced symptoms among previously depressed persons. Similarly, the short allele of the 5-HTTLPR has been linked to elevated risk of depression, though that link has often depended on the additional presence of adverse life experiences. Thus, there seems to be solid evidence that connects low serotonergic function to depression.

**Gender**

An issue that should be mentioned in this context concerns the gender difference in depression. Women are roughly twice as likely as men to experience clinical depression in their lifetimes (Kessler et al., 2003). Heritability estimates for depression are also higher among women than among men (Kendler, Gatz, Gardner, & Pedersen, 2006). This suggests that at least some of the gender difference in depression risk may reflect biological vulnerabilities.

In light of this difference, it is of considerable interest that gender differences also exist pertaining to serotonergic function. For example, there is considerable evidence that women are more susceptible than are men to negative mood changes after acute tryptophan depletion (Booij & Van der Does, 2007; Booij et al., 2002; Ellenbogen, Young, Dean, Palmour, & Benkelfat, 1996; Moreno, McGahuey, Freeman, & Delgado, 2006; K. A. Smith et al., 1997; Walderhaug et al., 2007). Also, as noted earlier, two studies relating the short allele of the 5-HTTLPR polymorphism in interaction with stress to depressive symptoms found this association among women but not men (Grabe et al., 2005; Sjoberg et al., 2006). A recent analysis pooling nine studies found that tryptophan depletion creates greater memory impairments among women than among men (Sambeth et al., 2007).

This is not to say that the short allele has behavioral consequences only among women. However, the pattern of consequences may differ by gender. For example, one study has found that the short allele predicts greater aggression in response to stress among men but not women (Verona, Joiner, Johnson, & Bender, 2006). Another found that the short allele of the serotonin transporter gene interacted with adverse childhood environment among men to produce violent adult behavior (Reif et al., 2007). Perhaps the effects of low serotonin function depend partly on interactions of serotonin with sex hormones (L. H. O’Connor & Fischette, 1986).

**Serotonin and Symptom Profiles**

Another important question is whether serotonergic function relates selectively to some particular aspect of depressive symptomatology or whether the relationship is more general. Several studies have provided evidence on this question. In research on persons who have recovered from depression, acute tryptophan depletion has been found to have an impact on both interview and self-report measures of anxiety and depressive symptoms as well as on self-reported somatic symptoms of depression (Booij, Swenne, et al., 2006; Spillmann et al., 2001). Other studies have found that acute tryptophan depletion produces effects of comparable size for mood and vegetative symptoms (Riedel, Klaassen, & Schmitt, 2002) and for REM sleep deficits that are typically associated with depression (Moore et al., 1998).

Evidence linking suicidality to low serotonergic function was described earlier. Some of the studies that related depression to the serotonin transporter polymorphism also examined suicidality. Caspi et al. (2003) found that the short allele was linked to higher risk of suicidality in interaction with life stressors. Zalsman et al. (2006) found that genotype did not relate to suicidality within the depressed group, but they did not test for an interaction with life events, nor did they compare the depressed group with the control group.

One might also ask whether there are differences between major depressive disorder and depression associated with bipolar disorder. Recent research (Sher et al., 2003; Sobczak, Honig, van Duinen, & Riedel, 2002) suggests that hormonal responses to fenfluramine and to tryptophan are blunted to a similar degree among people with unipolar and bipolar depression. Unaffected family members of persons with bipolar disorder have also shown atypical cognitive (Sobczak, Riedel, Booij, Het Rot, & Deutz, 2002) and mood responses (Quintin et al., 2001) to serotonin challenge; this suggests that deficits in serotonergic function in bipolar disorder do not reflect medication effects or scars of previous episodes.

In sum, there is substantial evidence of the involvement of serotonergic function as a contributor to risk of depression. The possibility that serotonergic function relates selectively to some symptoms but not others among depression-vulnerable persons is intriguing. However, support for that possibility is generally lacking.

**Differentiating Vulnerabilities: Impulse Problems Versus Depression**

An earlier section linked low serotonergic function to behavioral problems in which a salient feature was poor impulse control. Given various evidence reviewed prior to that section, it is not surprising that low serotonergic function relates to disorders involving impulsiveness. That is, in diverse contexts, markers of low serotonergic function have been linked to impulse expression, particularly impulsive reactions to emotion cues.

It is less obvious at first glance why low serotonergic function would be a vulnerability factor for depression. Although depression has sometimes been linked to impulsive qualities (Peluso et al., 2007), it is more frequently associated with lethargy, an absence of behavioral engagement. How can we account for this very substantial difference in presentation?

Earlier in the article, the balance of reflexive to deliberative control over behavior was characterized in terms of relative responsiveness to salient cues (as argued by Spoont, 1992). Deficits in deliberative control enhance responsiveness to cues of the moment, particularly emotion cues. What kind of behavior follows
from that responsiveness, however, depends on what the salient cues are. That, in turn, depends partly on the person’s emotional and motivational state (or sensitivity). Thus, low serotonergic function and the resulting deficits in effortful control may have divergent effects in people who vary systematically in other ways.

**Interaction Between Systems**

What other systematic variation might interact with serotonin function to yield such a divergence? A prime candidate is variation in the degree of sensitivity, or engagement, of a general incentive approach system. When poor executive oversight is combined with moderately high incentive sensitivity (a fairly reactive approach system), the result is impulsiveness. When poor executive oversight is combined with low incentive sensitivity (a nonreactive approach system), the result is inaction: lack of effort toward potential rewards. In both cases, the effects of variation in level of incentive sensitivity (high and low, respectively) are amplified by the absence of effortful override (Figure 2, earlier).

In the case of depression vulnerability, the intercession of low serotonergic function and nonreactive incentive system could create difficulties in two ways. First, a lack of incentive sensitivity means that the person is not strongly motivated to approach potentially rewarding contexts. A relative lack of effortful control will amplify this problem, such that the depression-vulnerable person will have greater difficulty overriding this lack of motivation so as to engage in actions. This combination thus should yield apathy, passivity, and fatigue, which characterize many cases of depression. Second, if a deficit in executive processes leaves the amygdala relatively uncontrolled and thus hyperactive, the salience of emotional associations should also be enhanced, which presumably are relatively negative in the person vulnerable to depression. Thus, low serotonergic function would have two adverse effects. It would hamper the ability to keep negative emotions at a distance. It would also hamper the ability to engage in effortful pursuit of incentives (Salame, Correa, Farrar, & Mingote, 2007).

**Blunted Incentive System in Depression**

Evidence that depression is associated with a blunted approach system comes from several sources. One of them is an earlier literature using EEG to examine anterior cortical asymmetry and motivational processes. This technique has far poorer spatial resolution than do newer imaging techniques, and thus it is far less specific about what areas are activated. However, it has been used to make a case concerning partial localization of approach and avoidance systems in left and right hemispheric areas, respectively (e.g., Davidson, 1998).

Previously depressed (Henriques & Davidson, 1990) and clinically depressed persons (Henriques & Davidson, 1991) have been found to have lower activation in left anterior cortical areas than do nondepressed persons, with no difference in right anterior activation. The lower activation was in areas that previous research had related to approach sensitivities (e.g., Coan & Allen, 2003; Harmon-Jones & Allen, 1997; Sutton & Davidson, 1997). This general pattern regarding depression has been replicated a number of times (J. J. Allen, Iacono, Depue, & Arbis, 1993; Gotlib, Ranganath, & Rosenfeld, 1998), though not all attempts have been successful (Reid, Duke, & Allen, 1998). These findings as a group suggest that depression relates to low engagement of areas pertaining to approach. One recent study has even linked early-onset depression to a deficit in left prefrontal responsiveness to a task that had been designed specifically to engage approach motivation (Shankman, Klein, Tenke, & Bruder, 2007).

Behavioral research also suggests that depression relates to blunted incentive sensitivity. For example, depressed persons were found to be less responsive to reward (though not to punishment) than were nondepressed persons (Henriques & Davidson, 2000; Henriques, Glowacki, & Davidson, 1994; see also Depue & Iacono, 1989). There is also evidence that dysphoric persons engage more effort at easy task levels than do nondysphoric persons, as though they overestimated the effort that is required by the task, but that dysphoric persons are more likely to disengage effort when the task actually is demanding (Brinkmann & Gendolla, 2008).

Other evidence relates self-reports of low incentive sensitivity to depression (Campbell-Sills, Liverant, & Brown, 2004; Finto-Meza et al., 2006). Indeed, three separate studies have found that self-reports of low incentive sensitivity predicted worse course of depression over time (Campbell-Sills et al., 2004; Kasch, Rottenberg, Arnow, & Gotlib, 2002; McFarland, Shankman, Tenke, Bruder, & Klein, 2006).

Blunted approach motivation may also be reflected in low dopaminergic function. Dopaminergic pathways are believed to be critical in the engagement of goal-directed effort (Farrar et al., 2007; Salamone et al., 2007; Salamone, Correa, Mingote, & Weber, 2005; Salamone, Correa, Mingote, Weber, & Farrar, 2006). A weakly functioning dopaminergic incentive system yields less “wanting” for appetitive outcomes (Berridge, 2007) and less engagement of effort in pursuit of them (Salamone et al., 2005, 2006, 2007). A recent comprehensive review reported a range of evidence for deficits in the function of dopamine among depressed persons, drawing from pharmacological studies, genetic studies, and dopamine challenge studies (Dunlop & Nemeroff, 2007).

Animal research has long implicated subcortical reward pathways in behaviors that mimic depressive symptoms, such as diminished motivation, apathy, and poor appetite (Nestler & Carlezon, 2006). Some brain studies in humans also link depression to deficits in subcortical regions involved in reward sensitivity, such as the caudate. For example, in a recent fMRI study, children diagnosed with depression displayed less activation of the caudate throughout a task involving reward, compared with children without depression (Forbes, Fox, Cohn, Galles, & Kovacs, 2006). Adults with depression have been found to show hyperreactivity of the caudate and the putamen to an amphetamine challenge that was designed to intensify dopaminergic activity (Tremblay et al., 2005); this pattern is consistent with there being a baseline deficit in dopaminergic activity.

**Incentive System in Impulsive Disorders**

What about impulse-related problems? A very substantial literature has addressed the role of dopamine and approach motivation (in interaction with serotonin) in externalizing disorders. Review of that literature, which includes studies using genetic and psychophysiological indices of various kinds, is well beyond the scope...
of this article. It is worth noting, however, that there is evidence linking externalizing disorders to elevated approach tendencies.

The concepts of reactive undercontrol (Eisenberg et al., 2004) and motivational inhibition (Nigg, 2001, 2003) are also relevant here. As noted earlier, evidence links reactive undercontrol to externalizing disorders in childhood (Rothbart, Ellis, & Posner, 2004; Valiente et al., 2003). Similarly, Nigg has held that the core issue in conduct disorder (and by implication psychopathy as well) is lack of a sufficiently reactive threat system. In principle, however, “sufficiently reactive” means only that the threat system is more reactive than is the incentive system, reactive enough to hold it in check. This could mean an insensitive threat system in combination with no special sensitivity of approach; it could also mean a relatively engaged incentive system with no special deficit of avoidance.

There is at least some evidence of both a weak threat system and a strong incentive system in psychopathy. In one study, various measures of psychopathy correlated with lower Anxiety facet scores on the NEO–PI–R and with higher scores on Assertiveness and excitement seeking facets of Extraversion, both reflecting incentive pursuit (Derefinko & Lynam, 2006). Another study relating self-reports of incentive and threat sensitivities to diverse externalizing symptoms found that self-reports of high reward sensitivity predicted primary and secondary psychopathy and impulsive ADHD symptoms (Hundt, Kimbrel, Mitchell, & Nelson-Gray, 2008). Low threat sensitivity predicted secondary psychopathy and inattentive ADHD symptoms.

Section Summary

The question of how to interpret the very different manifestations of low serotonergic function in impulsive disorders and depression prompts the search for an interaction. There is support for such an interaction: between the serotonergic system (and associated executive deficits) and a general system of incentive sensitivity, in the form of evidence of a blunted approach system in depression.

Two-Mode View of Depression: Additional Evidence Concerning Brain Function

The link between two-mode models and depression has been made here in terms of the involvement of serotonergic function in brain functions that are consistent with the two-mode models, and the involvement of serotonergic function in risk of depression. Evidence from other sources further supports that link. This section reviews one source: neurobiological research on depression. Specifically, evidence suggests involvement in depression of brain areas that are strongly influenced by serotonin function.

Elevated Amygdala Activity in Depression

Findings from both structural and functional imaging studies have indicated that depression involves hyperactivity in brain areas that have been related to the reflexive mode (Mayberg, 1997, 2002). More specifically, in various imaging studies, depression has been related to elevations in blood flow, blood oxygenation, and glucose metabolism in the amygdala (Drevets et al., 1992, 2002). Indeed, amygdala activation has been found to correlate positively with depression severity (Abercrombie et al., 1998; Drevets et al., 1992). Amygdala activation normalizes following antidepressant treatment (Sheline et al., 2001).

Depression has also been related to heightened amygdala responses to fearful or negative faces (e.g., Breiter et al., 1996), a pattern that seems to diminish with antidepressants (Drevets, 2001; Sheline et al., 2001; for an extended discussion, see Davidson, Pizzagalli, Nitschke, & Putnam, 2002). Similarly, processing of negative words has been related to greater amygdala activation among depressed persons than among control participants (Siegle et al., 2002). Finally, persons who report being ruminators, a quality that renders them vulnerable to depression (Nolen-Hoeksema, 2000; Nolen-Hoeksema, Morrow, & Fredrickson, 1993), show greater and longer amygdala activation when asked to temporarily increase their negative affect, compared with other people (Ray et al., 2005; Siegle et al., 2002).

Hypoactivation in Areas of PFC in Depression

Amygdala hyperactivity is one potential source of problems relevant to depression, by rendering existing negative emotions more salient (Spoon, 1992). A complementary source of problems, suggested by literatures reviewed earlier, is a link between depression and low activity in parts of the PFC believed to regulate affect by inhibiting activation in subcortical areas or activation in ventral parts of the PFC. Evidence fits this picture, as well (Levin, Heller, Mohanty, Harrington, & Miller, 2007; Mayberg, 2004).

A number of studies have found that dorsolateral PFC activation in depressed persons is low, both at baseline and during tasks (e.g., Baxter et al., 1989; Bench, Friston, Brown, Frackowiak, & Dolan, 1993; Davidson, 2000; Drevets, 1998; Siegle et al., 2002), while activation in ventral PFC and OFC are high (R. J. Dolan, Bench, Brown, Scott, & Frackowiak, 1994; R. J. Dolan et al., 1993). Antidepressant medication normalizes both of these differences (Kennedy et al., 2001; Mayberg, 1997). Other studies have found abnormally low cerebral blood flow and metabolism among depressed persons in areas of the lateral and dorsolateral PFC and the ACC (for reviews, see Drevets, 1999, 2000; Drevets & Raichle, 1998). These functional differences may be due to structural abnormalities associated with depression, such as lower neuron and glial count and size (e.g., Cotter et al., 2002; Rajkowska, 2000).

The key point here is the potential impact of this pattern on the affect-regulatory corticolimbic circuitry as a whole. Low activation in regions of the PFC that either project to subcortical structures, or project to other parts of the PFC that play an important role in inhibiting activity in subcortical regions, may reduce the constraining influence on those structures. The result can be a prolonged activation of the subcortical areas in response to challenges (cf. Siegle et al., 2002). This may be expressed as sustained negative affect. Indeed, Mayberg (1997, 2003) has proposed a model of depression involving disruption of corticolimbic pathways, based on evidence of abnormalities in these pathways among depressed patients and changes in them after treatment.

Section Summary

Neurobiological studies suggest that several aspects of brain function which are influenced by serotonergic function may also relate to depression. The effects include amygdala hyperreactivity
and diminished functionality of specific prefrontal areas. This evidence suggests multiple processes: Depression likely involves subcortical functions yielding high negative emotion. Depression may also involve deficits in executive control over these systems, minimizing the likelihood that they will be overridden.

Two-Mode View of Depression: Additional Behavioral Evidence

The perspective of two-mode models also provides a different angle on some of the behavioral phenomena known to be related to depression. From this perspective, people who are vulnerable to depression should be deficient in executive control of their actions or emotions (or aspects thereof). Additional evidence concerning psychological and behavioral manifestations of depression supports the plausibility of that view. Indeed, the idea that depression can be usefully viewed from a two-mode framework has already appeared in several places (Beevers, 2005; Conner & Barrett, 2005; Dalgleish et al., 2007; Haefel et al., 2007).

Implicit System and Vulnerability

Beevers (2005) proposed a two-mode model of depression vulnerability that assumes many of the same conceptual underpinnings as we suggest here but that has a rather different focus. Beevers focused primarily on how the implicit system might tend to induce and maintain depression. He argued that persons with many negative associations in the implicit self therefore have a “pattern-completion” mechanism (a reflexive, connectionist mechanism) that is negatively biased when processing information about the self. When experiences are assimilated to this schema, they evoke the negative feelings that are linked to these negative associations. For that reason, these people are likely to experience negative feelings more frequently than are other people.

Several findings from other sources are consistent with that position. For example, negative responses on implicit measures of self-constructs did better than responses on explicit measures at predicting negative emotional reactions to laboratory stressors (Haefel et al., 2007). Similarly, Conner and Barrett (2005) have confirmed that people who have negative implicit self-concepts experience more negative affect in daily life than do people whose implicit self-concepts are less negative.

Indeed, negative associations in the implicit self can multiply over time, if the person develops an automatic tendency to think negatively about the self (Verplanken, Friborg, Wang, Trafimow, & Woolf, 2007). Such a tendency leads to further encoding of negative associations. Reports of habitual negative thinking about the self have been shown to relate to lower implicit self-esteem and to predict variance in depressive symptoms over and above that predicted by explicit negative content of thoughts about the self (Verplanken et al., 2007).

Beevers (2005) argued that, because of the automatic pattern-completion tendency, people who are depressed must engage in an effortful (reflective) corrective process to counter the negative associations. If the effortful process fails to occur, the implicit self maintains control over experience. As experiences continue to be assimilated to the negative implicit self, negative affect continues to be evoked and depression becomes more likely. Maintenance of an existing depression also becomes more likely.

Deficits in Effortful Control

Effortful processing requires both a reason to do so and the capacity to do so (Barrett et al., 2004). With respect to the former, Beevers (2005) held that a negative association triggers an effortful correction only if the association is unexpected, and thus calls for attentional resources to sort things out. However, a negative association would be unexpected only if the person’s explicit self-view was positive. Inasmuch as depressed people typically describe themselves negatively on explicit self-view indices (Abramson et al., 2002), a negative implicit association would generally not violate their expectancies and would not trigger effortful override. Thus, a prompt for effortful processing may often be lacking.11

Other views of depression also emphasize the idea of cognitive control deficits. Hertel (2000), for example, viewed negative thoughts, biased processing of emotional information, and rumination thinking in depression as automatic patterns activated by negative mood states. Those patterns dominate unless overridden by effortful cognitive control (e.g., Hertel & Rude, 1991).

Deficits in effortful cognitive control can yield diverse outcomes. In some cases, it can keep the person from moving on. For example, depression has been related to impaired shift in behavior after error feedback (Compton et al., 2008; Holmes & Pizzagalli, 2007). Similarly, there is evidence that depressive rumination interferes with disengaging from previously relevant but no longer relevant information, as in task switching (Whitmer & Banich, 2007).

In other cases, deficits in effortful cognitive control can keep the person from remaining task focused. Thus, there is also broad evidence of mental restlessness and distractibility among depressed persons. Smallwood and his colleagues have repeatedly found that depression relates to mind wandering across diverse tasks (for a review, see Smallwood, O’Connor, Sudberry, & Onsawin, 2007; see also Smallwood & Schooler, 2006). Greater mind wandering has also been linked to lower levels of working memory capacity (Kane et al., 2007). The distraction of mind wandering among depressed persons interferes with completion of planful activities.

Other phenomena related to depression also fit this view of executive control deficits (see also Levin et al., 2007). For example, depression is well known to be associated with autobiographical recall of memories that are very general in character rather than specific (Williams et al., 2007). A recent set of studies has

---

11 Authors differ in where conceptually to assign the risk (and why). Haefel et al. (2007) argued that explicit interpretations are the ultimate risk factor, because they represent the final interpretation of the event. Thus, they suggested that people with positive implicit self-construal can be vulnerable if they have negative explicit self-construal. Beevers (2005) viewed the negative implicit self-construal as the real culprit, but its impact is great only if not overridden. On the other hand, he saw the override as unlikely if the explicit self-construal is negative. Thus, as a practical matter, those conceptions differ only with respect to the case of a positive implicit self and a negative explicit self. Our position is that persons who are vulnerable to depression are impaired in the ability to readily evoke the reflective viewpoint that the explicit self-construal depends on. Thus the positive self-construal is unlikely to be brought to bear. Our view therefore would seem somewhat closer to that of Beevers (2005) than to that of Haefel et al. (2007).
yielded convincing evidence that this effect stems from executive control deficits (Dalgleish et al., 2007).

There is also evidence that depressed persons have trouble ignoring negative information (Frings, Wentura, & Holtz, 2007) and disengaging from negative information once attended to (Cerasas, Garner, Bradley, & Mogg, 2007; Joormann & Gotlib, 2008; Joormann, Yoon, & Zetsche, 2007). Several measures have been developed to assess the capacity to effortfully suppress attention to negative material. Examples are the emotion Stroop task, measures of negative priming (Joormann, 2004; Wentura, 1999), and attention tasks with a long stimulus duration (Derryberry & Reed, 2002; Vasey, 2003, cited in Lonigan, Vasey, Phillips, & Hazen, 2004). Depressed persons show negative biases on the emotion Stroop, negative priming tasks, and attention tasks with long stimulus durations (Goeleven, De Raedt, Baert, & Koster, 2006; Gotlib et al., 2004; Gotlib & Krasnoperova, 1998; Joormann, 2004). All of this represents further evidence of deficits among depressed people in effortful cognitive control, in this case, specifically effortful control over thoughts.

Cognitive Effects of Tryptophan Depletion in Depression

Do these cognitive control deficits stem from low serotonergic function? Information on that question comes from studies of acute tryptophan depletion. As noted earlier, among healthy persons, acute tryptophan depletion affects cognitive processes involved in inhibiting responses to emotion-relevant stimuli (e.g., S. E. Murphy et al., 2006). Munafò et al. (2006) found that recovered but medicated depressed persons had particularly strong increases in Stroop interference for threat cues after acute tryptophan depletion; the effect of tryptophan depletion was not seen among healthy control participants nor among persons with remitted depression who were unmedicated.

Other studies have shown that such effects among formerly depressed persons are more pronounced for tasks with negatively valenced stimuli than for tasks without emotion-relevant stimuli. For example, Booij, Van der Does, et al. (2006) studied the effects of acute tryptophan depletion on processing of words in a Stroop task. Tryptophan depletion speeded processing of neutral and sad words, but it slowed processing of happiness-related words.

Hayward, Goodwin, Cowen, and Harmer (2005) also examined the effects of low-dose tryptophan depletion on persons with and without a history of depression, for emotion-relevant and other tasks. Recovered depressed people with acute tryptophan depletion displayed more elevation of emotion-potentiated startle and slower recognition of happy facial expressions, compared with other people. Tryptophan depletion and diagnostic group, however, did not interact in predicting non-emotion-relevant tasks such as a counting Stroop. Again, tryptophan depletion (and thus serotonin reduction) rendered formerly depressed people more responsive to negative emotional cues.

Restoring Effortful Control and Impediments to Doing So

If effortful cognitive control among depressed persons is poor, how can it be enhanced? Beever (2005) pointed out that much of cognitive therapy involves training the person to use effortful and rule-based processing to identify and correct underlying negative biases. Although it is common to think of cognitive therapy as being a simple self-corrective process, it has always been clear that more is involved than merely providing positive self-statements. The two-mode view provides a good context for understanding effects of that intervention (cf. Daw et al., 2005). The intervention is, in effect, an attempt to restore effortful management of what is currently automatic behavior (see also Wood & Neal, 2007). This view helps account for the fact that cognitive therapy is a substantially more arduous process than one might initially assume (cf. Strunk, DeRubeis, Chiu, & Alvarez, 2007). There is evidence that it is possible to increase working memory capacity (Friedman et al., 2008; Olesen, Westerberg, & Klingberg, 2004), and it may be that such changes themselves are an outcome of cognitive therapy techniques.

Even if a person with depression tries to use such corrective processes, there are reasons why their implementation may fail. For one, the person must have the attentional capacity to do so (Barrett et al., 2004). Evidence suggests that persons with depression do not easily deploy effortful control. If the person’s cognitive resources are stretched thin and thus unavailable, the correction is even less likely to take place.

Ironically, one thing that can deplete cognitive resources is negative affect. Thus, the emotional distress experienced in depression can help create the very conditions that make it difficult to counter further negative implicit processing. Beever (2005) noted that even when an effortful correction is attempted, it may be insufficient. He further reasoned that a downward spiral may be induced, as cognitive resources are stretched by the need to counter negative thoughts. Indeed, there is evidence that the attempt to suppress or inhibit information attaches negative valence to that information (Fenske & Raymon, 2006).

Consistent with that reasoning, Wenzlaff and Bates (1998) have shown that diminished working memory capacity (induced by a cognitive load) led depressed persons to display more negative bias on a scrambled sentence task. Intriguingly, negativity under cognitive load was a better predictor of depressive symptoms over time than was negativity without cognitive load. Hence, the cognitive resources available to correct implicit biases may be particularly important for the course of depressive symptoms.

Section Summary

A good deal of evidence suggests that persons with depression have difficulties exerting effortful cognitive control. Their attention is drawn to negative emotional cues (perhaps because of their many negative associations in memory), which they then have trouble disengaging from. They are prone to rumination about negativity but also prone to mind wandering rather than task engagement. Those recovered from depression experience recurrence of sensitivity to negative cues when their serotonin levels are artificially lowered. These various findings are consistent with a two-mode view in which depression relates to limited executive function.

Future Directions

An earlier section of the article suggests an interactive view of vulnerability to depression versus vulnerability to disorders involving impulsivity. It was suggested there that both tend to reflect low serotonergic function, and thus low effortful control, but that
they differ in level of engagement of the system responsible for
eager, effortful approach. Psychopathy and impulsive aggression
were characterized as reflecting low serotonergic function and
relatively high approach motivation (perhaps true of other
impulse-related problems as well; cf. Depue & Lenzenweger,
2005). Depression was characterized as reflecting low serotonergic
function and low approach motivation.

Further Interactions?

A broad question for the future is whether other interactions
should also be explored more fully (Depue & Lenzenweger, 2005;
Nigg, 2006). For example, perhaps overt expression of a vulner-
bility to anxiety disorders may also depend on poor executive
control (Lonigan et al., 2004). Consistent with this idea, serotonin
has been implicated in the development of anxiety disorders (see
Leonardo & Hen, 2006, for review). There is also evidence that the
5-HTTLPR polymorphism (in a triallelic design) interacts with
daily stress to influence reports of anxious mood (Gunhert et al.,
2007). Cools, Calder, and colleagues (2005) found that experi-
mental reduction of serotonin (via tryptophan depletion) led to in-
creased amygdala activation in response to fearful faces but did so
only among persons who had previously reported being high in
threat sensitivity.

These findings all suggest an interactive combination of a highly
sensitive threat system and low serotonergic functioning. How
the two-mode model might help in the understanding of anxiety-
based problems, and in distinguishing the etiology of anxiety
disorders from that of depression, would seem to be an important
topic for further investigation.

Low Serotonin Function as Amplifier

Many observers have noted that the attempt to link any given
neurotransmitter to the operation of a single behavioral system is
likely to be a great oversimplification; many have also noted that
effects of one neurotransmitter are likely to depend on the person’s
level of another neurotransmitter (Depue, 1995; Depue & Lenzen-
weger, 2005; Zuckerman, 2005). Indeed, what was known as
permissive theory (Prange, Wilson, Lynn, Alltop, & Stikleather,
1974) long ago argued specifically that low serotonin function
enhances risk of depression in interaction with other neurotransmit-
ners (see Howland & Thase, 1999). Characterizing one neuro-
transmitter as having an overall effect is hazardous.

Nonetheless, it does not seem too far an extrapolation from the
evidence to suggest that low serotonergic function promotes a
stronger manifestation of whatever tendencies the person has at the
reflexive or implicit level of functioning (for similar conclusions,
see Depue, 1995; Nigg, 2006; Spoont, 1992). In an incentive-
sensitive person, low serotonergic function amplifies the pursuit of
incentives. In an incentive-insensitive person, low serotonergic
function exaggerates the lack of effortful engagement. In a threat-
sensitive person, low serotonergic function may enhance vigilance
to threat.

Low serotonergic function may also promote a stronger mani-
festation of the pattern of conditioned associations the person has
acquired over time (S. E. Taylor et al., 2006). In a person who
experienced childhood maltreatment (and who thus has many
negative associations with the self in memory), low serotonergic
function may enhance the negative aftereffects of that maltreat-
ment (Caspì et al., 2003). In a person who had a very supportive
early family environment (and who thus has many positive ele-
ments in associative memory), low serotonergic function may
enhance the positive reverberations of that environment (S. E.
Taylor et al., 2006).

The specific cases of depression, impulsive disorders, and
anxiety disorders are only three possibilities, reflecting inter-
actions of a serotonergic system with two other systems. A
more complete understanding of the role of serotonin in behav-
ior will require a more elaborated understanding of how sero-
tonergic function interacts with effects of other neurotransmit-
ers. The idea that diverse disorders follow from diverse
combinations of system sensitivities (Depue & Lenzenweger,
2005; Lenzenweger & Willett, 2007) is very intriguing and
seems worthy of much more examination.

Effortful Control and Overcoming Inaction

Another sort of interaction also deserves more attention than it
has yet received. As noted earlier (Figure 2), the two-mode rea-
soning that guided this review suggests that high serotonergic
function, and the associated high degree of effortful control, does
more than restrain impulses. In theory (and in limited evidence),
effortful control can also cause a person who is unmotivated to ac-
to take action. In principle, then, effortful control should have
divergent effects, depending on the behavioral context. Effortful
control (and related constructs such as executive control, working
memory capacity, and serotonergic function) should be involved in
countermanding tendencies triggered by inhibitory reactive sys-
tems as well as impulsive reactive systems. This idea also deserves
closer attention.

Conclusion

Recent years have seen an explosion of interest in the role of
serotonergic function in depression and elsewhere. The developing
literatures use a wide array of research paradigms. The functions of
serotonin have been interpreted in terms of diverse constructs,
ranging from neuroticism and anxiety-related traits, to cortical
control over amygdala activity, to effortful cognitive processes.
Review of these literatures suggests that it is useful to conceptu-
alyze these various functions in terms of two-mode models of
self-regulation. Viewed through this lens, the evidence suggests
that serotonergic function can be linked to impulsivity versus
constraint in the personality literature, effortful control processes
in the cognitive and developmential literatures, and executive con-
tral over the amygdala and other subcortical areas in the neurobi-
ological literature. This two-mode picture helps organize what is
known about the experience of depression, and it may also be
useful in suggesting new areas of investigation.

This article was prompted by evidence that low serotonergic
function is a risk factor in depression. Though that evidence is
essentially neurobiological in nature, it has implications that are
cognitive, behavioral, and phenomenological. Ideas and evidence

12 It is perhaps noteworthy in this context that SSRIs are effective for
anxiety disorders as well as depression.
from several literatures that are psychological in nature appear to foster a deeper understanding of the meaning of this evidence regarding biological vulnerability. This confluence suggests another message as well. Increasingly it is held that biological concepts and methods are key resources to inform psychological theory (e.g., Cacioppo et al., 2007; Winkielman, Knutson, Paulus, & Trujillo, 2007). We would hold, however (see also Posner & Rothbart, 2007), that interpretation of neurobiological evidence also benefits from considering the findings through the lens of psychological principles.

References


system markers with adolescent depression. Molecular Psychiatry, 9, 908–915.


Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflec-


Murphy, F. C., Smith, K. A., Cowen, P. J., Robbins, T. W., & Sahakian, B. J. (2002). The effects of tryptophan depletion on cognitive and
sion: Relationship to cerebrospinal fluid monoamine metabolite levels. Biological Psychiatry, 50, 783–791.
Quay, H. C. (1988). Theories of ADDH.
Schafet, A., Collete, F., Wenniger, G., Van der Linden, M., Laureys, S.,


