Irritability in Youths: A Translational Model
Melissa A. Brotman, Ph.D., Katharina Kircanski, Ph.D., Argyris Stringaris, M.D., Ph.D., Daniel S. Pine, M.D., Ellen Leibenluft, M.D.

Although irritability is among the most common reasons that children and adolescents are brought for psychiatric care, there are few effective treatments. Developmentally sensitive pathophysiological models are needed to guide treatment development. In this review, the authors present a mechanistic model of irritability that integrates clinical and translational neuroscience research. Two complementary conceptualizations of pathological irritability are proposed: 1) aberrant emotional and behavioral responding to frustrative nonreward, mediated by reward-system dysfunction; and 2) aberrant approach responding to threat, mediated by threat-system dysfunction. The authors review the pathophysiological literature, including animal studies, as well as experimental psychology and clinical studies. Data suggest that, relative to healthy children, irritable children have deficient reward learning and elevated sensitivity to reward receipt and omission. These deficits are associated with dysfunction in the prefrontal cortex, striatum, and amygdala. Youths with irritability also show maladaptive orienting to, interpreting, and labeling of potential threats, associated with prefrontal cortical and amygdalar dysfunction. Abnormalities in reward and threat processing potentiate one another. Future work should test pathophysiological hypotheses and novel interventions targeting reward- and threat-related dysfunction to improve treatment for severe irritability in youths.


Irritability is among the most common reasons that children are brought for psychiatric evaluation and care (1, 2). While irritability has received increased research attention over the past two decades (3), few effective treatments are available. Given its public health importance (1, 2, 4–6) and the dearth of evidence-based treatments (7–9), developmentally sensitive pathophysiological models are needed to guide novel therapies for irritability. We present a mechanistic model of irritability that integrates clinical and neuroscientific research and has implications for treatment development.

First, we review the definition, longitudinal course, and prevalence of irritability. Next, we describe two conceptualizations of irritability based in translational neuroscience: 1) aberrant responding to blocked goal attainment (10) and 2) aberrant approach responding to threat (11). We link these two constructs in one overarching mechanistic model of irritability. Since irritability shows significant cross-species conservation of brain-behavior relationships, our translational model synthesizes neural, behavioral, and clinical data across human and animal research. Finally, after briefly reviewing data for existing treatments, we discuss ideas for novel, mechanism-based interventions.

We reviewed the literature prior to October 2016 using the search terms “irritability,” “anger,” and “frustrative nonreward.” Overall, 163 papers were included in the review: 14 animal studies, 71 human experimental studies, and 78 clinical studies. We focus on studies of irritability, severe mood dysregulation, and disruptive mood dysregulation disorder (DMDD). We also include studies on related clinical constructs, including externalizing or disruptive behavior disorders (defined here as attention deficit hyperactivity disorder [ADHD], oppositional defiant disorder, and conduct disorder). Thus, our model and its clinical implications are specifically targeted toward youths with DMDD or those with disruptive behavior disorders who exhibit chronic irritability.

CLINICAL DEFINITION, LONGITUDINAL COURSE, AND PREVALENCE

Irritability can be defined as an elevated proneness to anger relative to peers (12, 13). Proneness to anger is a trait that is distributed continuously in the population (14) and expressed stably across time (15). Proneness to anger is a dimension, and the precise cut-point demarcating normality from pathology varies developmentally (16). In this review, we focus on relatively severe manifestations of childhood irritability, that is, those that cause impairment and therefore necessitate treatment.

The DSM-5 criteria for DMDD represent one approach to operationalizing severe irritability in youths (17). Indeed, among children with oppositional defiant disorder, only those whose irritability severity is within the top 15% would meet criteria for DMDD (17). Youths with DMDD exhibit severe and recurrent temper outbursts that are more easily elicited, longer lasting, and contextually atypical relative to those of
their peers (16,18–20). Relative to peers, irritable youths have a lower threshold for expressing anger, leading to more frequent temper outbursts; when such outbursts occur, they have greater behavioral intensity (12, 13, 21). Outbursts are characterized by motor activity, prominent displays of anger and other negatively valenced emotions, and verbal as well as sometimes physical displays of reactive aggression (22). Between temper outbursts, severely irritable children also have a persistent angry mood, involving sullen nonverbal behaviors and reports of being annoyed over many days. Thus, DMDD includes affective and behavioral components, and our pathophysiological framework integrates both of these dimensions. However, it is important to note that the available data are equivocal as to whether temper outbursts and between-outburst negative mood are separable components (14).

The normative threshold, frequency, and behavioral manifestations of anger change over the course of development (23–27); for this reason, the criteria for DMDD stipulate that temper outbursts have to be inconsistent with developmental level (14, 16, 18, 23). For example, irritable behavior that would be normative in the preschool years or in early childhood (e.g., short-lived temper tantrums multiple times per week) could be abnormal in middle childhood. Retrospectively reported clinical data suggest that the average age at onset of DMDD symptoms is 5 years (3). However, because normative irritability peaks in the preschool years, the determination of a cut-point for psychopathology at that developmental phase is particularly challenging (14, 16, 18, 23). Therefore, DSM-5 does not allow the diagnosis to be assigned until age 6.

Clinically impairing irritability in children and adolescents began to gain more attention as interest grew in the diagnosis of pediatric bipolar disorder. Beginning in the 1990s, child psychiatry researchers suggested that while pediatric bipolar disorder can present with distinct episodes of mania or hypomania as in adults, the more typical pediatric presentation was chronic, severe irritability and hyperaffective symptoms (28, 29). However, a series of longitudinal (30), family (31), behavioral (32–34), and pathophysiological (35–38) studies differentiate classically defined episodic pediatric bipolar disorder from chronic irritability without distinct manic or hypomanic episodes (operationalized as severe mood dysregulation [28] (3)). For example, behavioral and functional MRI studies have found that while both youths with bipolar disorder and those with severe mood dysregulation have impairments in labeling face emotions (32, 33), the neural correlates of this deficit differ between the two groups (38–40). Indeed, more recent evidence indicates that the pathophysiological correlates of the trait of irritability itself differs between bipolar disorder and DMDD (40). Thus, the latter study addresses the pathophysiological specificity of DMDD and bipolar disorder while also demonstrating that the brain mechanisms mediating irritability may differ across diagnoses.

Among the several strands of research designed to differentiate pediatric bipolar disorder from chronic irritability, longitudinal studies provide the strongest evidence that these two phenotypes are distinct entities. Children with chronic irritability (including when it occurs in the context of oppositional defiant disorder) are at elevated risk for later depression and anxiety, but not manic episodes (12, 30, 41–45). Importantly, high levels of childhood irritability also predict increased risk for suicidality (4, 5) and functional impairment in adulthood (41, 46, 47). Such longitudinal data have heightened interest in the study of irritability, including research on clinical phenotyping, behavior, and pathophysiology. However, treatment research remains limited.

Prevalence estimates of severe irritability in community samples of children and adolescents range from 0.12% to 5% (48); 3% is the most common prevalence estimate for DMDD (46, 49) or severe mood dysregulation (43). The considerable disparity in frequency estimates results from the fact that studies vary in the extent to which they adhere to DSM-5 criteria regarding frequency of outbursts, duration of irritability, impairment, and the exclusion of individuals with mania or hypomania (4, 48). Of note, in these studies the criteria operationalizing irritability were generated post hoc. Not surprisingly, in clinically referred samples (26, 50), the rate of DMDD is higher than in epidemiological studies. To obtain valid, developmentally informed prevalence estimates of chronic, severe irritability, it is critical to assess youths prospectively over time (24, 25, 49).

**TRANSLATIONAL NEUROSCIENCE**

**CONCEPTUALIZATIONS OF IRRITABILITY**

Rewards and threats are defined as stimuli that organisms normatively approach or avoid, respectively (51). Such stimuli are intrinsic, evolutionarily adaptive motivators of behavior. Rewards and threats play a major role in current translational neuroscience conceptualizations of irritability. Neuroscientific definitions of irritability include 1) aberrant emotional and behavioral responding to frustrating nonreward (10) and 2) aberrant approach responding to threat (11, 52). These definitions have been operationalized across multiple species. In both conceptualizations, instrumental learning is a key concept. Instrumental learning is the process through which organisms learn to perform a behavior in order to obtain a reward or avoid punishment (e.g., in animal models, pressing a lever to obtain food; in human experiments, choosing a stimulus that gives the highest probability of monetary reward; in naturalistic settings, after punishment learning, a child complying with a parent’s request to stop playing a video game). As children develop, instrumental learning about rewards and threats enables them to adapt their behavior so that they respond most appropriately to environmental stimuli. Regarding irritability, different developmental phases are associated with unique symptom presentations and thresholds for psychopathology (e.g., the threshold for pathologic irritability differs between preschoolers and adolescents). Nonetheless, we hypothesize that across development, there is a shared final common pathway for the pathophysiology mediating pathologic irritability.
Aberrant Responding to Frustrative Nonreward
In rodents, Amsel (10) described frustrative nonreward as an adaptive, normative response to blocked goal attainment. Frustrative nonreward also is a negative valence construct in the National Institute of Mental Health (NIMH) Research Domain Criteria (RDoC) (53). The term describes the emotional state induced when an animal learns to expect a reward, such as food, and it is not delivered. The induction of this state elicits a range of responses (54). In rodents, the typical frustrative nonreward response consists of increased motor activity and aggression (10, 55). Evolutionarily, increased motor vigor allows for the mobilization of resources to overcome physical obstacles or other organisms and thereby acquire the reward. The utility of frustrative nonreward as the basis for translational research on irritability (53) is supported by the observation of increased motor activity following un-expected reward omission in rodents (10, 55), nonhuman primates (56), children (57), and adults (58). Frustration induction paradigms that involve deliberate, repeated blocked goal attainment can be administered cross-species.

While animal research describes normative responses to frustrative nonreward, human research suggests that higher levels of irritability are associated with a lower threshold for experiencing frustrative nonreward and a greater magnitude and duration of responding. Responses to frustration are mediated through neural circuits associated with reward processing, involving the prefrontal and anterior cingulate cortex, the striatum, and the amygdala (59–61). Brain imaging research links maturation of these circuits to improved executive functioning and emotion regulation (62–64). Such developmental processes may support the normative developmental trajectory of irritability, described above. Specifically, over the preschool to school-age period, healthy children learn to inhibit outbursts in response to frustration and to employ other, more effective strategies to achieve their goals (65). As healthy children develop more self-control and better frustration tolerance, the behavior of age-mates who persist in exhibiting significant irritability and oppositionality becomes increasingly aberrant from a developmental perspective.

Aberrant Approach Toward Threat
Threats are stimuli that signal circumstances with an increased possibility of harm for an organism. Threat responding is mediated by an established circuit involving the prefrontal cortex, the amygdala, the hypothalamus, and the periaqueductal gray. In healthy organisms, engagement of this circuit occurs in a graded fashion, based on the proximity of the organism to a threat (66). This normative behavioral cascade of threat responses involves vigilance for distal threats, freezing and other behaviors designed to avoid detection when threats are somewhat closer, and flight or attack behavior in response to proximal threats (67, 68). In this final stage, whether responding involves flight or attack depends on the context and characteristics of the threat (11, 69–71). Imminent inescapable threats elicit anger or rage (71), emotions associated with approach behavior and active engagement with the threatening stimulus in an attempt to neutralize it (72). Over development, both the stimuli deemed threatening and responses to those threats change; for example, while crying during separation from a caregiver is normative for babies, it is not for adolescents (73). Youths with irritability show a relatively low threshold for both threat detection and flight/attack (relative to flight/avoid) responding (74). Thus, while angry approach behavior may be adaptive in the context of unambiguous, inescapable, imminent threats, aberrant processing of threat stimuli (e.g., perceiving a benign stimulus as threatening, engaging in a limited range of responses) may lead to maladaptive aggressive responding.

Interactions of Frustrative Nonreward and Threat
Contexts that involve both frustrative nonreward and threat produce distinct behaviors relative to contexts that involve only frustrative nonreward or threat alone. Following frustrative nonreward, when faced with a threatening conspecific, rodents show decreased latency to attack and increased number of aggressive encounters relative to their behavior in the absence of frustrative nonreward (55). Indeed, frustrative nonreward and threat overlap conceptually; an animal who blocks another’s expected access to food both induces a state of frustrative nonreward in the deprived animal and constitutes a threat to the deprived animal’s survival. Consistent with this, rats emit ultrasonic vocalizations in the same frequency in response to both frustrative nonreward and threat (75). In humans, there is also altered processing of rewards in the context of threats, including increased expression of anger in response to frustration (76).

PATHOPHYSIOLOGICAL MODEL OF IRRITABILITY
We conceptualize irritability as aberrant responses to frustrative nonreward or threat, in the form of abnormally frequent, elevated, prolonged, and situationally inappropriate approach behaviors such as physical or verbal aggression. In the proposed model, these two neuroscience-based conceptualizations of irritability are articulated in two core pathophysiological constructs: reward-based dysfunction and threat-based dysfunction. Figure 1 depicts these two core constructs and their interaction with one another (panel C) as well as their links to the irritability phenotype (panel D).

Dysfunction in Reward Processing
Evidence suggests that irritable youths are particularly likely to experience frustrative nonreward due to aberrant reward processing. Below, we review evidence that irritable youths 1) have difficulties in instrumental learning (i.e., learning when to expect rewards and how to adjust their behavior to changing reward contingencies [see Figure 1, panel B]); 2) exhibit deficits in inhibiting responses and processing response errors; and 3) show particularly marked neural and behavioral responses to reward receipt and omission.
Deficits in reward learning. As described earlier, frustrative nonreward involves the omission of an expected reward. In both rodent and human frustrative nonreward paradigms, an initial period of instrumental reward learning establishes an expectation of reward, which is then violated when the behavior is performed but the reward is omitted. Thus, deficits in the process of reward learning could heighten susceptibility to frustrative nonreward and cause irritability (70). Indeed, youths with severe irritability exhibit deficits in learning, specifically in the initial learning of reward contingencies, and in reversal learning when reward contingencies change and behavior must be adjusted accordingly (34, 37).

There is significant cross-species preservation in the anatomic circuit mediating reward learning, which includes the prefrontal cortex, cingulate gyrus, inferior frontal gyrus, and caudate nucleus (70, 77–80) (Figure 2). Rodent, non-human primate, and human work demonstrates that the orbitofrontal cortex and ventromedial prefrontal cortex represent reward information and encode the value of actions; however, recent data raise questions as to how the orbitofrontal cortex mediates behavioral output in primates (81). The cingulate, inferior frontal gyrus, and caudate mediate the response to errors (70, 77, 78, 82, 83). Consistent with this, Adleman et al. (37) found inferior frontal gyrus and caudate dysfunction during reward learning in children with severe irritability. In addition, youths with irritability-related disorders (e.g., disruptive behavior disorders) show less orbitofrontal cortical responsiveness during reward learning relative to typically developing peers (84). Furthermore, relative to healthy youths, those with conduct problems, including oppositional defiant disorder, show reduced striatal and inferior frontal gyrus modulation as a function of the expected reward value of a stimulus (85).

One key process involved in reward learning is prediction error, that is, signaling that encodes the difference between expected outcome and received outcome (86). Animal research shows differential responses of midbrain dopamine neurons when a reward is expected but withheld and when a reward is unexpected but received (87). This differential response is thought to reflect a computation of the prediction error. Positive prediction errors occur when a reward is better than expected, whereas negative prediction errors occur when a reward is worse than expected, and these patterns are thought to manifest in associated neural circuitry. Relative to healthy youths, those with severe mood dysregulation exhibit striatal hypoactivation to the omission of an expected reward, possibly suggesting enhanced negative prediction error (88). Impaired prediction error signaling can lead to perturbed responses to frustrative nonreward because expectations of rewards are not appropriately computed (e.g., a reward continues to be expected when it should not be, based on previous nonreward outcomes) (89).

Adaptive behavior relies not only on appropriate signaling of response outcomes, but also on learning inhibitory control, or the ability to inhibit behaviors that are incompatible with task goals (62). Behavioral markers of impaired inhibitory control are associated with severe irritability and anger in children. In a toddler-aged twin sample, Gagne and Goldsmith (90) found that children with lower levels of inhibitory control displayed higher levels of anger. Furthermore, irritable (35) and aggressive (91) youths show anomalies in the neural correlates of inhibition measured using event-related
potentials, including aberrant P3 and N2 amplitudes. An increasingly large P3 reflects increasingly adaptive attention allocation (92), and an increasingly large N2 reflects increasingly strong inhibitory control (93); high trait anger has been related to decreased P3 amplitude (94) and decreased N2 activation (91). In addition, irritable youths show aberrant event-related potentials following motor errors when they do not appropriately inhibit a behavior in accordance with task demands.

In sum, research on reward learning indicates that youths with irritability show deficits in learning response associations, modifying behavior based on outcome feedback, and exhibiting inhibitory control. Both prediction error and inhibitory control are central to reward learning; whereas the neural substrates of prediction error appear to involve primarily the striatum, inhibitory control is more heavily dependent on cortical structures. Consistent with this, both cortical and subcortical circuits including the prefrontal cortex, cingulate, and striatum have been shown to be dysfunctional during reward learning in youths with severe irritability and externalizing problems (34, 37).

**Increased sensitivity to reward receipt and omission.** Increased sensitivity to reward receipt may contribute to the development and maintenance of irritability. For example, when receiving rewards, irritable school-aged children reported more positive mood than did nonirritable children, and there was increased activation in the middle frontal and anterior cingulate gyri (60). In emotionally and behaviorally dysregulated school-aged children, compared with healthy children, Bebko et al. (95) found greater middle but lower ventrolateral prefrontal cortical activity to rewards, which they interpreted as reflecting anomalous reward sensitivity. Finally, recent work in event-related potentials suggests that irritability in toddlerhood predicts enhanced neural processing of reward in preadolescence (96). Thus, heightened sensitivity to rewards may contribute to a lower threshold for frustrative nonreward in children with irritability.

Research also suggests a higher sensitivity to blocked rewards or goals (i.e., aberrant frustrative nonreward responses) in irritable youths. Indeed, irritability is a relatively tractable target for pathophysiological research because frustrative nonreward can be induced in both the clinic and the scanner. Frustration paradigms generally involve deliberate, repeated blocked goal attainment, operationalized by withholding an expected reward or by increasing task difficulty and thereby decreasing reward frequency (60, 61, 88). In such paradigms, higher trait irritability and aggression are associated with higher levels of self-reported frustration (35, 88, 97, 98). Moreover, in irritable youths, frustrative nonreward is associated with impaired attention (35, 88, 97). In healthy adults and youths, the absence of an expected reward is associated with prefrontal and striatal activity (99). Relative to healthy children, irritable children show diminished recruitment of regions mediating attention (frontal, parietal), salience (amygdala), and reward functioning (striatal, amygdala) (60, 88, 97, 98) during frustrative nonreward. For example, Deveney et al. (88) found that irritable children exhibited a decreased ability to shift their spatial attention following reward omission, along with amygdalar, striatal, parietal, and posterior cingulate dysfunction. Similarly, in school-aged irritable children, Perlman et al. (60) reported anterior cingulate cortical and striatal dysfunction during frustrative nonreward. Consistent with this, youths with externalizing problems, who often manifest irritability, show aberrant striatal activation to withheld rewards (100).

Our model suggests two other testable hypotheses. The first is that in irritable youths, a positive feedback loop occurs, such that frustration affects reward processing in ways that increase the propensity for future frustration. Specifically, we hypothesize that frustration-related affective and attentional dysfunction (88) prevents irritable youths from recognizing how their actions relate to outcomes. Thus, inappropriately expecting a reward that does not materialize causes
Hostile attribution bias. Considerable research in children has examined hostile attribution bias, defined as the tendency to interpret others’ behavior as having hostile intent. Relative to paradigms that assess attentional orienting to threat, those assessing hostile attribution bias use more complex stimuli that are shown for longer durations. For example, while attention-orienting paradigms use brief exposures of faces, paradigms assessing hostile attribution bias may use complex pictures or social vignettes. Hostile attribution bias has been associated with high levels of self-reported anger, relational aggression, and physical aggression (74, 119, 120). Hostile attribution bias is present in both anxiety (121, 122) and aggression; whereas anxious youths flee, aggressive youths attack. While minimal research examines the neural circuitry mediating hostile attribution bias, it is a complex cognitive process and hence might be expected to engage cortical mechanisms.

Face emotion processing deficits. Irritable children tend to interpret ambiguous faces as more hostile than do non-irritable children (123). Moreover, compared with healthy subjects, youths with severe irritability rate neutral faces as more fear producing and exhibit amygdalar dysfunction when making these behavioral ratings (39). In addition to these biases, irritable youths exhibit generalized deficits in labeling positively and negatively valenced face emotions. They make more errors than healthy youths and need more intense emotional information to make correct identifications (32, 33). Consistent with this, studies have shown that compared with nonirritable youths, irritable youths show poor modulation of amygdalar activity in response to

Dysfunction in Threat Processing
Fear and anger can both be adaptive responses to threat, depending on the imminence of danger (11, 69, 70, 101). Given the clinical importance of anxiety, most human studies examining pathological responses to threat have focused on exaggerated fear and avoidance behavior. However, maladaptive approach responses to threat, in the form of reactive aggression, are also clinically important (74). As noted earlier, the neurobiological circuit mediating threat processing includes the ventromedial prefrontal cortex, amygdala, hypothalamus, and periaqueductal gray (Figure 2) (11, 70). Three different types of behavioral paradigms elicit threat processing dysfunction in irritable youths. The first probes attentional orienting to threat. The second probes hostile attribution bias, that is, the tendency to interpret ambiguous or neutral social stimuli as threatening. The third type of paradigm examines explicit and implicit processing of faces displaying a range of emotions. These maladaptive forms of threat processing could lead to a lowered threshold for threat detection and aggressive responding in youths with irritability.

Increased orienting to threat. Attention biases toward threats (e.g., angry faces) occur in anxiety disorders (102). Emerging research finds similar biases in irritable youths (103). Hommer et al. (104) demonstrated that youths with severe irritability, like those with anxiety, tend to direct more attention toward threatening, angry faces than to neutral faces. This shared bias might contribute to clinical associations between irritability and anxiety (105, 106). In addition, multiple studies link trait anger to altered attention to threatening faces (107, 108), words (109–111), and images (112). Together, these studies document associations between irritability or anger and preferential allocation of attention toward threat.

Research in anxious subjects demonstrates amygdalar and prefrontal cortical hyperactivity during attention to threatening stimuli (113, 114). Studies of irritability have yet to utilize the same paradigms as in anxiety research; this is an important future direction. However, in one study probing early attentional processing of angry faces, Tseng et al. (115) found that, compared with healthy children, irritable children showed hyperactivation in regions that encode stimulus salience, such as the parahippocampal gyrus. In two small studies of individuals with intermittent explosive disorder (i.e., phasic anger and aggression), viewing angry faces was associated with greater amygdalar (116) and lower orbitofrontal cortical (117) activity relative to healthy controls. These findings are consistent with another study in which individuals with high trait anger exhibited reduced amygdalar-orbitofrontal cortical resting state connectivity relative to healthy individuals (118). Thus, several studies suggest that aberrant amygdalar and prefrontal cortical engagement may mediate irritable youths’ disproportionate attention to threat relative to healthy youths (104).
increasingly intense face emotions, specifically anger and possibly happiness (38, 40). When processing emotional faces implicitly (e.g., labeling the gender of an emotional face), children with severe irritability demonstrate amygdalar hyperactivity relative to healthy subjects (124). In addition, a recent study using an implicit face emotion processing task in youths with an anxiety disorder, DMDD, or ADHD found that the brain mechanisms mediating irritability did not differ primarily by diagnosis, although across diagnoses, the two traits of irritability and anxiety did interact to determine amygdalar-prefrontal connectivity (125). This is in contrast to another study, referenced above (40), in which the brain mechanisms mediating irritability differed by diagnosis—specifically, between DMDD and bipolar disorder. Thus, the answer to the important question regarding whether the same pathophysiology underlies irritability across psychiatric diagnoses may vary depending on the diagnoses in question.

NATURE VERSUS NURTURE: GENETIC AND ENVIRONMENTAL FACTORS

Genetic Factors
The heritability of irritability is approximately 30% – 40% (126, 127), similar to estimates for unipolar depression and anxiety (128). Twin studies suggest that genetic effects on irritability are developmentally dynamic (129), with both additive genetic and unique environmental effects (130). Of note, genetic effects may increase with development in males but decrease in females (129). Longitudinal clinical studies demonstrate associations between early irritability and later depression and anxiety (12), and there are reciprocal longitudinal relationships between maternal depression and child irritability (23, 131). Two twin-based studies suggest that the association between irritability and depression may, in part, be genetically mediated (127, 132). There have been few genetic studies of irritability, with no significant genome-wide findings. Studies should explore genetic and environmental mechanisms mediating the longitudinal and cross-sectional associations between irritability and depression.

Environmental Influences
In twin studies, 70% of the variance in irritability is explained by “non-shared” environmental factors, that is, environmental circumstances that make twins less like each other, such as adverse events that happen to one but not the other twin (127). For irritability, one important environmental factor is the content of youths’ instrumental learning from parents (see Figure 1, panel A). Irritable children often live in environments that deliver inconsistent rewards and punishments, which may unintentionally reward disruptive behavior (133–135). Inconsistent parenting behaviors have been associated with anger, aggression, and externalizing problems in children (133–135). Parenting interventions train parents to stop rewarding maladaptive behavior and start rewarding adaptive behavior (7, 8, 136–139). Results from randomized controlled trials of such interventions (140–143) demonstrate that they are among the most effective treatments for disruptive behavior disorders, such as oppositional defiant disorder. A meta-analysis of 77 studies found that a critical aspect of these treatments was increased parental consistency in the delivery of consequences for child behavior (144). Furthermore, parenting interventions have been shown to affect downstream pathophysiology, as evidenced by decreased stress reactivity and cortisol levels in youths (145).

Interplay of Genes and Environment
There is also an interplay of genetic and environmental influences on the manifestation of irritability in youths (146). Reward- and threat-related processing deficits, which may have a genetic component, evoke certain behaviors in parents. As noted above, these parental behaviors may include inconsistent parental reward and punishment of behavior. Such gene-environment correlation may further perpetuate the difficulties with irritability in youths.

TREATMENT IMPLICATIONS

This pathophysiological model suggests novel approaches to treating irritability by addressing the two domains of dysfunction described here. It remains unclear whether the pathophysiology of irritability differs transdiagnostically and in the context of co-occurring symptoms (e.g., anxiety). However, early evidence suggests that the neural correlates of irritability differ in the context of DMDD and bipolar disorder (40) and across differing levels of anxiety (125). This may have significant treatment implications.

Based on our model, we propose that treatments could target reward processing dysfunction, with a focus on correcting deficits in the content and process of instrumental learning, and on decreasing sensitivity to reward omission (138, 147). Indeed, pilot data suggest that interventions targeting aberrant responding to threat may decrease irritability (123). Pharmacological interventions for irritability, including stimulants (148), selective serotonin reuptake inhibitors (SSRIs) (149, 150), and atypical antipsychotics (151), have shown promise in the treatment of aggression and irritability. Current NIMH-funded studies (ClinicalTrials.gov identifiers: NCT00794040, NCT01714310) are examining the combination of a stimulant plus an SSRI in the treatment of severe mood dysregulation. Below, we focus on novel psychological approaches suggested by our pathophysiological model; however, such work could be extended to pharmaceutical treatments (152).

Pilot data support further testing of computer-based cognitive interventions for irritability. Based on work documenting irritable youths’ propensity to interpret faces and social situations as hostile and a treatment trial by Penton-Voak et al. (153), Stoddard et al. (123) trained irritable youths to report a more positive, less hostile interpretation of ambiguous faces. This intervention was associated with decreased irritability. A randomized trial is under way (ClinicalTrials.gov identifier: NCT02531893) that includes imaging of face emotion processing before and after treatment. Also, in anxiety
disorders, attention bias training can decrease the automatic attentional processes to threat (154, 155); whether such training would decrease irritability is unknown. Similar to computer-based interventions designed to alter subjects’ perceptions of ambiguous faces, the theory behind cognitive-behavioral therapy (CBT) for aggression and irritability evolved from social learning and social information processing models. From this perspective, irritability arises from maladaptive and hostile interpretations of social situations with peers and adults (9, 156, 157). Using learning principles and structured strategies, CBT is designed to decrease hostile interpretations of situations, provide alternative coping skills, reduce aggressive behaviors, and increase self-efficacy. Sukhodolsky and Scahill (156) developed a manualized CBT for anger and aggression in children that emphasizes social problem-solving skills; while its focus is on individual treatment with the child, there are some parent-focused sessions as well (158). Consistent with the model we describe, these investigators predict that a reduction in reactive aggression after CBT will be associated with decreased amygdala activation and increased dorsal anterior cingulate and ventromedial prefrontal cortical activation during frustration (157). In addition, Waxmonsky et al. (8) recently developed a parent and child group psychotherapy to improve irritable children’s ability to assess the potential consequences of their behavior before responding to social situations and thus select more adaptive behaviors. Early work demonstrates the feasibility of this approach.

Given threat processing anomalies in irritable youths, the pathophysiological overlap of irritability with anxiety, and the efficacy of exposure techniques in treating anxiety, another novel treatment for irritability would test whether exposure to frustrating contexts may be an effective intervention for irritability. In the treatment of anxiety disorders, exposure techniques are used to extinguish fear responses. Such techniques help patients gradually confront and tolerate stimuli that they perceive as threatening. In children and adolescents with irritability, exposure might include stimuli or situations that evoke frustration, anger, and temper outbursts. Exposure for irritability in youths has yet to be tested and will likely lead to complex clinical challenges in safely evoking anger in irritable patients. Of note, while some have questioned the rationale and efficacy of exposure techniques in the treatment of anger (159), early open pilot work in small samples of adults (160) is promising. Given the dearth of effective treatments for irritability in youths, and based on the pathophysiological model proposed here, we suggest the importance of studying a novel intervention for irritable youths that would incorporate parenting training and therapeutic exposure to anger-inducing stimuli.

CONCLUSIONS

This translational neuroscientific model synthesizes the literature relevant to irritability in youths, organizing neurobiological and behavioral findings into two overarching domains, namely, deficits in reward and in threat processing. Our model suggests how deficits in these two domains may interact to produce the clinical phenotype in youths.

First, we propose that impairments in reward processing are associated with aberrant responses to frustrating nonreward and with irritability (see Figure 1, panel C). Both neurobiological and behavioral research has shown that irritable youths have difficulties modifying their behavior in accordance with stimulus-reward associations, learning from their behavioral errors, and adapting their behavior in response to changing contingencies in order to optimize reward receipt. Irritable children’s difficulties predicting and adapting to their external environment puts them at high risk for experiencing frustrating nonreward. Moreover, rewards and reward omissions themselves may be more salient to irritable children than they are to healthy peers. When frustrating nonreward does occur, irritable youths may exhibit a heightened magnitude and duration of responding.

Second, we assert that irritability is associated with heightened orientation to threat in the environment, stronger threat-based interpretation of ambiguous and neutral stimuli, and deficits in correctly identifying face emotions (see Figure 1, panel C). That is, youths with irritability have a lower threshold for perceiving stimuli as threatening, in addition to a lower threshold for perceiving stressful responses. These maladaptive behaviors are associated with amygdalar and prefrontal cortical dysfunction. Thus, we encourage clinicians working with highly irritable children and adolescents to explore associations between perceived threats and temper outbursts. We present evidence that deficits in threat and reward processing interact (55) (see Figure 1, panel C). For example, in animals, frustrative nonreward potentiates threat approach, and threat potentiates frustrative nonreward. Clinically, an irritable child may be particularly prone to perceiving a stimulus (e.g., a parent’s neutral facial expression) as threatening in the context of frustrative nonreward (e.g., as the parent is saying “no”), increasing the propensity to respond with a temper outburst.

Our review suggests specific research directions. First, although the measurement of irritability has improved (161), there are important areas for development. For example, the “peak-end” rule, a well-known phenomenon in clinical assessment, may bias reports of irritability. That is, children and their parents may tend to recall and rate symptoms based on their most severe (i.e., “peak”) and most recent (i.e., “end”) presentations (162), thus introducing systematic bias. Therefore, current methods for characterizing irritability may be improved through the use of ecological momentary assessment techniques, in which informants report on symptoms in real time within their naturalistic settings. Ecological momentary assessment studies could also clarify whether temper outbursts and between-outburst negative mood are indeed distinct constructs. Although the DSM-5 criteria for DMDD require both clinical presentations, the existing data as to whether they are dissociable are limited and equivocal (14, 148).
Second, the study of irritability provides tractable targets for novel behavioral and neuroimaging paradigms. Paradigms should be developed to examine the consequences of acute frustration on reward and threat processing. For example, baseline instrumental learning deficits in chronically irritable youths may be exacerbated during the acute state of frustration. Similarly, when frustrated, irritable youths may perceive ambiguous or even positive stimuli as more threatening or hostile (39, 123). Of course, while there are clear advantages to probing frustration directly in children with irritability, there are also inherent difficulties. It is challenging to develop affectively salient tasks (35, 60, 97), and it can be difficult to disentangle frustration-related deficits from various confounders (e.g., time from baseline spent on a task, fatigue). In addition, such paradigms typically require deception in order for frustration to occur, raising both ethical and logistical issues.

Finally, the conceptualization of two core pathophysiological mechanisms in irritability, reward-based and threat-based, raises intriguing questions regarding subtypes that would have treatment implications. Possibly, irritable youths could be characterized as those with relatively more prominent reward-based dysfunction versus more prominent threat-based dysfunction. Tasks designed to test interactions of threat and reward processing would also help to clarify the extent to which threat stimuli potentiate dysfunction in reward processing and vice versa. Personalized psychological treatments could address reward and/or threat processing deficits, with varying “doses” of a specific treatment component prescribed on the basis of pretreatment assessments (163). For example, youths with heightened reward (relative to threat) sensitivity may benefit from more targeted parenting interventions, whereas those with more profound threat processing deficits may find computerized attention or interpretation bias interventions (155) or exposure-based treatments more effective. Based on the current literature demonstrating the efficacy of parenting interventions (140–143), we suggest that a combination of parenting and threat-based targeted interventions will be most effective in the comprehensive treatment of severe irritability in youths. Future work is encouraged to test the hypothesized pathways in the model and novel targeted interventions. Precise phenotypic and pathophysiological characterization will ultimately lead to more effective treatments for severe and impairing irritability in youths.

Dr. Stringaris has received funding from the Wellcome Trust and the U.K. National Institutes of Health Research, funds from University College London for a joint project with Johnson & Johnson, and royalties from Cambridge University Press and Oxford University Press. The other authors report no financial relationships with commercial interests.

Received July 28, 2016; revision received Oct. 24, 2016; accepted Oct. 31, 2016; published online Jan. 20, 2017.

REFERENCES


AUTHOR AND ARTICLE INFORMATION

From the Emotion and Development Branch, NIMH, Bethesda, Md.
Address correspondence to Dr. Brotman (brotmann@mail.nih.gov).
Supported by the NIMH Intramural Research Program, conducted under NIH Clinical Study Protocols 15-M-0182 (ClinicalTrials.gov identifier: NCT02531893), 02-M-0021 (ClinicalTrials.gov identifier: NCT00006177), and 00-M-0198 (ClinicalTrials.gov identifier: NCT00006177).

The authors gratefully acknowledge R. James R. Blair, Ph.D., for his thoughtful suggestions and insights on early drafts of the model.


