Irritability in Youth: A Critical Integrative Review

Ellen Leibenluft, M.D., Laura E. Allen, M.D., Robert R. Althoff, M.D., Ph.D., Melissa A. Brotman, Ph.D., Jeffrey D. Burke, Ph.D., Gabrielle A. Carlson, M.D., Daniel P. Dickstein, M.D., Lea R. Dougherty, Ph.D., Spencer C. Evans, Ph.D., Katharina Kircanski, Ph.D., Daniel N. Klein, Ph.D., Eleanor P. Malone, B.A., Carla A. Mazefsky, Ph.D., Joel Nigg, Ph.D., Susan B. Perlman, Ph.D., Daniel S. Pine, M.D., Amy Krain Roy, Ph.D., Giovanni A. Salum, M.D., Ph.D., Amy Shakeshaft, Ph.D., Jamilah Silver, M.A., Joel Stoddard M.D., MAS, Anita Thapar, M.D., Ph.D., F.R.C. Psych, Wan-Ling Tseng, Ph.D., Pablo Vidal-Ribas, Ph.D., Lauren S. Wakschlag, Ph.D., Argyris Stringaris, M.D., Ph.D., F.R.C. Psych

Author and Article Information
Primary section authors: Introduction, Conclusion, Translational Research: Leibenluft, Stringaris; Definition and Conceptualization: Leibenluft, Carlson, Stringaris; Measurement: Althoff, Evans; Epidemiology: Vidal-Ribas, Burke; DSM-5 and beyond: Salum, Pine; Early Childhood: Perlman, Wakschlag; Middle Childhood/Adolescence: Klein, Silver, Dougherty; Comorbidity with Neurodevelopmental Disorders: Mazefsky, Nigg; Treatment: Stoddard, Brotman, Dickstein, Allen; Neuroimaging: Kircanski, Roy, Tseng; Genetics: Thapar, Shakeshaft.

Corresponding author: Ellen Leibenluft, M.D. leibs@mail.nih.gov

Affiliations:
Emotion and Development Branch, National Institute of Mental Health, Bethesda MD (Leibenluft, Brotman, Kircanski, Malone, Pine); Departments of Psychiatry, Pediatrics, & Psychological Science, University of Vermont, Burlington, VT (Althoff); Department of Psychological Sciences, University of Connecticut, Storrs, CT (Burke); Division of Child and Adolescent Psychiatry, Department of Psychiatry, Stony Brook University, Stony Brook, NY (Carlson); Division of Child and Adolescent Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA (Dickstein); Department of Psychology, University of Maryland, College Park, MD (Dougherty); Department of Psychology, University of Miami, Coral Gables, FL (Evans); Departments of Psychology and Psychiatry, Stony Brook University, Stony Brook, NY (Klein); Departments of Psychology, Psychiatry, and Clinical and Translational Science, University of Pittsburgh, Pittsburgh PA (Mazefsky); Department of Psychiatry, Oregon Health and Science University, Portland OR (Nigg); Department of Psychology, Washington University at St. Louis, St. Louis, MO (Perlman); Department of Psychology, Fordham University, Bronx, NY (Roy); Child Mind Institute, New York, NY (Salum); Division of Psychological Medicine and Clinical Neurosciences, Wolfson Centre for Young People’s Mental Health, Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, Wales, United Kingdom (Shakeshaft, Thapar);
Disclosures:
Dr. Althoff is a founder and receives royalties from WISER Systems, LLC. He receives honoraria from the MGH Psychiatry Academy and the Journal of the American Academy of Child and Adolescent Psychiatry.
Dr. Roy receives book royalties from Oxford University Press, Springer Publishing
Prof Stringaris receives royalties from Oxford and Cambridge University Press and co-developed the Affective Reactivity Index (ARI), an open-source irritability measurement scale.

Funding Acknowledgment:
Dr. Leibenluft, Dr. Kircanski, Ms. Malone: NIMH Intramural Research Program (IRP) ZIAMH002786; Dr. Brotman: NIMH IRP ZIA MH002969; Drs. Carlson and Klein: NIMH R01MH069942; Dr. Dickstein: NIMH R01MH110379, NIMH R01MH111542, NIMH K24MH110402, Charles H. Hood Foundation Major Grant; Dr. Mazefsky: NICHD R01HD079512 and NIMH P50 MH130957; Dr. Nigg: NIMH R01MH1159105; Dr. Perlman: NIMH R01MH124266 and R01MH130007; Dr. Pine: NIMH IRP ZIAMH002781; Drs. Shakeshaft and Thapar: Wolfson Foundation; Dr. Silver: National Science Foundation Graduate Research Fellowship Program 16-588; Dr. Stoddard: NIMH K23MH113731, Brain and Behavioral Research Foundation, Children’s Colorado Hospital Foundation; Dr. Tseng: NIMH R00MH110570), Charles H. Hood Foundation, Fund to Retain Clinical Scientists from Yale School of Medicine and Yale Center for Clinical Investigation; Dr. Vidal-Ribas: Grant RYC2021-033369-I from MCIN/AEI/10.13039/501100011033 and the European Union «NextGenerationEU/PRTR»; Dr. Wakschlag: NIMH R01MH107652; Dr. Stringaris: NIMH IRP ZIA MH002957, University College London
Abstract

Irritability, defined as proneness to anger that may reach an impairing extent, is common in youth. There has been a recent upsurge in relevant research. We combine systematic and narrative review approaches to integrate the latest clinical and translational findings and provide suggestions to address research gaps.

Clinicians and researchers should assess irritability routinely; specific assessment tools are now available. Informant effects are prominent, stable, and vary by age and gender. The prevalence of irritability is particularly high in attention deficit hyperactivity disorder, autism spectrum disorder, and mood and anxiety disorders. Irritability is associated with impairment and suicidality risk independent of co-occurring diagnoses.

Irritability trajectories have been identified that are differentially associated with clinical outcomes; some begin early in life. Youth irritability is associated with increased risk later in life for anxiety, depression, behavioral problems, and suicidality. Irritability is moderately heritable and genetic associations differ based on age and comorbid illnesses. Parent management training is effective for constructs related to irritability, but its efficacy in irritability should be tested rigorously, as should novel mechanism-informed interventions (e.g., those targeted to frustration exposure).

Associations between irritability and suicidality and the impact of cultural context are important, under-researched topics. Large, diverse, longitudinal samples that extend into
adulthood are needed. Data from both animal and human research indicate that aberrant responses to frustration and threat are central to the pathophysiology of irritability, thus affording important translational opportunities.
Introduction

Here we present a review of the literature on irritability in youth by clinical researchers with a broad range of expertise. Why is pediatric irritability important, and why is an integrative review warranted now?

Irritability can be defined as proneness to anger that may reach an impairing extent. It is one of the most common reasons that youth present for mental health evaluation and care (1). The most salient feature of clinically relevant irritability is frequent temper outbursts that are developmentally inappropriate and out of proportion to a precipitating event (phasic irritability (2)). Such outbursts are typically accompanied by chronic grouching and angry mood (tonic irritability); for case descriptions, see (3). Severe irritability is associated with impairment at home, in school, and with peers (2). It is a central symptom of three Diagnostic and Statistical Manual for Mental Disorders (DSM-5) diagnoses (DSM-5 and beyond) as well as a transdiagnostic, dimensional construct that is especially common in attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), and anxiety and mood disorders (4–6).

Irritability is a negative valence phenotype characterized by approach behavior (7). Etiologically, it has moderate heritability (Genetics), is associated with environmental adversities including peer victimization (8) and adverse childhood experiences (9,10) and can follow several different developmental trajectories (Epidemiology). One model of irritability proposes two potential mechanisms: difficulty tolerating frustration (i.e., the emotional and behavioral response to
omission of an expected or desired reward; frustrative non-reward, (FNR) in the Research Domain Criteria (RDoC)) and aggressive responses to threat (7). This model has been elaborated to specify intervention targets (11) and can be studied in animals (Translational Research) (7).

The number of publications focused on irritability per year has quadrupled since 2000 (12). Nonetheless, irritability remains understudied relative to its clinical importance. We present the latest clinical and translational findings and identify research gaps. Depending on the content of each section of the review, we performed systematic reviews, relied on prior reviews and metanalyses, and/or used narrative methodology (Supplement: Search methods). Four sections focus on foundational topics; three on specific developmental periods or comorbidities; one on treatment; and three on methods and mechanisms. Each section concludes with prioritized suggestions for future research (also see Table 1).

**Definition and conceptualization**

Irritability has both mood (13) and behavioral components. It can be differentiated from related constructs, specifically emotion dysregulation (ED, (2)); the latter is a broader construct encompassing multiple emotional phenotypes including surgency, anxiety, mania, depression, and others (5). Questionnaire data indicate that irritability can be differentiated psychometrically from ED (14), anger and aggression (15). Irritability is also unique in that it is associated with both externalizing (16) and internalizing problems (Epidemiology; Genetics).

---

1This review focuses on literature studying irritability. However, under Comorbidity with Neurodevelopmental Conditions and Treatment, we also include literature studying ED or externalizing disorders because, in these subfields, irritability is usually not differentiated from these other constructs.
Beyond defining irritability as distinct from other constructs, it is important to test its external validity i.e., its ability to reliably provide relevant clinical information (17,18). Kendler (17) suggests assessing external validity using familial aggregation; illness course; stability; and treatment response. As detailed below, the main indicators of external validity for irritability are its moderate heritability (19–21); illness course data (i.e., longitudinal associations with anxiety, depressive disorders, behavioral problems and suicidality (5,22,23); and its ability to predict future impairment and psychopathology beyond that of comorbid disorders (4,24–26).

Irritability waxes and wanes; research on diagnostic criteria for bipolar disorder (BD) in youth demonstrates that such temporal heterogeneity is clinically important (27) (DSM-5 and beyond). There are two complementary approaches to characterizing the temporal course of irritability. The first references a lengthy time scale and differentiates episodic vs. chronic (trait) irritability. Episodic irritability is present for a delimited period. Its onset and offset are contemporaneous with changes in other characteristic symptoms (e.g., for mania, changes in sleep, activity, etc.). In contrast, chronic irritability is relatively constant.

The second subtyping approach has a shorter time scale. It differentiates acute temper outbursts (phasic irritability) vs. chronic angry mood (tonic irritability). Phasic and tonic irritability are highly correlated but may differ in comorbidities, genetic mechanisms, longitudinal course, and treatment response (28,29). Research into these associations is ongoing (Middle Childhood/Adolescence) as are studies of acute vs. sustained emotional
responses during paradigms inducing anger or frustration (30,31); these could elucidate mechanisms underlying phasic vs. tonic irritability.

Finally, there are individual differences in proneness to experiencing and exhibiting anger. For example, research suggests that inter-individual differences in the extent to which an environmental stimulus (e.g., another young person) is perceived as threatening influences angry, aggressive responses (32). However, such differences have not been explored from a first-person perspective with a deep qualitative understanding of how prior experiences and current context impact the experience of anger and subsequent behaviors. Regarding context, evidence suggests that individuals’ anger expression is causally impacted by environmental affordances i.e., when the latter support fighting, anger is more likely to be expressed (33). This supports the importance of considering cultural and social demands. Indeed, the efficacy of parent management training (PMT) in clinical constructs related to irritability (Treatment) demonstrates that these are relevant (5,34,35). The mechanisms of these different effects could be studied using real-time digital phenotyping techniques (e.g., ecological momentary assessment (EMA), actigraphy) as well as fMRI paradigms evoking frustration and anger in a social context.

The highest priority research questions regarding the conceptualization of irritability involve extending current work on the external and incremental validity of the construct and refining its definition based on these data. This effort should be cross-cultural and overlaps significantly with other research suggestions below.
Measurement

Initial advances in irritability research relied on secondary analyses of symptom data from large datasets (36). Post-hoc irritability scales were derived from structured (e.g., Development and Well-Being Assessment, DAWBA (37)) and semi-structured DSM-based interviews (e.g., Schedule for Affective Disorders and Schizophrenia for School-Age Children, K-SADS (38)), parent- and youth-report checklists (e.g., Child Behavior Checklist and Youth Self-Report, CBCL/YSR) and temperament rating scales, but these have psychometric limitations. Irritability has also been measured using items designed to assess specific diagnoses (e.g., disruptive behavior, mood, anxiety, stress disorders) (Table S1).

It is essential to have instruments that measure irritability specifically. Notable examples include ((5,39,40), Table S1): the Affective Reactivity Index (ARI); Clinician ARI; Multi-dimensional Assessment of Preschool Disruptive Behavior; Disruptive Behavior Diagnostic Observation Schedule; Emotion Dysregulation Inventory (EDI) Reactivity scale; and Emotional Outburst Inventory (EMO-I). When measuring irritability, it is important to consider “near neighbor” constructs (e.g., ED, anger, aggression) and co-occurring symptoms and disorders (e.g., depression, ADHD, traumatic stress, sleep disturbance) (5,41).

Multi-informant irritability studies show evidence of validity, reliability, and stability for both parent- and self-report; however, interrater agreement is modest and yields item- and scale-level differences (15,42). Such low agreement is common in youth psychopathology, likely
reflecting differences in perspective rather than measurement error (43). This evidence argues against averaging scores across informants or using a single informant. Researchers examining multiple informant data should parse informant variance (e.g., using bifactor or informant-specific models). Clinicians should also view informant discrepancies as potentially meaningful, explore them in detail, and consider parent and child reports as equally important sources of information (5,39). Finally, it is not uncommon for irritability to be reported by one informant, typically a parent, in the absence of corroborating reports. In such cases, the treating clinician should assess the reliability of the historian, clinical impairment and intervention options.

Inconsistent results have been reported in the neuroimaging and genetic literatures on irritability, possibly because of different assessment instruments (Neuroimaging, Genetics). Given the relatively nascent state of irritability research, it is premature to designate “gold standard” measures. However, a reasonable assessment might include an irritability-specific measure (e.g., ARI), outburst-specific measure (e.g., EMO-I), and a broad measure of psychopathology and regulation (e.g., CBCL and its dysregulation profile (44,45)). As data emerge on the reliability and external validity of different measures, the field could adopt consistent measures.

Future studies should investigate how symptom measures and their variation among informants relate to external validators, including cognitive and psychophysiological measures, longitudinal course, treatment response (5,10,15), passive sensor data, and in vivo real-time EMA from parent and child. Self-report measures differentiating tonic vs. phasic irritability, along with...
developmentally- and culturally-sensitive idiographic and qualitative tools, would enable young people’s perspectives to be better incorporated in care and research and help elucidate factors underlying observed informant effects (1). Scant literature describes how vulnerable populations (e.g., those marginalized or with neurodevelopmental conditions) experience or express irritability. Research combining qualitative and quantitative methods is key to addressing such issues of epistemic injustice.

**Epidemiology**

The frequency and intensity of normative irritable mood and temper outbursts varies developmentally, generally peaking in early childhood and declining into adulthood.

International data suggest that, from early childhood to early adolescence, 44% to 82% of individuals show low, stable levels of irritability (Table 2, (35,46–50)) while 2-5% of children show elevated and persistent or increasing irritability. Non-stable trajectories also exist; in one study, 5% of children experienced a curvilinear trajectory peaking in late childhood (48). Vulnerable populations may show greater variability in developmental trajectories.

Evidence suggests complex interactions among age, gender, and informant in rating a child’s irritability (42,48). Comparisons across studies are difficult because informants vary by subject age (e.g., parent ratings of young children vs. self, with or without parent, ratings in adolescence) and individual studies generally do not cover broad developmental periods.
Cross-sectionally, chronic irritability clusters with depression, anxiety, oppositional behavior, conduct disorder (CD), ADHD, and ASD (51). However, a meta-analysis of longitudinal studies found that irritability has the strongest links with depression (OR=1.80), anxiety (OR=1.72), and oppositional defiant disorder (ODD, OR=2.62); the latter might be inflated by item overlap (22). Irritability in youth is associated with suicidal thoughts and behaviors in community samples, both cross-sectionally (OR=1.08-4.40) and longitudinally, the latter especially if accompanied by depressive symptoms (OR= 1.15-2.22) (52).

Recent research identifies developmental trajectories of irritability. The highest priority for epidemiological research on irritability is related mechanistic studies and examinations of the continuity of irritability into adulthood. Studies should have multiple informants, including self-report in adolescents to avoid missing less overt symptoms, especially in girls. Finally, few studies have included underrepresented populations (for exceptions, see (35,53)).

**DSM-5 and beyond**

“Irritability” or “anger” appears in 20 conditions in the DSM-5, including disorders characterized by aggression. Episodic irritability occurs during manic episodes, depressive episodes in youth, withdrawal syndromes, and Premenstrual Dysphoric Disorder. Chronic irritability manifests in multiple disorders, including Generalized Anxiety Disorder and Post-Traumatic Stress Disorder; is a cardinal symptom of Intermittent Explosive Disorder (IED), ODD, and Disruptive Mood Dysregulation Disorder (DMDD); and may be a major feature of neurodevelopmental disorders (*Comorbidity with neurodevelopmental conditions*).
We discuss three questions regarding irritability in DSM-5. One concerns the boundary between DMDD and mania. In the early 2000’s, an emerging school of thought maintained that BD did not present in youth with distinct manic episodes, as in adults, but instead with non-episodic irritability and hyperarousal (54). Since the latter phenotype overlaps considerably with ADHD, there was an upsurge in the rate at which youth were assigned the diagnosis of BD (55). However, subsequent research revealed that chronic, non-episodic irritability in youth does not confer increased risk for mania later in life, indicating that the diagnosis of BD should be reserved for youth with distinct manic episodes (22,27). DMDD was included in DSM-5 to provide a diagnosis for children receiving the diagnosis of BD inappropriately. ICD-11 did not adopt DMDD, opting instead for a subtype of ODD (36). DMDD has had rapid uptake in clinical practice and was associated with reduced rates of BD diagnosis (56). However, patients with DMDD, like those with BD, have high rates of antipsychotic and polypharmacy prescriptions (56), perhaps reflecting the lack of specific, evidence-based therapeutic options.

A second question relates to boundaries among the three DSM-5 diagnoses in which chronic irritability is a core symptom i.e., IED, ODD, and DMDD. DMDD is classified as a depressive disorder, whereas ODD and IED are classified as disruptive/impulse-control disorders. Questions exist about the uniqueness of these syndromes, particularly given item overlap. Few studies compare them using external validators, including familial aggregation, outcome, and treatment response.
Third, the optimal threshold for differentiating pathological and subclinical irritability is unknown and likely to vary developmentally (57). Confusion about the threshold and the use of different assessment instruments likely contribute to the varying prevalence of DMDD across studies i.e., the reported prevalence varies 10-fold, with lower bounds below 0.5% and upper bounds above 8% (Table 3, (4,51,57–65)); notably, variation is also found in more established psychiatric disorders (66).

Within the DSM, the most important research priority is for cross-cultural epidemiological studies to derive data-driven thresholds for irritability (4) and facilitate harmonizing IED, ODD, and DMDD. Longitudinal studies could identify criteria for non-normative irritability based on outcome. External validity could be tested using family measures, physiological measures and treatment response.

Outside the DSM framework, a high priority research goal is integrating a dimensional perspective into diagnosis because the boundary between normative and pathological irritability is unlikely to be sharp. Dimensionalizing the multiple symptoms comprising, or frequently associated with, irritability (e.g., temper outbursts, reactive aggression) might help characterize the heterogeneous presentations of irritability and facilitate mechanistic studies. Dimensional approaches (i.e., Hierarchical Taxonomy of Psychopathology (HiTOP, (67), RDoC, (68)) presume that irritability-related symptoms are expressions of a shared liability or are hierarchically organized, with general psychopathology factors and lower-level domains and subfactors.
Early childhood

While all young children exhibit temper outbursts, individual differences in outbursts can be a transdiagnostic risk marker for later psychopathology. Studying such differences can aid in prevention. A recent meta-analysis found a small association between infant irritability and later internalizing and externalizing symptoms; for preschool irritability the associations were small-moderate (69). Specific markers of atypical irritability can be identified as early as 12 months of age and include dysregulation (e.g., tantrums till exhausted), occurrence in unexpected contexts (e.g., grumpy during fun activities), and high frequency (daily or more) (70). In one longitudinal study of children recruited at pediatric visits and oversampled for irritability (N=425; mean baseline age=4.7 years, mean follow-up= 2.9 years, 51% girls), low frustration tolerance and destructive tantrums at baseline were associated with impairment and DSM-5 diagnoses cross-sectionally and longitudinally (26). Repeated assessment was important because single assessments yielded frequent false positives but few false negatives (57). Secondary data analyses suggest that early irritability responds to interventions designed to promote early self-regulation e.g., (71,72).

Studying brain-behavior substrates of early irritability is challenging, but methodological advances have yielded interesting pilot data. One study using functional near-infrared spectroscopy (fNIRS) linked irritability to inter-dyadic differences in parent-child synchrony neurally and behaviorally (73); another found that irritability was inversely correlated with striatal and anterior cingulate activity during frustration (74).
The highest priority research aim would be developing tools that combine behavior patterns, neural biomarkers and parent-child interactional processes to identify which highly irritable young children will develop clinical problems and what malleable protective processes can be leveraged for prevention. Such tools could allow the integration of personalized prediction of risk into pediatric primary care (75). Improved early identification will enable the development of interventions designed to increase adaptive parent-child co-regulation.

**Middle childhood/adolescence**

Multiple community-based studies describe longitudinal developmental trajectories of irritability (*Epidemiology*, Table 1). One large study (N=4,462) followed a community sample of youth, oversampled for children of unmarried parents, from ages 3 to 15. Results showed that pre-adolescence and adolescence are important inflection points in several trajectories, and that harsh/neglectful parenting and internalizing symptoms differentiated groups (35).

Irritability in middle childhood and adolescence is often associated with co-occurring symptoms that complicate clinical care. Irritability can obscure internalizing psychopathology, raising questions about the primary treatment focus and whether treating the co-occurring problem could worsen the irritability. When irritability co-occurs with another syndrome or diagnosis, it is associated with increased impairment and a worse course (24,76).
Irritability confers risk for suicidality and non-suicidal self-injury, above major depressive disorder and other risk factors. Increased irritability often precedes suicidal behavior (46,77), particularly in the context of high irritability during middle childhood (78), and may be useful in detecting near-term risk. Mechanisms mediating associations between irritability and suicidality are unknown, although several models have been proposed (e.g., irritability as an endophenotype of suicidality, a risk factor for suicidality due to associations with other psychopathology, and/or a risk factor for the transition from suicidal ideation to attempt) (77).

In adolescence, irritability may play a role in increased mood problems and risk-taking behaviors including substance abuse. For instance, irritability mediates the association between ADHD and alcohol use problems (79), and temper outbursts (i.e., phasic irritability) predict substance use and risky sexual behaviors (29).

The highest-priority research questions revolve around suicidality. It is important to test whether interventions targeting irritability reduce suicidality and other clinical problems e.g., substance abuse, and to study mechanistic links between irritability and depression. Further research is warranted on longitudinal outcomes of tonic and phasic irritability (29,80), including mechanisms mediating those predictions. Given high correlations between tonic and phasic irritability, large samples will be needed. Finally, researchers should examine whether minority stressors are associated with increased youth irritability and suicidality, thus contributing to mental health disparities in sexual and gender minority youth and other marginalized populations.
Comorbidity with neurodevelopmental conditions

Youth with ADHD and ASD have high rates of irritability, often occurring in the context of more generalized ED i.e., excitability, anxiety, and lability. Over half of youth with ADHD or ASD have ED (6,16) and approximately 40% of youth with ADHD aged 7-12 experience extreme irritability; a similar proportion experience extreme surgency with more moderate irritability (16). ED, including irritability, is related to genetic liability for ADHD (16) (Genetics). Thus, ED may be a central feature of ADHD, complicating both assessment and treatment. Data suggest that variation in ED may form a basis for ADHD subtypes (16,81).

The clinical significance of irritability in ASD is illustrated by 45 clinical trials with irritability as the outcome (82). However, irritability is often defined to include self-injury and aggression, with insufficient focus on propensity to anger. There are few mechanistically focused studies on irritability in ASD. Some studies use tasks from ADHD research to examine frustration response and associated physiological arousal in ASD (83), but it is unknown whether factors influencing irritability are shared across ASD and ADHD. Irritability might stem from core symptoms such as inattention contributing to blocked goal attainment and frustration in ADHD, or violations of preference for sameness in ASD. However, given broad ED in both populations, cross-domain self-regulation difficulties (e.g., difficulties in top-down regulatory mechanisms) may play a larger role than diagnostic-specific features.
ADHD, ASD, and irritability all have an early onset (84). The highest-priority research question is whether irritability is a feature of those conditions versus an early indicator or risk factor of a secondary condition. Do alterations in the development of top-down control in early life influence both impulsivity and irritability? It is important to study the longitudinal course of irritability into adulthood in ADHD and ASD, including the impact of treatment and of the stresses of neurodevelopmental conditions.

**Treatment**

Evidence for the treatment of irritability is increasing, yet few randomized controlled trials (RCTs) use irritability as a primary or secondary outcome. Our systematic review (Supplement: Systematic review of studies on treatment for irritability, Fig. S1, Table S2) found 14 medication and/or psychological intervention studies in the past 5 years on the treatment of irritability (excluding those for ASD). However, most are at high risk of bias with either no control, or controls where unblinding is likely (Fig. S2). Here we also include information from publications prior to 2018; for these we used a narrative review approach (Supplement: Search Methods).

Parent management training (PMT) teaches parents reinforcement-based methods, reduces externalizing problems in youth younger than 13 years ($d \sim 0.5$), (34) and is scalable. Indeed, PMT’s efficacy is substantial and comes from many trials of good quality. Considering PMT’s efficacy across methods, cultures, and treatment settings (85), it should be studied explicitly in young people with irritability.
Cognitive Behavioral Therapy (CBT) shows medium effect sizes in children and adolescents with externalizing symptoms, disruptive and antisocial behavior, and anger-related problems; however, the studies were small and often included inactive comparators (86). Transdiagnostic CBT programs (e.g., MATCH, Unified Protocol) show promise (87), as do behavioral analysis and skills enhancement programs (88,89). Preliminary data support the potential efficacy of dialectical behavior therapy (82,90), interpersonal psychotherapy (83,91), and exposure-based treatments for irritability (84,92). Irritability may also be reduced as a secondary outcome in treatment whose primary target is another condition e.g., children who received exposure and response prevention CBT for their primary diagnosis of obsessive-compulsive disorder (93).

When irritability does not respond to behavioral interventions, medication augmentation is common. Risperidone and aripiprazole have an FDA indication for irritability in those with ASD. While extending these findings to those without ASD is often done, practitioners should do so with caution and ensure that non-pharmacological interventions have been maximized. Concerns regarding generalizing the ASD literature include first, the irritability measure in these trials was designed for those with developmental disabilities (e.g., including self-harm items (94)). Second, medication side-effects (e.g., rapid weight gain) can lead to unblinding and inflated effect sizes and/or cause differences from placebo due to non-specific effects e.g., sedation. Finally, antipsychotic efficacy should be balanced against the high frequency and severity of side effects (95).
Secondary data analyses show that stimulants may reduce irritability in youth with ADHD (96). Thus, in children with ADHD and irritability, ADHD medication should be optimized first (5). Antidepressants may also be useful. A small trial in youth with DMDD and ADHD found citalopram plus methylphenidate reduced irritability more than placebo plus methylphenidate (97).

Clinically, the top priority is to conduct rigorous RCTs for children with severe irritability that does not improve when the primary diagnosis has been treated. Careful assessment of both efficacy and side effects is important in both pharmacological and psychological trials. Increased mechanistic work is essential. For example, studies testing whether irritability and ED respond differently to treatment would be clinically important and informative regarding the external validity of both constructs. As another example, intervention that exposes youth to anger-inducing events (92,98) builds on work showing predictive associations between irritability and aberrant brain reorganization after frustration (30), but the relevance of this mechanistic work to treatment response has not been tested. In-session psychological mechanisms (e.g., enhancing distress tolerance, habituation, extinction) should also be studied.

Neuroimaging

Neuroimaging research on irritability has increased but faces the challenges of reproducibility and small effect sizes seen in clinical neuroscience broadly ((99), Tables S2-4). A meta-analysis of 30 task-based fMRI studies showed no spatial convergence of neural activation associated with irritability across or within neurocognitive domains including emotional reactivity, cognitive...
control, and reward processing (100). Studies of irritability and functional connectivity (FC) of brain regions during task or rest have yielded inconsistent results (101). The amygdala has been the focus of the largest number of FC studies, exhibiting mixed results. These inconsistent findings may reflect methodological variability and/or heterogeneity in associations between irritability and brain function. The literature on structural MRI is much smaller but has also shown inconsistent findings in grey or white matter volume or microstructure (100).

A few studies of irritability in preschoolers have used event-related potential (ERP) metrics of executive functioning, reward, or error processing (102). Further work should test constructs and ERP indices more consistently across development. A few fNIRS investigations of frustration and cognitive control find that preschoolers with irritability have lower prefrontal activity, suggesting more difficulty regulating frustration (103) (*Early Childhood*).

In sum, the neuroimaging literature on irritability is small, young, and inconclusive. Cross-sectional studies have not yielded neurobiological markers. Methodological adaptations could yield more promising results. First, neuroimaging researchers should employ developmentally-sensitive, multi-informant, and valid measures of irritability (*Measurement*). Second, existing cross-sectional studies can inform longitudinal studies of neural mechanisms underlying intra-individual changes in irritability due to development or treatment. Third, for fMRI studies, task designs that evoke relevant affective states (i.e., frustration) have increased ecological validity and evoke stronger emotional responses, and multivariate measures identified using machine learning are more reliable than those identified using univariate or single-region approaches,
improving rigor and reproducibility. For example, a recent pilot study using pre- and post-frustration resting-state scans found associations between irritability and post-frustration connectivity across limbic, reward, and sensorimotor networks (30), a finding that merits follow-up in future studies. Fourth, future work should capitalize on technological and analytic advances that integrate multiple modalities (e.g., EEG and fMRI) to probe interactions between underlying structural and functional connectivity and task-based responses at multiple timescales. Multisite collaborations will enable large sample sizes, internal and external validation, and more robust investigation of socioenvironmental factors including race/ethnicity, cultural context, and psychiatric comorbidity. Such robust investigations could identify mediators and moderators and differentiate more homogeneous subgroups. These methodological advancements could yield enhanced understanding of specific neurobiological mechanisms underlying irritability and guide development and testing of treatments with tractable neural targets (e.g., real-time fMRI neurofeedback or transcranial magnetic stimulation).

Genetics

Twin studies suggest that irritability is moderately heritable, with estimates ranging from 22% to 51% (21) (Table 4, (19–21,104,105)). Genetic contributions to irritability appear to be dynamic with stable genetic influences from childhood to adolescence and later, novel genetic and environmental influences (21). What is distinctive about genetic and environmental influences on irritability, and how do they inform our conceptualization?
First, the magnitude of heritability appears to differ across sexes, which is unusual. Heritability also increases for males and decreases for females from childhood to adolescence (21).

Second, genetically, a central question is whether irritability is best conceptualized as a) a mood problem, as in DSM-5; b) an ODD subtype, as in ICD-11; or c) closely related to ADHD and other neurodevelopmental disorders. An early twin study (19) showed that irritability predicted later depression and, unlike headstrong/hurtful ODD symptoms, showed genetic correlation with depression (0.70, 95% CI=0.59-0.82). This has been replicated, highlighting the distinction of irritable vs. headstrong/hurtful ODD symptoms. Non-shared environmental influences make a substantial contribution to irritability but shared environmental factors do not (19), while shared environment makes a substantial contribution to behavioral problems (106), again suggesting a distinction between the latter and irritability. However, a twin study of emotional lability, a construct comprised of irritability and mood volatility, showed substantial genetic overlap with ADHD (20), while other studies suggest irritability predicts and shows genetic overlap with antisocial behaviour (105). Therefore, the data are currently mixed and further research is needed.

Recent genetic investigations use composite measures of common genetic risk variants (polygenic scores, PGS). In a longitudinal, population-based study, an early-onset, persistent form of irritability was associated with ADHD PGS, male preponderance, and ADHD (48). A later, adolescent-onset increasing trajectory was associated with depression PGS, female
preponderance and adolescent depression. Also, a large cross-sectional study found that the
ADHD PGS was associated with irritability, particularly global ED in youth with ADHD (107).
These findings, together with twin studies, suggest that irritability is related to ADHD genetic
liability, and that early-onset, male irritability may be more neurodevelopmental in origin and
later-onset irritability more akin to other mood problems, although the groups overlap.

Finally, gene-environment correlations e.g., when genetically encoded child characteristics
evoke parental irritability (evocative correlation), or when a child inherits not only genes but
also the environment of irritable parents (passive correlation), are probably abundant in
developmental psychology but are difficult to study (108). However, both clinical experience and
developmental theory suggest that person-environment evocative correlations exist such that
parent-child dyads engage in mutually coercive, angry cycles (109). Indeed, PMT’s efficacy is
thought to be predicated on breaking this cyclic reinforcement of aberrant behaviors, including
irritability.

The highest priority goal would be to use longitudinal genetic designs to test whether irritability
represents the same construct by gender, across ages, and in the context of different disorders
(e.g., major depressive disorder vs ADHD). A related question is whether irritability indexes
biological /genetic heterogeneity and could be a useful stratifier of treatment and prognosis. A
third topic involves examining the genetic architecture of phasic vs. tonic irritability, given twin
evidence suggesting differences (110). Detailed clinical enquiry among diverse patient groups
and large-scale genetic data can address clinically important questions on how best to conceptualize and treat irritability.

**Translational research**

Translational models guiding work bridging preclinical and clinical research on irritability center on two constructs: frustrative non-reward (FNR) and reactive aggression (RA).

FNR was first described by Amsel in rodents and since observed in multiple other species, including humans (7,111). FNR is the normative response to frustration, or blocked goal attainment; it can be viewed as one form of negative prediction error (11). The FNR response is characterized by increased activity, increased aggression (typically measured using the resident intruder test (112)), and resistance to extinction of conditioned responses. Data suggest that aberrant behavioral and neural responses to frustration are central to pediatric irritability (30,113), supporting the clinical relevance of FNR paradigms in animals.

Rodent studies induce FNR using operant conditioning followed by omission or diminution of the expected reward. Studies of the FNR response in adult animals implicate the central amygdala, as well as dopaminergic pathways in frontal cortex and ventral striatum (114,115) (98, 99). However, these paradigms often require lengthy training and mature motor skills, making them less applicable to juvenile animals (116,117). One new paradigm addresses these limitations by capitalizing on the mouse’s tendency to explore two places alternately. This paradigm elicits the expected FNR responses in juvenile mice without affecting behavior on
tasks modelling anxiety, depression, or non-aggressive social behavior (118). Using this paradigm and C-fos expression, investigators found FNR-induced activation in multiple regions and a brain-wide shift toward a more integrated, network organization, broadly consistent with one human study (30). Future FNR studies in rodent models using techniques such as fiber photometry optogenetics (119) could guide human fMRI FNR studies.

In certain contexts, reactive aggression, (RA), or aggressive responses to threat, are adaptive (120). However, youth with irritability often have maladaptive or excessive aggressive responses to perceived threat and other negative stimuli (121–125). RA is defensive, characterized by impulsivity, and associated with emotional and physiological hyperarousal. In contrast, proactive or instrumental aggression is goal-driven, rewarding, and associated with hypoarousal (126); it is not a model for irritability. A hyperaroused RA rodent model can be created using alcohol or anabolic steroid exposure, frustration, or priming by exposure to an opponent (120). In non-human primates, the Human Intruder Paradigm can be used to model RA (127). Human data finding associations between irritability and autonomic hyperarousal at baseline and after frustration demonstrate the translational potential of such models (83,98). Human research could be extended using other arousal measures, including actigraphy and pupillometry. Rodent studies of RA implicate the amygdala, and possibly the medial PFC, findings generally consistent with human literature (120,123).

In irritability and other clinical domains, one barrier to clinically-relevant cross-species research is that clinical pathophysiological studies focus on inter-individual differences, i.e. why are some
people more irritable than others? However, such inter-individual differences remain understudied in animals; exceptions include differences in aggressive or anxiety-related threat responses (128,129). Given the central role of frustration in irritability, the highest priority studies would examine inter-individual differences in FNR in animals. Human experimental medicine studies that pair a therapeutic intervention with mechanistic approaches parallel animal studies of mean intra-individual changes in response to an experimental (vs. control) intervention; these provide another possible translational bridge. More cross-talk between clinical researchers using neuroimaging to study irritability and animal researchers interested in the topic could foster the development of novel interventions.

**Future research directions**

Irritability research has progressed. Significant advances have occurred in defining the scope and nature of the clinical problem; creating developmentally-sensitive tools that measure irritability specifically; identifying developmental trajectories; confronting diagnostic challenges; leveraging clear clinical connections between frustration and irritability into mechanistic studies; obtaining preliminary treatment data; and elucidating genetic associations.

However, given the relatively recent focus on irritability, it is unsurprising that much work remains (Table 1); five overarching research priorities emerge. The first is the need for studies in large, diverse samples of the impact of cultural and social context on the presentation of irritability, parental responses, and children’s experience. There has been related research on
the broader construct of ED (130), and international efforts are being organized to undertake research in irritability.

A second major research need concerns the external validity of the irritability construct and the measurement tools used to assess it. This requires going beyond self- and parent-report and confronting the general lack of agreement in psychology between subject report and biomarkers such as brain imaging or physiology (131). However, since self- and parent-report will always be essential to the clinical endeavor, a third research need is for greater understanding of the factors driving differences in informant reports (30,42).

A fourth major research priority encompasses understanding the transdiagnostic nature of irritability. One important and tractable question is whether treatments targeting irritability also diminish co-occurring symptoms such as ADHD, depression, or anxiety. An underlying mechanistic question regards the extent to which irritability present in different clinical contexts is mediated by similar brain mechanisms.

Finally and importantly, the evidence base for treatment needs to be expanded by rigorously conducted, well-powered RCTs that either re-purpose existing treatments or test novel interventions. More mechanistic work is needed to guide the development of novel interventions. Regarding mechanism, neuroimaging studies would benefit from the adoption of multiple methodological advances, and increased cross-talk is needed between clinical researchers and those doing relevant basic and translational research (111,114). While this lays
out an ambitious research agenda, the scope and importance of the clinical problem, coupled with recent advances laying groundwork for future studies, warrant the investment.
References


55. Moreno C, Laje G, Blanco C, Jiang H, Schmidt AB, Olfson M. National Trends in the Outpatient Diagnosis and Treatment of Bipolar Disorder in Youth. Arch Gen Psychiatry. 2007 Sep 1;64(9):1032.


69. Finlay-Jones AL, Ang JE, Brook J, Lucas JD, MacNeill LA, Mancini VO, et al. Systematic Review and Meta-analysis: Early Irritability as a Transdiagnostic Neurodevelopmental


