Development and Psychopathology (2021), **33**, 1734–1746
doi:10.1017/S0954579421000717

CAMBRIDGE UNIVERSITY PRESS

Special Issue Article

Using ecological momentary assessment to enhance irritability phenotyping in a transdiagnostic sample of youth

Reut Naim, Ashley Smith, Amanda Chue, Hannah Grassie, Julia Linke, Kelly Dombek, Shannon Shaughnessy, Cheri McNeil, Elise Cardinale, Courtney Agorsor, Sofia Cardenas, Julia Brooks, Anni R. Subar, Emily L. Jones, Quyen B. Do, Daniel S. Pine, Ellen Leibenluft, Melissa A. Brotman and Katharina Kircanski

Emotion and Development Branch, National Institute of Mental Health, Bethesda, MD, USA

Abstract

Irritability is a transdiagnostic symptom dimension in developmental psychopathology, closely related to the Research Domain Criteria (RDoC) construct of frustrative nonreward. Consistent with the RDoC framework and calls for transdiagnostic, developmentally-sensitive assessment methods, we report data from a smartphone-based, naturalistic ecological momentary assessment (EMA) study of irritability. We assessed 109 children and adolescents (M_{age} = 12.55 years; 75.20% male) encompassing several diagnostic groups – disruptive mood dysregulation disorder (DMDD), attention-deficit/hyperactivity disorder (ADHD), anxiety disorders (ANX), healthy volunteers (HV). The participants rated symptoms three times per day for 1 week. Compliance with the EMA protocol was high. As tested using multilevel modeling, EMA ratings of irritability were strongly and consistently associated with in-clinic, gold-standard measures of irritability. Further, EMA ratings of irritability were significantly related to subjective frustration during a laboratory task eliciting frustrative nonreward. Irritability levels exhibited an expected graduated pattern across diagnostic groups, and the different EMA items measuring irritability were significantly associated with one another within all groups, supporting the transdiagnostic phenomenology of irritability. Additional analyses utilized EMA ratings of anxiety as a comparison with respect to convergent validity and transdiagnostic phenomenology. The results support new measurement tools that can be used in future studies of irritability and frustrative nonreward.

Keywords: disruptive mood dysregulation disorder, ecological momentary assessment, frustrative nonreward, irritability, transdiagnostic

(Received 13 February 2021; revised 28 May 2021; accepted 5 June 2021)

Introduction

Irritability – an elevated proneness to anger relative to peers (Brotman, Kircanski, Stringaris, Pine, & Leibenluft, 2017) – is increasingly the focus of clinical and translational research (Brotman et al., 2017; Leibenluft, 2017). The inclusion of frustrative nonreward, a related construct, in the National Institute of Mental Health's Research Domain Criteria (RDoC) (Insel et al., 2010) has facilitated recent work in irritability: preliminary evidence suggests that behavioral and neural function in the context of frustrative nonreward is impaired in youth with high levels of irritability (Meyers, DeSerisy, & Roy, 2017; Tseng et al., 2019). Consistent with the RDoC principles, irritability is a transdiagnostic symptom in developmental psychopathology that spans a continuum of severity. Irritability is common in many pediatric mood, anxiety, and disruptive behavior disorders and can confer risk for multiple negative outcomes in adulthood (Vidal-Ribas, Brotman, Valdivieso, Leibenluft, & Stringaris, 2016). Severe, impairing irritability is the defining symptom of disruptive mood dysregulation disorder (DMDD) (American Psychiatric Association, 2013) and predicts later unipolar depression, anxiety, suicidality, and functional impairment (Brotman et al., 2006; Copeland, Angold, Costello, & Egger, 2013; Orri, Perret, Turecki, & Geoffroy, 2018; Stringaris, Vidal-Ribas, Brotman, & Leibenluft, 2018; Vidal-Ribas et al., 2016).

Notwithstanding research progress on the brain and behavioral mechanisms of irritability, important gaps in the measurement and phenotyping of irritability remain and may impact further advancements. The emergence of new technologies assessing irritability symptoms naturalistically in real time allows investigators to answer phenomenological questions and to link laboratory measures and clinical interventions with youth's real-world functioning. This is particularly important for pediatric irritability as studies suggest that irritable youth may retrospectively underreport levels of symptoms relative to their parents (Pan & Yeh, 2019; Stoddard et al., 2014). In the current study we developed a naturalistic, smartphone-based measurement of pediatric irritability using ecological momentary assessment (EMA) (Myin-Germeys et al., 2009; Russell & Gajos, 2020) to answer specific questions about irritability. Here, we evaluate EMA-assessed irritability with respect to feasibility, convergent validity with other levels of analysis, and transdiagnostic phenomenology in a clinical sample.

Author for Correspondence: Reut Naim, National Institute of Mental Health, Bldg. 15K, MSC 2670, Bethesda, MD 20892-2670, USA; E-mail: reut.naim@nih.gov

Cite this article: Naim R *et al* (2021). Using ecological momentary assessment to enhance irritability phenotyping in a transdiagnostic sample of youth. *Development and Psychopathology* **33**, 1734–1746. https://doi.org/10.1017/S0954579421000717

Irritability is a multifaceted clinical construct. Many conceptualizations of irritability, including its formulation in DMDD, include both a "phasic" component of temper outbursts and a "tonic" component of irritable mood. Temper outbursts refer to acute expressions of anger, verbal aggression, and/or physical aggression, typically in response to triggering stimuli (Avenevoli, Blader, & Leibenluft, 2015). Irritable mood refers to lower-intensity but longer-lasting crankiness, grouchiness, or annoyance (Cardinale et al., 2021; Moore et al., 2019). Recent studies have compared tonic and phasic irritability; for example, Cardinale et al. (2021) explored the expression of phasic and tonic irritability in a sample of youth with varying levels of attention-deficit/hyperactivity disorder (ADHD) symptoms. The findings showed that phasic, but not tonic, irritability was significantly associated with ADHD symptoms. In another study, Moore et al. (2019) reported that phasic and tonic irritability were influenced by different, unrelated subsets of genetic variants. EMA may be well-suited to further examine phasic versus tonic manifestations by capturing this multifaceted, complex symptomatology of irritability through the use of multiple assessment items.

As noted, frustrative nonreward is a translational research construct in the RDoC framework that is closely related to irritability (Insel et al., 2010). Frustrative nonreward was originally conceptualized in animal research (Amsel, 1958), defined as a response to the omission of an expected reward involving increased motor activity and aggression (e.g., Burokas, Gutierrez-Cuesta, Martin-Garcia, & Maldonado, 2012; Deveney et al., 2013; Leibenluft, 2011). Irritability relates to frustrative nonreward as irritable symptoms are often triggered by blocked goal attainment (Brotman et al., 2017; Kircanski et al., 2019). In multiple laboratory studies, higher levels of irritability in youth have been linked to altered neural (e.g., Deveney et al., 2013; Grabell et al., 2018; Perlman et al., 2015; Tseng et al., 2019) and subjective (e.g., Deveney et al., 2013; Rich et al., 2007, 2011) responses to frustration.

To date, irritability has been quantified predominantly using retrospective questionnaires and clinician inventories (e.g., Haller, Kircanski, et al., 2020; Lindgren & Koeppl, 1987; Stringaris, Goodman, et al., 2012; Wakschlag et al., 2014). While useful, the limitations of these methods include potential memory biases and social desirability (Bradburn, Rips, & Shevell, 1987). The naturalistic, momentary nature of EMA may help to improve ecological validity and reduce memory biases and social desirability. EMA also has the ability to capture more fine-grained within-person variability over time (Russell & Gajos, 2020). For example, previous EMA studies in pediatric samples have successfully probed within-person variability across complex clinical phenomena such as reactive and proactive aggression in youth with negative emotional lability (Slaughter, Leaberry, Fogleman, & Rosen, 2020), anger following violence exposure (Odgers & Russell, 2017), and emotion regulation in youth with ADHD (Babinski & Welkie, 2020).

Previous research links EMA-based and converging laboratory findings, underscoring EMA as a compelling methodology to probe irritability in line with the RDoC initiative. In a recent study from our group, Smith et al. (2019) examined EMA ratings in relation to functional magnetic resonance imaging (fMRI) data in a pediatric anxiety sample ($n = 18$). Participants rated the subjective valence of recent interactions with peers and completed fMRI task probing error monitoring in social versus nonsocial contexts. The results indicated significant associations between EMA ratings and neural responses to social errors. Other studies have linked EMA-acquired data with laboratory-based measures of pupillary reactivity (Silk et al., 2012), attention to threat (Price et al., 2016), and neural responses to monetary reward (Flores et al., 2018). As described later in the paper, this is the first study to test associations between EMA-assessed irritability and laboratory-assessed frustrative nonreward.

In the present study, we evaluated the feasibility and validity of EMA measures of pediatric irritability. We utilized a transdiagnostic sample of youth aged 8–18 years diagnosed with primary DMDD, ADHD, or anxiety disorders (ANX), along with healthy volunteers (HV). Irritability is common in both ADHD (e.g., Mulraney et al., 2016; Shaw, Stringaris, Nigg, & Leibenluft, 2014) and ANX (e.g., Comer, Pincus, & Hofmann, 2012; Cornacchio, Crum, Coxe, Pincus, & Comer, 2016; Stoddard et al., 2014). Thus, there may be some shared mechanisms of irritability as a symptom dimension across diagnoses (Hommer et al., 2014; Kircanski et al., 2018; Stoddard et al., 2017; Tseng et al., 2019).

First, we evaluated the convergent validity of EMA reports of irritability with gold-standard questionnaires and clinician instruments. We predicted that EMA ratings of irritability symptoms would correlate significantly with gold-standard assessments, to a similar degree as has been shown for anxiety symptoms (Smith et al., 2019). Second, as an initial test of convergence across levels of analysis, we tested associations of EMA-reported irritability symptoms with frustration ratings obtained during a laboratory Stop Signal Task (SST) (Logan, Cowan, & Davis, 1984). The SST is a canonical assessment of inhibitory control (Buzzell et al., 2017; Cardinale et al., 2019; Nigg, 2017), and modifications of this task have been used to elicit frustrative nonreward (Scheinost et al., 2021; Tseng et al., 2019). We predicted that EMA ratings of irritability symptoms would correlate significantly with SST frustration ratings. Together, these analyses serve as a critical first step toward validating the current EMA as a tool for quantifying irritability symptoms, which can then be further examined in relation to behavioral, physiological, and neural circuit levels of analysis. Third, to provide naturalistic evidence for irritability as a transdiagnostic construct, we leveraged the sample composition to compare irritability symptoms across diagnostic groups. We focused on the levels of and interrelations among irritability symptoms (e.g., temper outbursts, irritable mood) across diagnoses. We hypothesized that irritability symptoms would show a graduated pattern across diagnostic groups, being highest and most strongly interrelated in participants with DMDD, followed by participants with ADHD and ANX, and lastly by HV.

Given our prior use of EMA in pediatric ANX (Smith et al., 2019) and the documented links between irritability and anxiety, we included several EMA-reported anxiety symptoms as relevant comparisons with which to benchmark the irritability results. Thus, parallel to the irritability analyses, we evaluated the convergent validity of EMA-assessed anxiety symptoms and levels of anxiety symptoms across diagnostic groups.

Method

Participants

To examine irritability transdiagnostically, recruitment focused on youth aged 8–18 years ($M_{age} = 12.55$ years, $SD = 2.53$ years; 75.20% male) meeting criteria for a primary diagnosis of DMDD ($n = 26$), ADHD ($n = 28$), or ANX (generalized, social, and/or separation anxiety disorder) ($n = 28$), and youth with no psychiatric diagnosis (HV) ($n = 27$). A total of 109 youth thus participated in the study. Demographic and clinical characteristics by diagnostic group are

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

provided in Table 1 (see Table S1 of the Supplementary Material for details of all psychiatric comorbidities).

Participants were recruited via direct mailings and online advertisements. Participants were evaluated for eligibility and diagnostic status by a doctoral- or master's-level clinician using the Schedule for Affective Disorders and Schizophrenia for School-Age Children – Present and Lifetime version (Kaufman et al., 1997). Primary diagnosis was based on the chief presenting complaint and clinician judgment of the most severely impairing diagnosis. Recruitment focused on youth whose irritability was chronic and not clearly related to another ongoing or episodic diagnosis (e.g., major depressive disorder, bipolar disorder). Exclusion criteria were: IQ < 70, assessed using the Wechsler Abbreviated Intelligence Scale (Wechsler, 1999); a diagnosis of posttraumatic stress disorder, schizophrenia, neurological disorder, developmental disorder, bipolar disorder, or obsessive–compulsive disorder; a current major depressive episode; or substance abuse within 3 months of participation. Participants with a primary anxiety disorder were additionally excluded if they were taking psychotropic medication, based on their simultaneous recruitment for a study of pediatric anxiety that involved treatment. HV participants were free of any current or past psychiatric disorder. Two participants were excluded from the current study due to EMA compliance below the predetermined cutoff (less than five prompts completed). For the current study, ANX and HV participants were drawn from a larger sample that completed the same EMA and were reported in Smith et al. (2019) (sample overlap: ANX n = 15; HV n = 7). ANX and HV participants were randomly selected from this larger sample with the constraint that age and distribution by sex did not differ significantly between either group and the DMDD or ADHD group. Participants and their parents provided written assent and consent, respectively. Participants were compensated and offered a monetary bonus for completing ≥ 75% of prompts. The study was approved by the National Institute of Mental Health Institutional Review Board.

Procedure

Participants and parents completed a clinical evaluation visit before being enrolled in the research protocol. Participants subsequently completed a standardized EMA training session during which a research assistant familiarized the participant with the smartphone and protocol and reviewed each EMA item by guiding the participant through a practice prompt. To enhance feasibility and compliance, for each day during the upcoming 7 days of EMA, participants preselected 60-min periods during standardized time windows within which prompts would be delivered: morning/before school (6:00–9:00 a.m.), afternoon/after school (3:00–6:00 p.m.), and evening/before bedtime (7:00–10:00 p.m.). The actual prompt times were randomized within these time periods. EMA was administrated via ReTAINE technology (http://retaine.org/). Participants used either a personal or a study-provided smartphone (iPhone 7). Access on study-provided smartphones was limited to the website that delivered the items.

Following the training session, participants were prompted three times per day for seven consecutive days. At each prompt, participants received a text message with a link to the website through which the items were delivered. Once the prompt was received, participants had 60 min to complete the assessment before it expired and was considered incomplete. At the end of the 7-day period, participants and their parents completed retrospective questionnaires assessing psychopathology symptoms, including irritability. With the aim of evaluating convergent validity of the EMA items, every attempt was made to assess symptoms via questionnaires as close as possible to the EMA period. Of note, parents of participants with DMDD and ADHD also completed EMA as part of a separate project. However, to enable direct comparisons across diagnostic groups, only youth-report EMA data were included in the current study.

Measures

EMA measures

The full EMA protocol assessed various dimensions of mood and anxiety symptoms and their situational context. The current analyses examined irritability and anxiety symptoms.

Irritability symptoms

Irritability symptoms were assessed using four items (*temper outburst*, *irritable mood*, *frustration*, and *momentary anger*), with a focus on irritability chronometry (e.g., phasic vs. tonic irritability). Aiming to capture irritability throughout the entire day, three items assessed symptoms since the previous prompt (thereby encompassing the entire day in total): (a) temper outburst – "Since the last beep, I felt really, really angry and out of control" (categorical *yes* or *no*); (b) irritable mood – "Since the last beep, aside from being really, really angry and out of control, I was feeling generally grouchy or cranky" (5-point Likert scale, 1 = *none of the time*; 5 = *the whole time*); (c) frustration – "Since the last beep, I felt frustrated" (5-point Likert scale, 1 = *not at all*; 5 = *extremely*). One item assessed irritability at the time of the prompt: (d) momentary anger – "At the time of the beep, I felt annoyed or angry" (5-point Likert scale, 1 = *not at all*; 5 = *extremely*).

To enable further evaluation of convergent validity, we also included one item querying impairment related to irritability since the previous prompt, as an important component of in-clinic assessments. This item was *irritability-related impairment* – "Since the last beep, my grouchy mood, or being angry and out of control, got me in trouble" (categorical "*with my parent*," "*at school*," and/or "*with other kids*" [multiple selections allowed], or "*none of the above*"). The total number of domains of impairment was calculated for each participant (range = 0–3).

Anxiety symptoms

For the present analyses, anxiety symptoms were indexed using three items (*anxious affect*, *anxious avoidance*, and *momentary anxiety*). Parallel to the irritability items, anxiety symptoms were assessed either since the last prompt or at the time of the prompt. Two items assessed anxiety symptoms since the previous prompt: (a) anxious affect – "Since the last beep, I felt worried or scared" (5-point Likert scale, 1 = *not at all*; 5 = *extremely*); (b) anxious avoidance – "Since the last beep, I avoided doing things because I felt worried or scared" (categorical *yes* or *no*). One item assessed anxiety at the time of the prompt: (c) momentary anxiety – "At the time of the beep, I felt worried or scared" (5-point Likert scale, 1 = *not at all*; 5 = *extremely*).

Youth- and parent-report questionnaires

Irritability symptoms

After completing the EMA protocol, irritability symptoms over the past week were assessed using Affective Reactivity Index

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

Table 1. Demographic and clinical characteristics by diagnostic group

	M (*SD*) or % (*n*)			
	DMDD (*n* = 26)	ADHD (*n* = 28)	ANX (*n* = 28)	HV (*n* = 27)
Demographics				
Age	12.01 (1.99)	12.66 (2.38)	12.87 (2.89)	12.65 (2.80)
Sex (male)	65.38% (17)	85.71% (14)	75.00% (21)	74.07% (20)
IQ[a]	113.04 (11.04)	113.48 (12.95)	115.29 (12.79)	113.68 (14.23)
Race				
Black or African American	11.54% (3)	10.71% (3)	3.57% (1)	14.81% (4)
White or Caucasian	76.92% (20)	60.71% (17)	57.14% (16)	74.07% (20)
Asian or Asian American	3.84% (1)	3.57% (1)	7.14% (2)	0.00% (0)
American Indian or Alaskan Native	0.00% (0)	3.57% (1)	3.57% (1)	0.00% (0)
Multiple races	3.84% (1)	17.85% (5)	25.00% (7)	3.70% (1)
Not reported	3.84% (1)	3.57% (1)	3.57% (1)	7.41% (2)
Ethnicity				
Not Latino or Hispanic	88.46% (23)	75.00% (21)	75.00% (21)	96.29% (26)
Latino or Hispanic	3.84% (1)	17.85% (5)	14.28% (4)	3.70% (1)
Not reported	7.69% (2)	7.14% (2)	10.71% (3)	0.00% (0)
Symptom measures				
ARI youth-report 1-week total	5.22 (3.32)	2.41 (2.67)	2.50 (2.95)	0.78 (1.40)
ARI parent-report 1-week total	7.69 (3.02)	3.93 (3.25)	2.61 (2.75)	0.41 (0.75)
CL-ARI 1-week total	36.01 (15.42)	18.87 (18.22)	–	–
SCARED youth-report total	19.93 (16.93)	17.03 (15.42)	27.25 (15.24)	8.78 (11.41)
SCARED parent-report total	18.13 (14.65)	17.80 (14.18)	29.05 (12.83)	3.08 (3.74)
PARS Total	7.12 (5.73)	3.08 (3.71)	14.25 (3.69)	1.56 (2.14)
EMA compliance				
Percentage of prompts completed	82.23 (16.85)	80.44 (16.41)	72.62 (16.93)	80.78 (16.02)

Note: ADHD = attention-deficit/hyperactivity disorder; ANX = anxiety disorder; ARI = Affective Reactivity Index; CL-ARI = Clinician Affective Reactivity Index; DMDD = disruptive mood dysregulation disorder; EMA = ecological momentary assessment; HV = healthy volunteers; PARS = Pediatric Anxiety Rating Scale; SCARED = Screen for Child Anxiety Related Emotional Disorders

[a]IQ data were missing for 12 participants

(ARI) 1-Week version completed separately by youth and parents (Stringaris, Goodman, et al., 2012). The ARI includes six items related to irritable feelings and behaviors that are computed in the total score, and one additional item assessing impairment due to irritability. Each item uses a 3-point Likert scale ranging from 0 = *not true* to 2 = *certainly true*. The ARI has demonstrated strong construct validity and reliability (DeSousa et al., 2013; Mulraney, Melvin, & Tonge, 2014; Stringaris, Goodman, et al., 2012). In the current study, internal consistency reliability was .89 for the ARI youth-report and .92 for the ARI parent-report. The ARI was completed as close as possible to the completion of EMA (number of days from completing EMA: ARI youth-report *Mdn* = 3.00, *M* = 8.14, *SD* = 14.68; ARI parent-report *Mdn* = 3.00, *M* = 7.50, *SD* = 13.70).

Anxiety symptoms

After completing the EMA protocol, anxiety symptoms were assessed using the Screen for Child Anxiety Related Emotional Disorders (SCARED) parent- and youth-report forms, which probe symptoms over the past 3 months (Birmaher et al., 1999). The SCARED includes 41 items grouped into categories of panic/somatic symptoms, generalized anxiety, social anxiety, separation anxiety, and school avoidance. Each item uses a 3-point Likert scale ranging from 0 = *not true* to 2 = *very true*. The SCARED has demonstrated strong construct validity and reliability (Birmaher et al., 1997, 1999). In the current study, internal consistency reliability was .85 for the SCARED youth-report and .84 for the SCARED parent-report. The SCARED was completed as close as possible to the completion of EMA (number of days from completing EMA: SCARED youth-report *Mdn* = 3.00, *M* = 8.17, *SD* = 14.64; SCARED parent-report *Mdn* = 3.00, *M* = 7.87, *SD* = 13.85). One participant (ADHD) was missing the SCARED youth-report and two participants (HV) were missing the SCARED parent-report.

Clinician-administered ratings

Irritability symptoms

After completing the EMA protocol, participants in the DMDD and ADHD groups were further assessed for irritability symptoms

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

over the past week using the Clinician Affective Reactivity Index (CL-ARI) (Haller, Kircanski, et al., 2020). Only the DMDD and ADHD groups completed the CL-ARI due to constraints of the research protocols for the different diagnostic groups. The CL-ARI is a 12-item semi-structured interview with both parent and child assessing irritability along three subscales: temper outbursts, irritable mood, and impairment. Items are scored on Likert scales. A total score is completed by weighting each subscale equally (transforming to a proportion of the total possible score for the subscale), averaging across the three subscales, and multiplying by 100 (score range = 0–100). The CL-ARI has demonstrated strong validity and adequate reliability (Haller, Kircanski, et al., 2020). The CL-ARI was completed as close as possible to the completion of EMA (number of days from completing EMA: *Mdn* = 2.00, *M* = 3.37, *SD* = 9.14). Three participants (ADHD) were missing the CL-ARI.

Anxiety symptoms

After completing the EMA protocol, all participants were further assessed for anxiety symptoms over the past week using the clinician-rated Pediatric Anxiety Rating Scale (PARS) (Research Units on Pediatric Psychopharmacology Anxiety Study Group, 2002). The PARS is a 50-item symptom checklist encompassing categories of physical symptoms, generalized anxiety, social anxiety, separation anxiety, and specific phobia. Each symptom is scored as present or absent. Endorsed symptoms are then rated by the clinician on dimensions of severity, frequency, avoidance, and interference using Likert scales. A total score is calculated by summing most items. The PARS has demonstrated strong psychometric properties (Research Units on Pediatric Psychopharmacology Anxiety Study Group, 2002). The PARS was completed within 3 months of the completion of EMA (number of days from completing EMA: *Mdn* = 3.00, *M* = 15.65, *SD* = 29.32). Four participants (two HV and two ADHD) were missing the PARS.

Laboratory task frustration ratings

To begin to probe the associations of EMA-assessed symptoms with laboratory-assessed frustrative nonreward, the current study utilized data from a subset of participants ($n = 79$; M_{age} = 12.60 years, *SD* = 2.43 years; 73.70% male) across the groups (DMDD $n = 16$, ADHD $n = 23$, ANX $n = 15$, HV $n = 25$, total $n = 79$) who completed a behavioral SST in the laboratory (Verbruggen, Logan, Liefooghe, & Vandierendonck, 2008) (number of days from completing EMA: *Mdn* = 40.00, *M* = 75.86, *SD* = 99.66). The SST has been used widely to assess response inhibition in youth (Alderson, Rapport, & Kofler, 2007; Lipszyc & Schachar, 2010) and, as noted, has been leveraged as one way to elicit frustration (Scheinost et al., 2021; Tseng et al., 2019). Briefly, the task includes "go" trials and "stop" signal trials. On go trials, participants are presented with a standard two-choice reaction time task (e.g., an X or O) and instructed to press a corresponding key for each target as it appears on the screen. Stop signal trials randomly occur on 25% of the total trials, in which participants are instructed to inhibit pressing a key in response to the target if they hear an auditory stop signal. Critically, the onset time of the stop signal is adjusted based on the participant's performance in order to titrate accuracy on inhibition trials as close as possible to 50% (for details, see Logan, Schachar, & Tannock, 1997). The SST consisted of five experimental blocks. After each block, participants rated their subjective level of frustration on a 9-point Likert scale (1 = *not at all*; 9 = *extremely*). Average frustration level across the blocks was then computed.

Data Analysis

Given the nested structure of the EMA data (prompts within participants), statistical analyses were performed using multilevel modeling (HLM software, version 8.0; Raudenbush, Bryk, Cheong, & Congdon, 2019). Level 1 included the within-subject repeated data and Level 2 included the between-subject variables. Level 1 continuous predictors were person-centered and Level 2 continuous predictors were grand-mean centered. Level 1 and Level 2 categorical predictors were uncentered. The models specified outcome variables as continuous with the exception of temper outbursts and anxious avoidance, which were analyzed using logistic multilevel modeling. Where noted below, standardized variables at Level 1 and Level 2 were computed to enable more direct interpretation of coefficients (Snijders & Bosker, 2012). All models included a random intercept and, where applicable, a random slope. To handle missing data, the HLM software performed listwise deletion of missing data at Level 1 when running each analysis (i.e., within-person missing data were deleted when the analysis was conducted, as opposed to when the data set was constructed). This enabled us to include all available time points for each participant for each analysis.

(a) *Convergent validity of EMA items:* A series of analyses tested the associations between each EMA item assessing irritability or anxiety and the external measure of irritability (e.g., ARI) or anxiety (e.g., SCARED), respectively.
(b) *Associations with laboratory task frustration ratings:* A series of analyses tested the associations between each EMA item and frustration during the SST.
(c) *Transdiagnostic phenomenology of symptoms:* A series of means-as-outcomes models (Raudenbush & Bryk, 2002) was conducted for each EMA item to examine group differences in levels of EMA-assessed irritability and anxiety. In addition, a series of analyses evaluated within-prompt associations between the different irritability symptoms.

All the models are described in full later in the paper. Supplementary analyses evaluating within-prompt associations between irritability and anxiety symptoms are detailed in the Supplementary Material.

Given the multiple tests conducted, the results for each series of analyses were subjected to false discovery rate (FDR) correction (Benjamini–Hochberg procedure) with the expected proportion of false positives set to $q = .05$. Thus, the results below represent raw coefficients and standard errors from the multilevel modeling and FDR-corrected *p* values. Note that the Benjamini–Hochberg method is a step-up procedure in which all corrected values below $q = .05$ are considered significant.

Results

Participant characteristics

As shown in Table 1, there were no significant differences across the diagnostic groups in age, $F(3, 106) = 0.56$, $p = .644$, IQ, $F(3, 93) = 0.15$, $p = .928$, or distribution by sex, $\chi^2(3) = 3.03$, $p = .388$, race, likelihood ratio $\chi^2(9) = 7.77$, $p = .557$, or ethnicity, $\chi^2(3) = 5.13$, $p = .162$. We examined potential associations between each

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

of the irritability- or anxiety-related EMA items and age. Table S2 of the Supplementary Material presents the full results of these models, which were nonsignificant.

As expected, symptom questionnaire scores and clinical ratings differed significantly across groups – ARI youth-report, F (3, 105) = 12.50, $p < .001$; parent-report, F (3, 105) = 35.41, $p < .001$. Post-hoc tests indicated that the DMDD group had consistently higher scores than all other groups, and the ADHD and ANX groups had consistently higher scores than the HV group (all p values ≤ .025). The CL-ARI was completed with the DMDD and ADHD groups only; the DMDD group again exhibited higher scores than the ADHD group, t (49) = 3.63, $p = .001$. For the SCARED youth-report, omnibus F (3, 104) = 7.33, $p < .001$, SCARED parent-report, omnibus F (3, 103) = 19.73, $p < .001$, and PARS, omnibus F (3, 104) = 53.36, $p < .001$, post-hoc tests indicated that the ANX group had consistently higher scores than the ADHD and HV groups, and the DMDD group had consistently higher scores than the HV group (all p values ≤ .010). Of note, the ANX group had higher scores on the SCARED parent-report and PARS than the DMDD group (both p values ≤ .002) but did not differ significantly from the DMDD group in SCARED youth-report scores ($p = .074$).

Supporting feasibility of the EMA protocol, participant compliance with the protocol was high (proportion of prompts completed: M = 78.94%, SD = 16.02%). Compliance rate did not differ significantly across groups, F (3,106) = 2.08, $p = .107$.

Convergent validity of EMA items

To evaluate the convergent validity of each EMA item, a series of models specified each EMA item as the outcome variable at Level 1 and each external measure (youth-report, parent-report, or clinician-report) as the predictor at Level 2. Variables were standardized at both levels. However, the categorical variables "temper outburst" and "anxious avoidance" were not standardized. Consistent with the work of Smith et al. (2019), the models also included the number of days between completion of EMA and the external measure as a covariate at Level 2.

$$\text{Level 1 (prompt level): EMA item rating}_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Level 2 (participant level): } \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{score on external measure}) + \gamma_{02}(\text{days between}) + u_{0j}$$

At Level 1, EMA item rating$_{ij}$ denotes the rating for participant j at prompt i, β_{0j} represents the within-person mean rating, and r_{ij} denotes the within-person random effect. At Level 2, γ_{00} denotes the average rating in the sample, γ_{01} denotes the between-person association of the score on the external measure with the mean rating, and u_{0j} denotes the between-person random effect.

Irritability symptoms

Supporting convergent validity, as shown in Table 2, the ratings on all four EMA items assessing irritability symptoms (temper outburst, irritable mood, frustration, momentary anger) were significantly associated with ARI youth-report total score, ARI parent-report total score, and CL-ARI total score (all standardized coefficients ≥ .25, all corrected p values ≤ .015). Further, ratings on the EMA item irritability-related impairment were significantly associated with scores on the impairment item on both the ARI youth- and parent-report and the impairment subscale of the CL-ARI (all standardized coefficients ≥ .13, all corrected p values ≤ .014).

Anxiety symptoms

To a similar degree, as shown in Table S3 of the Supplementary Material, ratings on all three EMA items assessing anxiety (anxious affect, anxious avoidance, momentary anxiety) were significantly associated with SCARED youth- and parent-report total scores and PARS total score (all standardized coefficients ≥ .20, all corrected p values ≤ .009).

Associations with laboratory task frustration ratings

A series of models was conducted to examine associations between EMA ratings and frustration ratings obtained while a subset of participants (n = 79) completed the SST. The Level 1 and Level 2 equations were analogous to those examining convergent validity, and continuous variables were again standardized at both levels. As shown in Table S4 of the Supplementary Material, average frustration ratings during the SST were significantly associated with mean levels of all EMA-assessed irritability symptoms, including temper outburst, irritable mood, frustration, and momentary anger (all standardized coefficients ≥ .24, all corrected p values ≤ .007). In addition, average frustration ratings were significantly associated with mean levels of all EMA-assessed anxiety symptoms, including anxious affect, anxious avoidance, and momentary anxiety (all standardized coefficients ≥ .23, all corrected p values ≤ .012).

Transdiagnostic phenomenology of symptoms

Symptom levels across diagnoses

To examine group differences in levels of EMA-assessed symptoms, a series of means-as-outcomes models (Raudenbush & Bryk, 2002) was conducted. For each EMA item, the Level 1 equation was the same as presented earlier in the paper. At Level 2, diagnostic group was entered as a categorical predictor using dummy coding (0 = *participant not in that group*; 1 = *participant in that group*):

$$\text{Level 2 (participant level): } \beta_{0j} = \gamma_{00} + \gamma_{01}(\text{DMDD}) + \gamma_{02}(\text{ADHD}) + \gamma_{03}(\text{ANX}) + u_{0j}$$

With respect to the new notation, γ_{00} denotes the mean rating in the HV group, while γ_{01}, γ_{02}, and γ_{03} denote the differences in mean rating between the HV group and the DMDD, ADHD, and ANX groups, respectively. For each EMA item, parallel Level 2 equations tested the remaining group comparisons (e.g., DMDD vs. ADHD) in which the reference group was changed. FDR-corrected p values were calculated separately for irritability symptoms and anxiety symptoms.

Irritability symptoms. The full results for irritability symptoms across diagnostic groups are presented in Table 3. In general, irritability symptom levels exhibited a graduated pattern across the

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

Table 2. Convergent validity of EMA-assessed irritability symptoms

	ARI youth-report 1-week total score	ARI parent-report 1-week total score	CL-ARI 1-week total score[a]
Temper outburst			
Coefficient (*SE*)	0.78 (0.15)	0.84 (0.18)	0.50 (0.19)
t	5.07	4.80	2.59
FDR-corrected *p* value	.004	.003	.014
Irritable mood			
Coefficient (*SE*)	0.39 (0.08)	0.29 (0.07)	0.30 (0.10)
t	5.00	3.94	3.11
FDR-corrected *p* value	.015	.002	.004
Frustration			
Coefficient (*SE*)	0.33 (0.07)	0.27 (0.07)	0.25 (0.09)
t	5.02	3.72	2.86
FDR-corrected *p* value	.002	.002	.007
Momentary anger			
Coefficient (*SE*)	0.34 (0.06)	0.30 (0.06)	0.25 (.0.08)
t	5.25	4.66	3.20
FDR-corrected *p* value	.001	.007	.003
	ARI youth-report 1-week impairment score	ARI parent-report 1-week impairment score	CL-ARI 1-week impairment score[a]
Irritability-related impairment			
Coefficient (*SE*)	0.30 (0.08)	0.22 (0.06)	0.13 (0.05)
t	3.89	3.47	2.55
FDR-corrected *p* value	.002	.005	.014

Note: ARI = Affective Reactivity Index; CL-ARI = Clinician Affective Reactivity Index; EMA = ecological momentary assessment; FDR = false discovery rate; *SE* = standard error. Results reflect models in which all continuous variables were standardized

[a]Models included the disruptive mood dysregulation disorder (DMDD) and attention-deficit/hyperactivity disorder (ADHD) groups only; three participants (ADHD) were missing the CL-ARI

groups. As expected, the DMDD group endorsed a temper outburst on a significantly greater proportion of prompts than the ANX and HV groups. However, the DMDD group did not differ significantly from the ADHD group in the frequency of temper outbursts. In turn, both the ADHD and ANX groups endorsed a temper outburst on a significantly greater proportion of prompts than the HV group. With respect to both irritable mood and frustration, the DMDD group reported significantly higher levels than all three of the other groups (ADHD, ANX, and HV). The ANX group reported significantly higher irritable mood and frustration than the HV group. No other group differences were significant. For momentary anger, again, the DMDD group reported significantly higher levels than all the other groups (ADHD, ANX, and HV). In turn, the ADHD and ANX groups reported significantly higher momentary anger than the HV group. No other group differences were significant.

Anxiety symptoms. The full results for anxiety symptoms across groups are presented in Table S5 of the Supplementary Material. Similar to irritability symptoms, anxiety symptom levels exhibited a graduated pattern across groups. As expected, the ANX group reported significantly higher levels of anxious affect and endorsed anxious avoidance on a significantly greater proportion of prompts than both the ADHD and HV groups. However, the ANX group did not differ significantly from the DMDD group on anxious affect or anxious avoidance. Consistent with this, the DMDD group reported significantly higher anxious affect than the HV group and endorsed significantly greater anxious avoidance than both the ADHD and HV groups. For momentary anxiety, the ANX group reported a significantly higher level than the HV group, and the DMDD group reported a significantly higher level than the HV group. No other group differences were significant.

Within-prompt associations between irritability symptoms

A series of analyses was conducted to examine within-prompt associations between irritability symptoms across groups. At Level 1, each EMA irritability item was entered as a person-centered, continuous predictor to test whether and how within-person fluctuations in that symptom predicted fluctuations in the other irritability symptom. Analyses were constrained such that symptoms assessed since the previous prompt were tested as predictors of other symptoms since the previous prompt (i.e., temper outburst, irritable mood, frustration) and of momentary symptoms (i.e., momentary anger). Momentary symptoms were never predictors of symptoms since the previous prompt. Further, no association was tested twice. At Level 2, diagnostic group was again entered as a categorical predictor:

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

Table 3. Irritability symptom levels across diagnostic groups

	DMDD γ_{00} (*SE*)	ADHD γ_{00} (*SE*)	ANX γ_{00} (*SE*)	HV γ_{00} (*SE*)	Significant group comparisons (FDR-corrected *p* value)
Temper outburst[a]	11.66% (2.36%)	6.06% (2.43%)	2.98% (1.08%)	0.41% (0.28%)	DMDD > ANX (.002), HV (.002); ADHD > HV (.005); ANX > HV (.020)
Irritable mood	1.90 (0.15)	1.38 (0.09)	1.47 (0.08)	1.19 (0.05)	DMDD > ADHD (.005), ANX (.017), HV (.040); ANX > HV (.006)
Frustration	2.20 (0.17)	1.46 (0.10)	1.62 (0.09)	1.34 (0.08)	DMDD > ADHD (.002), ANX (.005), HV (.002); ANX > HV (.027)
Momentary anger	1.86 (0.13)	1.45 (0.09)	1.30 (0.06)	1.15 (0.04)	DMDD > ADHD (.019), ANX (.002), HV (.002); ADHD > HV (.004); ANX > HV (.043)

Note: ADHD = attention-deficit/hyperactivity disorder; ANX = anxiety disorder; DMDD = disruptive mood dysregulation disorder; FDR = false discovery rate; HV = healthy volunteers; *SE* = standard error. All group means for all variables were significantly greater than 0 (all corrected *p* values < .05)
[a]For illustration purposes only, γ00 (*SE*) reflects multilevel modeling in which categorical outcome variable was run as continuous, thus representing the estimated % of prompts on which the item was endorsed; group comparisons were computed using logistic multilevel modeling

$$\text{Level 1 (prompt level): EMA irritability item } Y \text{ rating}_{ij} = \beta_{0j} + \beta_{1j}(\text{EMA irritability item } X \text{ rating}) + r_{ij}$$

Level 2 (participant level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{DMDD}) + \gamma_{02}(\text{ADHD}) + \gamma_{03}(\text{ANX}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{DMDD}) + \gamma_{12}(\text{ADHD}) + \gamma_{13}(\text{ANX}) + u_{1j}$$

Here, β_{1j} denotes the association between irritability item ratings rating for participant *j*. At Level 2, γ_{10} denotes the association between irritability ratings in the HV group, while γ_{11}, γ_{12}, and γ_{13} denote the differences in this association between the HV group and the DMDD, ADHD, and ANX groups, respectively. FDR-corrected *p* values were calculated separately for group means and comparisons.

Table 4 presents the full results for within-irritability associations across groups. In general, all four groups showed significant associations among irritability symptoms. There were no significant group differences in within-irritability, within-prompt associations (i.e., no difference in slopes across groups).

Supplementary analyses: within-prompt associations between irritability and anxiety

For a full description and results of these analyses, which were generally nonsignificant, see the Supplementary Material.

Discussion

The aim of the present study was to use smartphone-based EMA to measure pediatric irritability expressed naturalistically and transdiagnostically. Supporting feasibility, rates of compliance with the EMA protocol were high and similar across diagnostic groups (Babinski & Welkie, 2020; Glenn et al., 2020; Rodríguez-Blanco, Carballo, de León, & Baca-García, 2020). These compliance rates are comparable with those reported in previous similar studies. According to a recent meta-analysis that included 42 EMA studies of youth across clinical and nonclinical settings (Wen, Schneider, Stone, & Spruijt-Metz, 2017), most studies defined compliance as the proportion of prompts to which participants responded. In these studies, the weighted average compliance rate was 78.3%, with a similar average compliance across studies with clinical (76.9%) and nonclinical (79.2%) samples.

Overall, the results indicated significant, consistent correlations between youth's EMA reports of irritability symptoms in their natural environments and in-clinic measures of irritability as rated by youth, parents, and clinicians. EMA-assessed irritability symptoms were also related to laboratory task frustration ratings, documenting links between EMA and laboratory-elicited frustrative nonreward. The diagnostic groups differed in levels of irritability symptoms, as expected, while the interrelations among irritability symptoms appeared comparable across groups. Below, we discuss each of these findings in detail and propose key directions for future research.

All four EMA items assessing irritability symptoms were significantly associated with established youth-, parent-, and clinician-report measures of irritability. The strength of these associations was in the small to medium range, the same range as that found for the EMA anxiety items (see also Smith et al., 2019), and, importantly, was comparable across the different irritability items. These results suggest that the EMA items used in this study might serve as similarly strong estimates of the overall construct of irritability, while also reflecting its multifaceted nature and fluctuations over time. Convergent validity was also shown for irritability-related functional impairment, which is an important outcome measure in treatment studies. The feasibility and convergent validity of these items build upon previous EMA research in child and adolescent anxiety (e.g., Oppenheimer et al., 2020; Smith et al., 2019), ADHD (e.g., Babinski & Welkie, 2020; Slaughter et al., 2020), and mood disorders (Gershon, Kaufmann, Torous, Depp, & Ketter, 2019; Hamilton et al., 2020), providing naturalistic indices that can be used in future pathophysiological and treatment research on irritability and, as discussed below, could be integrated with real-time interventions.

Additional evidence of convergence across levels of analysis was found in the subset of transdiagnostic participants who completed the SST. Specifically, the findings documented consistent links between EMA-assessed irritability symptoms and frustration levels reported during the SST. This was shown despite the relatively long average time span between the completion of EMA and the SST (median of 40 days). Of note, however, EMA-rated anxiety symptoms were also significantly associated with SST-reported frustration. The associations of both irritability and anxiety symptoms with SST frustration ratings may reflect

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

Table 4. Within-prompt associations between irritability symptoms

	DMDD γ_{10} (*SE*)	ADHD γ_{10} (*SE*)	ANX γ_{10} (*SE*)	HV γ_{10} (*SE*)	Significant group means (FDR-corrected *p* value)
Irritable mood predictor					
Temper outburst	1.10 (0.19)	1.10 (0.31)	1.17 (0.24)	0.47 (1.74)	DMDD (.012); ADHD (.008); ANX (.006)
Frustration	0.31 (0.05)	0.25 (0.05)	0.39 (0.05)	0.29 (0.05)	DMDD (.002); ADHD (.002); ANX (.002); HV (.003)
Momentary anger	0.42 (0.10)	0.41 (0.15)	0.25 (0.10)	0.29 (0.10)	DMDD (.024); ADHD (.012); ANX (.018); HV (.003)
Frustration predictor					
Temper outburst	1.01 (0.16)	1.03 (0.20)	1.76 (0.30)	0.62 (0.18)	DMDD (.004); ADHD (.003); ANX (.003); HV (.005)
Momentary anger	0.41 (0.07)	0.46 (0.09)	0.30 (0.07)	0.22 (0.05)	DMDD (.001); ADHD (.002); ANX (.002); HV (.002)
Temper outburst predictor					
Momentary anger	0.76 (0.31)	1.01 (0.47)	0.54 (0.35)	1.77 (1.28)	DMDD (.018); ADHD (.041)

Note: ADHD = attention-deficit/hyperactivity disorder; ANX = anxiety disorder; DMDD = disruptive mood dysregulation disorder; FDR = false discovery rate; HV = healthy volunteers; *SE* = standard error. There were no significant group comparisons

a common element of clinical severity or negative affectivity that cuts across symptom dimensions. These results also may be consistent with accumulating evidence on the clinical (e.g., Cornacchio et al., 2016; Stoddard et al., 2014), genetic (Savage et al., 2015; Stringaris, Zavos, Leibenluft, Maughan, & Eley, 2012), behavioral (e.g., Hommer et al., 2014), and neural (Cardinale et al., 2019; Kircanski et al., 2018) links between irritability and anxiety. That is, extant studies suggest that there are both shared and specific correlates of irritability and anxiety. The current findings for laboratory-assessed frustration underscore the value of further examination of common versus specific mechanisms of these two symptom dimensions. Future studies should examine associations between EMA measures and neural circuit, physiological, and behavioral levels of analysis, in line with the RDoC framework.

As predicted, the findings for irritability symptom levels by primary diagnosis evidenced a graduated pattern across the clinical groups. First, the DMDD, ADHD, and ANX groups all exhibited higher levels of temper outbursts and momentary anger than youth with no psychiatric diagnosis, and the DMDD and ANX groups also exhibited higher levels of irritable mood and frustration than healthy youth. These results highlight the transdiagnostic nature of irritability in clinical populations. Second, the DMDD group exhibited higher levels of all irritability symptoms relative to both the ADHD and ANX groups, with the exception of a nonsignificant difference in temper outbursts between the DMDD and ADHD groups. These findings support the higher threshold of overall symptomatology required for DMDD, as well as the role of temper outbursts in ADHD (Cardinale et al., 2021; Gisbert et al., 2019; Karalunas, Gustafsson, Fair, Musser, & Nigg, 2019). The observed discrepancy for "phasic" (behavior) versus "tonic" (mood) irritability in the ADHD group indicated that the ADHD group reported temper outbursts as frequently as the DMDD group, but the ADHD group did not report similarly elevated levels of irritable mood as the DMDD group. This relatively stronger association between ADHD and phasic irritability is consistent with recent findings reported by Cardinale et al. (2021). Third, the ADHD and ANX groups both showed levels of irritability intermediate between DMDD and HV, replicating and extending previous reports on the significance of irritability within primary ADHD and ANX (e.g., Cornacchio et al., 2016; Haller, Stoddard, et al., 2020; Shaw et al., 2014).

Interestingly, somewhat less diagnostic specificity was suggested for EMA-assessed anxiety symptoms between the ANX and DMDD groups, which did not differ significantly in levels of any anxiety item. An equivalent pattern was observed for SCARED youth-report scores, although the ANX group did exhibit higher SCARED parent-report and PARS scores than the DMDD group. One potential source of non-specificity in youth reports might involve the influence of ANX comorbidity in the DMDD group (see Table S1 of the Supplementary Material), consistent with previous findings (Brotman et al., 2006; Hommer et al., 2014; Mulraney et al., 2016). In addition, the overall clinical threshold for DMDD is high, reflecting a severe presentation that includes comorbid symptoms more often than not (Copeland et al., 2013; Freeman, Youngstrom, Youngstrom, & Findling, 2016). Thus, results might relate to the fact that the DMDD group was, by definition, clinically severe.

While irritability symptom *levels* differed significantly across groups, the *interrelations* among irritability symptoms were similar across groups. Specifically, irritability symptoms within prompts were significantly associated with one another in all groups, with one exception in the ANX group and two exceptions in the HV group. Thus, despite divergent levels of symptoms across diagnoses, the interrelations between different facets of irritability were preserved across diagnoses, providing insight into the transdiagnostic phenomenology of irritability. Future EMA studies may usefully explore the dynamics of irritability symptoms over time, for example, exploring how reporting irritable mood or frustration at one time point may influence the occurrence of a temper outburst at a subsequent time point. Future work might also consider moderating factors such as parental interactions. For example, parental responses to temper outbursts, such as increased attention, might influence the likelihood or intensity of future outbursts. Interrogating the antecedents and consequences of irritability symptoms in youth's daily lives will be valuable in both understanding and clinically targeting these symptoms.

In addition to using EMA as a phenotyping tool, recent and ongoing efforts are using EMA in conjunction with information processing technologies to develop "just in time" interventions tailored to psychiatric symptoms reported in daily life (e.g., Bidargaddi, Schrader, Klasnja, Licinio, & Murphy, 2020; Nahum-Shani, Hekler, & Spruijt-Metz, 2015; Nahum-Shani et al., 2018). As empirically-supported measures of daily and momentary irritability symptoms, EMA items from the current

https://doi.org/10.1017/S0954579421000717 Published online by Cambridge University Press

study could be employed in future interventions for pediatric irritability, such as informing clinicians as to the clinical status of their patients to tailor interventions as needed. EMA reports could also be paired with real-time monitoring of youth's physiology and behavior to derive a more complete phenotype of irritability and assess the impact of interventions across levels of analysis (Chaspari et al., 2014; Fletcher et al., 2010). This may be of a particular value in pediatric irritability, which is underserved with respect to evidence-based psychosocial treatments (Kircanski et al., 2018). Overall, the feasibility, convergent validity, and transdiagnostic phenomenology of EMA-assessed irritability in the current study lay important groundwork for intervention efforts.

There are several limitations of the current study. First, we focused on irritability as it presents in DMDD, ADHD, and ANX. However, irritability is common in other disorders, such as major depression (Perlis et al., 2005). We did not recruit participants with major depression in the current study because we were interested in chronic, not episodic, irritability. Future studies might examine these EMA items across additional diagnoses to further assess generalizability. Second, as participants were simultaneously being recruited for multiple ongoing studies with different inclusion criteria, there were several discrepancies across groups in relation to recruitment and procedures. In particular, only the DMDD and ADHD groups completed the CL-ARI, and participants in the ANX group had to be medication-free. Third, we matched the ANX and HV groups to the DMDD and ADHD groups based on age and sex distribution. Because males were particularly predominant in the ADHD group, this matching strategy resulted in the sample being predominately male. In addition, the sample was largely composed of White non-Hispanic participants. As such, future research should increase sample diversity with respect to race and ethnicity in order to further examine generalizability. Fourth, based largely on the constraints of school and smartphone use, the youth were prompted three times per day and the majority of the EMA items assessed symptoms that had occurred since the previous prompt. While this enhanced our ability to capture irritability symptoms (relative to querying only momentary experience), this protocol may have not fully captured the intensity of symptoms at their peak. Future EMA studies might also consider event-contingent reporting. Finally, while this study utilized frustration ratings obtained during a laboratory task in an initial examination of frustrative nonreward, future studies might include other external validators such as biological and behavioral indices.

In sum, in accordance with the call for transdiagnostic, developmentally-sensitive assessment methods consistent with the RDoC framework, the present study leveraged digital technology and EMA methodology to assess pediatric irritability in daily life. Demonstrating convergent validity with established measures and laboratory-elicited frustrative nonreward, as well as transdiagnostic phenomenology across a range of pediatric conditions, the results provide tools that can be used in future research. In particular, pathophysiological and treatment studies should prioritize examinations across EMA, behavioral, and neural circuit levels of analysis in order to better understand and intervene on irritability symptoms as they naturalistically occur.

Supplementary Material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579421000717

Acknowledgments. This work was supported by the Intramural Research Program (IRP) of the National Institute of Mental Health/National Institutes of Health (NIMH/NIH), ZIAMH002781 (Pine), ZIAMH002786 (Leibenluft), ZIAMH002778 (Leibenluft), and conducted under NIH Clinical Study Protocols 01-M-0192 and M-00-M-0021 (ClinicalTrials.gov identifiers: NCT00018057 and NCT00025935).

Funding Statement. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflicts of Interest. None.

References

Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: A meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology*, *35*, 745–758. doi:10.1007/s10802-007-9131-6

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition: DSM-5*. Arlington, VA: Author.

Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, *55*, 102–119. doi:10.1037/h0043125

Avenevoli, S., Blader, J. C., & Leibenluft, E. (2015). Irritability in youth: An update. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 881–883. doi:10.1016/j.jaac.2015.08.012

Babinski, D. E., & Welkie, J. (2020). Feasibility of ecological momentary assessment of negative emotion in girls with ADHD: A pilot study. *Psychological Reports*, *123*, 1027–1043. doi:10.1177/0033294119838757

Bidargaddi, N., Schrader, G., Klasnja, P., Licinio, J., & Murphy, S. (2020). Designing m-health interventions for precision mental health support. *Translational Psychiatry*, *10*, 222. doi:10.1038/s41398-020-00895-2

Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A replication study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*, 1230–1236. doi:10.1097/00004583-199910000-00011

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 545–553. doi:10.1097/00004583-199704000-00018

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*, 157–161. doi:10.1126/science.3563494

Brotman, M. A., Kircanski, K., Stringaris, A., Pine, D. S., & Leibenluft, E. (2017). Irritability in youths: A translational model. *American Journal of Psychiatry*, *174*, 520–532. doi:10.1176/appi.ajp.2016.16070839

Brotman, M. A., Schmajuk, M., Rich, B. A., Dickstein, D. P., Guyer, A. E., Costello, E. J., … Leibenluft, E. (2006). Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. *Biological Psychiatry*, *60*, 991–997. doi:10.1016/j.biopsych.2006.08.042

Burokas, A., Gutierrez-Cuesta, J., Martin-Garcia, E., & Maldonado, R. (2012). Operant model of frustrated expected reward in mice. *Addiction Biology*, *17*, 770–782. doi:10.1111/j.1369-1600.2011.00423.x

Buzzell, G. A., Troller-Renfree, S. V., Barker, T. V., Bowman, L. C., Chronis-Tuscano, A., Henderson, H. A., … Fox, N. A. (2017). A neurobehavioral mechanism linking behaviorally inhibited temperament and later adolescent social anxiety. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*, 1097–1105. doi:10.1016/j.jaac.2017.10.007

Cardinale, E. M., Freitag, G. F., Brotman, M. A., Pine, D. S., Leibenluft, E., & Kircanski, K. (2021). Phasic versus tonic irritability: Differential associations with ADHD symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, doi:10.1016/j.jaac.2020.11.022

Cardinale, E. M., Subar, A. R., Brotman, M. A., Leibenluft, E., Kircanski, K., & Pine, D. S. (2019). Inhibitory control and emotion dysregulation: A framework for research on anxiety. *Development and Psychopathology*, *31*, 859–869. doi:10.1017/S0954579419000300

Chaspari, T., Goodwin, M., Wilder-Smith, O., Gulsrud, A., Mucchetti, C. A., Kasari, C., … Narayanan, S. (2014). A non-homogeneous Poisson process model of skin conductance responses integrated with observed regulatory behaviors for autism intervention. In *2014 IEEE international conference*

on acoustics, speech and signal processing (ICASSP). pp. 1630-1634. Piscataway, NJ: IEEE.

Comer, J. S., Pincus, D. B., & Hofmann, S. G. (2012). Generalized anxiety disorder and the proposed associated symptoms criterion change for DSM-5 in a treatment-seeking sample of anxious youth. *Depression & Anxiety*, *29*, 994–1003. doi:10.1002/da.21999

Copeland, W. E., Angold, A., Costello, E. J., & Egger, H. (2013). Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *American Journal of Psychiatry*, *170*, 173–179. doi:10.1176/appi.ajp.2012.12010132

Cornacchio, D., Crum, K. I., Coxe, S., Pincus, D. B., & Comer, J. S. (2016). Irritability and severity of anxious symptomatology among youth with anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*, 54–61. doi:10.1016/j.jaac.2015.10.007

DeSousa, D. A., Stringaris, A., Leibenluft, E., Koller, S. H., Manfro, G. G., & Salum, G. A. (2013). Cross-cultural adaptation and preliminary psychometric properties of the affective reactivity index in Brazilian youth: Implications for DSM-5 measured irritability. *Trends in Psychiatry and Psychotherapy*, *35*, 171–180. doi:10.1590/s2237-60892013000300004

Deveney, C. M., Connolly, M. E., Haring, C. T., Bones, B. L., Reynolds, R. C., Kim, P., … Leibenluft, E. (2013). Neural mechanisms of frustration in chronically irritable children. *American Journal of Psychiatry*, *170*, 1186–1194. doi:10.1176/appi.ajp.2013.12070917

Fletcher, R. R., Dobson, K., Goodwin, M. S., Eydgahi, H., Wilder-Smith, O., Fernholz, D., … Picard, R. W. (2010). Icalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity. *IEEE Transactions on Information Technology in Biomedicine*, *14*, 215–223. doi:10.1109/TITB.2009.2038692

Flores, L. E., Jr., Eckstrand, K. L., Silk, J. S., Allen, N. B., Ambrosia, M., Healey, K. L., & Forbes, E. E. (2018). Adolescents' neural response to social reward and real-world emotional closeness and positive affect. *Cognitive, Affective, & Behavioral Neuroscience*, *18*, 705–717. doi:10.3758/s13415-018-0598-0

Freeman, A. J., Youngstrom, E. A., Youngstrom, J. K., & Findling, R. L. (2016). Disruptive mood dysregulation disorder in a community mental health clinic: Prevalence, comorbidity and correlates. *Journal of Child and Adolescent Psychopharmacology*, *26*, 123–130. doi:10.1089/cap.2015.0061

Gershon, A., Kaufmann, C. N., Torous, J., Depp, C., & Ketter, T. A. (2019). Electronic ecological momentary assessment (EMA) in youth with bipolar disorder: Demographic and clinical predictors of electronic EMA adherence. *Journal of Psychiatric Research*, *116*, 14–18. doi:10.1016/j.jpsychires.2019.05.026

Gisbert, L., Vilar, L., Rovira, P., Sanchez-Mora, C., Pagerols, M., Garcia-Martinez, I., … Ribases, M. (2019). Genome-wide analysis of emotional lability in adult attention deficit hyperactivity disorder (ADHD). *European Neuropsychopharmacology*, *29*, 795–802. doi:10.1016/j.euroneuro.2019.04.004

Glenn, C. R., Kleiman, E. M., Kearns, J. C., Santee, A. C., Esposito, E. C., Conwell, Y., & Alpert-Gillis, L. J. (2020). Feasibility and acceptability of ecological momentary assessment with high-risk suicidal adolescents following acute psychiatric care. *Journal of Clinical Child & Adolescent Psychology*, 1–17. doi:10.1080/15374416.2020.1741377

Grabell, A. S., Li, Y., Barker, J. W., Wakschlag, L. S., Huppert, T. J., & Perlman, S. B. (2018). Evidence of non-linear associations between frustration-related prefrontal Cortex activation and the normal:abnormal spectrum of irritability in young children. *Journal of Abnormal Child Psychology*, *46*, 137–147. doi:10.1007/s10802-017-0286-5

Haller, S., Kircanski, K., Stringaris, A., Clayton, M., Bui, H., Agorsor, C., … Brotman, M. A. (2020). The Clinician Affective Reactivity Index: Validity and reliability of a clinician-rated assessment of irritability. *Behavior Therapy*, *51*, 283–293. doi:10.1016/j.beth.2019.10.005

Haller, S., Stoddard, J., Pagliaccio, D., Bui, H., McGillivrary, C., & Brotman, M. (2020). Computational modeling of attentional impairments in disruptive mood dysregulation and attention deficit/hyperactivity disorder. *Biological Psychiatry*, *87*, S155–S156.

Hamilton, J. L., Chand, S., Reinhardt, L., Ladouceur, C. D., Silk, J. S., Moreno, M., … Bylsma, L. M. (2020). Social media use predicts later sleep timing and greater sleep variability: An ecological momentary assessment study of youth at high and low familial risk for depression. *Journal of Adolescence*, *83*, 122–130. doi:10.1016/j.adolescence.2020.07.009

Hommer, R. E., Meyer, A., Stoddard, J., Connolly, M. E., Mogg, K., Bradley, B. P., … Brotman, M. A. (2014). Attention bias to threat faces in severe mood dysregulation. *Depression & Anxiety*, *31*, 559–565. doi:10.1002/da.22145

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., … Wang, P. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*, 748–751. doi:10.1176/appi.ajp.2010.09091379

Karalunas, S. L., Gustafsson, H. C., Fair, D., Musser, E. D., & Nigg, J. T. (2019). Do we need an irritable subtype of ADHD? Replication and extension of a promising temperament profile approach to ADHD subtyping. *Psychological Assessment*, *31*, 236–247. doi:10.1037/pas0000664

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children – Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 980–988. doi:10.1097/00004583-199707000-00021

Kircanski, K., Craske, M. G., Averbeck, B. B., Pine, D. S., Leibenluft, E., & Brotman, M. A. (2019). Exposure therapy for pediatric irritability: Theory and potential mechanisms. *Behaviour Research and Therapy*, *118*, 141–149. doi:10.1016/j.brat.2019.04.007

Kircanski, K., White, L. K., Tseng, W. L., Wiggins, J. L., Frank, H. R., Sequeira, S., … Brotman, M. A. (2018). A latent variable approach to differentiating neural mechanisms of irritability and anxiety in youth. *JAMA Psychiatry*, *75*, 631–639. doi:10.1001/jamapsychiatry.2018.0468

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *American Journal of Psychiatry*, *168*, 129–142. doi:10.1176/appi.ajp.2010.10050766

Leibenluft, E. (2017). Pediatric irritability: A systems neuroscience approach. *Trends in Cognitive Sciences*, *21*, 277–289. doi:10.1016/j.tics.2017.02.002

Lindgren, S., & Koeppl, G. (1987). Assessing child behavior problems in a medical setting: Development of the pediatric behavior scale. In R. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. *3*, pp. 57–99). Greenwich, CT: JAI Press.

Lipszyc, J., & Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the Stop Signal Task. *Journal of the International Neuropsychological Society*, *16*, 1064–1076. doi:10.1017/S1355617710000895

Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 276–291. doi:10.1037//0096-1523.10.2.276

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science*, *8*, 60–64. doi:10.1111/j.1467-9280.1997.tb00545.x

Meyers, E., DeSerisy, M., & Roy, A. K. (2017). Disruptive mood dysregulation disorder (DMDD): An RDoC perspective. *Journal of Affective Disorders*, *216*, 117–122. doi:10.1016/j.jad.2016.08.007

Moore, A. A., Lapato, D. M., Brotman, M. A., Leibenluft, E., Aggen, S. H., Hettema, J. M., … Roberson-Nay, R. (2019). Heritability, stability, and prevalence of tonic and phasic irritability as indicators of disruptive mood dysregulation disorder. *Journal of Child Psychology and Psychiatry*, *60*, 1032–1041. doi:10.1111/jcpp.13062

Mulraney, M., Melvin, G. A., & Tonge, B. J. (2014). Psychometric properties of the affective reactivity index in Australian adults and adolescents. *Psychological Assessment*, *26*, 148–155. doi:10.1037/a0034891

Mulraney, M., Schilpzand, E. J., Hazell, P., Nicholson, J. M., Anderson, V., Efron, D., … Sciberras, E. (2016). Comorbidity and correlates of disruptive mood dysregulation disorder in 6-8-year-old children with ADHD. *European Child + Adolescent Psychiatry*, *25*, 321–330. doi:10.1007/s00787-015-0738-9

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine*, *39*, 1533–1547. doi:10.1017/S0033291708004947

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive

interventions: A pragmatic framework. *Health Psychology*, *34S*, 1209–1219. doi:10.1037/hea0000306

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*, 446–462. doi:10.1007/s12160-016-9830-8

Nigg, J. T. (2017). Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *58*, 361–383. doi:10.1111/jcpp.12675

Odgers, C. L., & Russell, M. A. (2017). Violence exposure is associated with adolescents' same- and next-day mental health symptoms. *Journal of Child Psychology and Psychiatry*, *58*, 1310–1318. doi:10.1111/jcpp.12763

Oppenheimer, C. W., Silk, J. S., Lee, K. H., Dahl, R. E., Forbes, E., Ryan, N., & Ladouceur, C. D. (2020). Suicidal ideation among anxious youth: A preliminary investigation of the role of neural processing of social rejection in interaction with real world negative social experiences. *Child Psychiatry & Human Development*, *51*, 163–173. doi:10.1007/s10578-019-00920-6

Orri, M., Perret, L. C., Turecki, G., & Geoffroy, M. C. (2018). Association between irritability and suicide-related outcomes across the life-course. Systematic review of both community and clinical studies. *Journal of Affective Disorders*, *239*, 220–233. doi:10.1016/j.jad.2018.07.010

Pan, P. Y., & Yeh, C. B. (2019). Irritability and maladaptation Among children: The utility of Chinese versions of the Affective Reactivity Index and Aberrant Behavior Checklist-Irritability subscale. *Journal of Child and Adolescent Psychopharmacology*, *29*, 213–219. doi:10.1089/cap.2018.0070

Perlis, R. H., Fraguas, R., Fava, M., Trivedi, M. H., Luther, J. F., Wisniewski, S. R., & Rush, A. J. (2005). Prevalence and clinical correlates of irritability in major depressive disorder: A preliminary report from the sequenced treatment alternatives to relieve depression study. *Journal of Clinical Psychiatry*, *66*, 159–166. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15705000

Perlman, S. B., Jones, B. M., Wakschlag, L. S., Axelson, D., Birmaher, B., & Phillips, M. L. (2015). Neural substrates of child irritability in typically developing and psychiatric populations. *Developmental Cognitive Neuroscience*, *14*, 71–80. doi:10.1016/j.dcn.2015.07.003

Price, R. B., Allen, K. B., Silk, J. S., Ladouceur, C. D., Ryan, N. D., Dahl, R. E., … Siegle, G. J. (2016). Vigilance in the laboratory predicts avoidance in the real world: A dimensional analysis of neural, behavioral, and ecological momentary data in anxious youth. *Developmental Cognitive Neuroscience*, *19*, 128–136. doi:10.1016/j.dcn.2016.03.001

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2019). *HLM 8 for Windows*. Skokie, IL: Scientific Software International, Inc.

Research Units on Pediatric Psychopharmacology Anxiety Study Group. (2002). The Pediatric Anxiety Rating Scale (PARS): Development and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*, 1061–1069. doi:10.1097/00004583-200209000-00006

Rich, B. A., Carver, F. W., Holroyd, T., Rosen, H. R., Mendoza, J. K., Cornwell, B. R., … Leibenluft, E. (2011). Different neural pathways to negative affect in youth with pediatric bipolar disorder and severe mood dysregulation. *Journal of Psychiatric Research*, *45*, 1283–1294. doi:10.1016/j.jpsychires.2011.04.006

Rich, B. A., Schmajuk, M., Perez-Edgar, K. E., Fox, N. A., Pine, D. S., & Leibenluft, E. (2007). Different psychophysiological and behavioral responses elicited by frustration in pediatric bipolar disorder and severe mood dysregulation. *American Journal of Psychiatry*, *164*, 309–317. doi:10.1176/ajp.2007.164.2.309

Rodríguez-Blanco, L., Carballo, J. J., de León, S., & Baca-García, E. (2020). User profiles of electronic ecological momentary assessment in outpatient child and adolescent mental health services. *Revista de Psiquiatría y Salud Mental* (Engl Ed). May 20:S1888-9891(20)30025-2. English, Spanish. doi:10.1016/j.rpsm.2020.04.001. Epub ahead of print. PMID: 32446867.

Russell, M. A., & Gajos, J. M. (2020). Annual research review: Ecological momentary assessment studies in child psychology and psychiatry. *Journal of Child Psychology and Psychiatry*, *61*, 376–394. doi:10.1111/jcpp.13204

Savage, J., Verhulst, B., Copeland, W., Althoff, R. R., Lichtenstein, P., & Roberson-Nay, R. (2015). A genetically informed study of the longitudinal relation between irritability and anxious/depressed symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*, 377–384. doi:10.1016/j.jaac.2015.02.010

Scheinost, D., Dadashkarimi, J., Finn, E. S., Wambach, C. G., MacGillivray, C., Roule, A. L., … Tseng, W. L. (2021). Functional connectivity during frustration: A preliminary study of predictive modeling of irritability in youth. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, *46*(7), 1300–1306. doi:10.1038/s41386-020-00954-8. Retrieved from https://escholarship.org/uc/item/4ck8h8sx

Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *171*, 276–293. doi:10.1176/appi.ajp.2013.13070966

Silk, J. S., Stroud, L. R., Siegle, G. J., Dahl, R. E., Lee, K. H., & Nelson, E. E. (2012). Peer acceptance and rejection through the eyes of youth: Pupillary, eyetracking and ecological data from the chatroom interact task. *Social Cognitive and Affective Neuroscience*, *7*, 93–105. doi:10.1093/scan/nsr044

Slaughter, K. E., Leaberry, K. D., Fogleman, N. D., & Rosen, P. J. (2020). Reactive and proactive aggression in children with and without ADHD and negative emotional lability. *Social Development*, *29*, 320–338. doi:10.1111/sode.12402

Smith, A. R., Kircanski, K., Brotman, M. A., Do, Q. B., Subar, A. R., Silk, J. S., … Pine, D. S. (2019). Advancing clinical neuroscience through enhanced tools: Pediatric social anxiety as an example. *Depression & Anxiety*, *36*, 701–711. doi:10.1002/da.22937

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage.

Stoddard, J., Stringaris, A., Brotman, M. A., Montville, D., Pine, D. S., & Leibenluft, E. (2014). Irritability in child and adolescent anxiety disorders. *Depression & Anxiety*, *31*, 566–573. doi:10.1002/da.22151

Stoddard, J., Tseng, W. L., Kim, P., Chen, G., Yi, J., Donahue, L., … Leibenluft, E. (2017). Association of irritability and anxiety with the neural mechanisms of implicit face emotion processing in youths with psychopathology. *JAMA Psychiatry*, *74*, 95–103. doi:10.1001/jamapsychiatry.2016.3282

Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., & Brotman, M. A. (2012). The Affective Reactivity Index: A concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry*, *53*, 1109–1117. doi:10.1111/j.1469-7610.2012.02561.x

Stringaris, A., Vidal-Ribas, P., Brotman, M. A., & Leibenluft, E. (2018). Practitioner review: Definition, recognition, and treatment challenges of irritability in young people. *Journal of Child Psychology and Psychiatry*, *59*, 721–739. doi:10.1111/jcpp.12823

Stringaris, A., Zavos, H., Leibenluft, E., Maughan, B., & Eley, T. C. (2012). Adolescent irritability: Phenotypic associations and genetic links with depressed mood. *American Journal of Psychiatry*, *169*, 47–54. doi:10.1176/appi.ajp.2011.10101549

Tseng, W. L., Deveney, C. M., Stoddard, J., Kircanski, K., Frackman, A. E., Yi, J. Y., … Leibenluft, E. (2019). Brain mechanisms of attention orienting following frustration: Associations with irritability and age in youths. *American Journal of Psychiatry*, *176*, 67–76. doi:10.1176/appi.ajp.2018.18040491

Verbruggen, F., Logan, G. D., Liefooghe, B., & Vandierendonck, A. (2008). Short-term aftereffects of response inhibition: Repetition priming or between-trial control adjustments? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 413–426. doi:10.1037/0096-1523.34.2.413

Vidal-Ribas, P., Brotman, M. A., Valdivieso, I., Leibenluft, E., & Stringaris, A. (2016). The status of irritability in psychiatry: A conceptual and quantitative review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*, 556–570. doi:10.1016/j.jaac.2016.04.014

Wakschlag, L. S., Briggs-Gowan, M. J., Choi, S. W., Nichols, S. R., Kestler, J., Burns, J. L., … Henry, D. (2014). Advancing a multidimensional,

developmental spectrum approach to preschool disruptive behavior. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*, 82–96 e83. doi:10.1016/j.jaac.2013.10.011

Wechsler, D. (1999). *Manual for the Wechsler Abbreviated Intelligence Scale (WASI)*. San Antonio, TX: The Psychological Corporation.

Wen, C K F, Schneider, S, Stone, A A, & Spruijt-Metz, D. (2017). Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis. *J Med Internet Res*, *19*(4), e132. doi: 10.2196/jmir.6641. PMID: 28446418; PMCID: PMC5425774.