



Published in final edited form as:

*J Am Acad Child Adolesc Psychiatry*. 2018 March ; 57(3): 191–199.e2. doi:10.1016/j.jaac.2017.12.008.

## Identifying Clinically Significant Irritability in Early Childhood

**Dr. Jillian Lee Wiggins, PhD,**

San Diego State University and the San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology

**Dr. Margaret J. Briggs-Gowan, PhD,**

University of Connecticut School of Medicine, Farmington

**Dr. Ryne Estabrook, PhD,**

Feinberg School of Medicine and Institute for Innovations in Developmental Sciences, Northwestern University, Evanston, IL

**Dr. Melissa A. Brotman, PhD,**

Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Rockville, MD

**Dr. Daniel S. Pine, MD,**

Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Rockville, MD

**Dr. Ellen Leibenluft, MD, and**

Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Rockville, MD

**Dr. Lauren S. Wakschlag, PhD**

Feinberg School of Medicine and Institute for Innovations in Developmental Sciences, Northwestern University, Evanston, IL. Institute for Policy Research at Northwestern

### Abstract

**Objective**—Advances in developmentally sensitive measurement now enable differentiation of normative vs. clinically salient irritability in early childhood. Yet, clinical application of these measures is still nascent. Here, we first developed an optimized model of clinically salient irritable behaviors at preschool age. Based on this model, we next derived an empirically-based cutoff in relation to concurrent *DSM-5* irritability-related disorders (i.e., oppositional defiant disorder [ODD], disruptive mood dysregulation disorder [DMDD], other depressive disorders) and used

---

Correspondence to: Jillian Lee Wiggins, PhD, 6363 Alvarado Ct, Ste. 103 San Diego, CA 92120; jillian.wiggins@sdsu.edu. Drs. Wiggins and Estabrook served as the statistical experts for this research.

Disclosure: Drs. Wiggins, Briggs-Gowan, Estabrook, Brotman, Pine, Leibenluft, and Wakschlag report no biomedical financial interests or potential conflicts of interest.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

longitudinal models to test the predictive validity of the cutoff for impairment and irritability trajectories and later *DSM* disorders.

**Method**—Preschool children oversampled for irritability were followed over three time points into early school age (N=425; mean age at baseline=4.7 years, mean follow-up=2.9 years). Mothers reported on children’s irritability via the developmentally-validated Multidimensional Assessment of Profile of Disruptive Behavior (MAP-DB) Temper Loss scale, impairment via the Family Life Impairment Scale, and *DSM* categories via the Preschool Age Psychiatric Assessment and Kiddie Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version.

**Results**—Of 22 MAP-DB Temper Loss behaviors, two behaviors—one normative (easily frustrated), one rare, dysregulated (destructive tantrums)—uniquely related to cross-domain impairment. At baseline, these two irritability items identified diagnostic status (ODD, DMDD, other depressive disorders) with good sensitivity (70–73%) and specificity (74–83%). Also, children above the irritability cutoff at baseline exhibited more persistent irritability and impairment and greater likelihood of *DSM* disorders in early school-age.

**Conclusion**—Clinical identification of early-onset irritability can be enhanced via brief, developmentally optimized indicators. Further research to apply these findings to tiered early intervention is important.

### Keywords

irritability; child; developmental psychopathology; disruptive mood dysregulation disorder

## INTRODUCTION

Frustration, the normative response to blocked goal attainment, is typically associated with anger. However, there are substantial individual differences in the angry responses evoked by frustration, e.g., in terms of ease of elicitation, intensity, recovery, and responsiveness to social supports.<sup>1–3</sup> Atypical irritability can be defined as a tendency to respond to frustration with increased levels of anger, relative to normative patterns within a developmental period. Thus, whereas brief, mild temper tantrums represent normative misbehavior of early childhood,<sup>4</sup> tantrums and irritable mood can become clinically salient when they are pervasive, dysregulated, and occur in developmentally unexpected contexts.<sup>5–7</sup> Clinical levels of irritability are common, impairing, and predictive of mood and behavioral disorders through adulthood.<sup>8–10</sup> Irritability is also important transdiagnostically; atypical irritability is represented in over a dozen *DSM-5* psychiatric disorders, especially oppositional defiant disorder (ODD), pediatric depression, and the new disruptive mood dysregulation disorder (DMDD). However, irritability symptoms in the *DSM* are not developmentally defined, impeding differentiation of normative misbehavior from clinical problems, e.g., how to determine when “often losing temper” is of clinical concern, as 83.7% of preschoolers regularly have temper tantrums.<sup>4</sup> Efficient identification of those irritable behaviors that sensitively and specifically differentiate normal from abnormal irritability in young children (i.e., preschoolers) is a foundational step toward effective screening for the onset of irritability-related syndromes in the early phase of the clinical sequence. Prior work from our group has identified an early irritability severity continuum,<sup>6, 11</sup> yet these findings have

not been clinically applied. No study has identified specific features of irritability that relate to impairment nor derived cutoffs. Thus, our goal is to use developmentally specified features of irritability to identify preschool-age irritable behaviors, with empirically-based frequency thresholds that most sensitively and specifically identify concurrent and early school-age *DSM* irritability-related disorders and longitudinal patterns of impairment. Specifically, we (1) develop an optimized model of clinically salient irritable behaviors at preschool age by identifying features of irritability that uniquely relate to cross-domain impairment, and, based on this model, we (2) derive an empirically-based cutoff in relation to *DSM* irritability-related syndromes, and (3) test predictive validity of the cutoff for impairment and irritability trajectories from preschool to school age as well as *DSM* disorders in early school age.

Increasingly, early childhood studies draw on developmentally sensitive methods of irritability assessment that account for qualitative features and context of irritability to facilitate normal:abnormal differentiation.<sup>3, 4, 6, 12</sup> Our focus here is on the Multidimensional Assessment Profile of Disruptive Behavior (MAP-DB) Temper Loss scale, which we used to define a developmentally-based irritability spectrum in preschoolers. Variation along the Temper Loss Scale predicts mood, anxiety, and disruptive symptoms as well as subsequent impairment.<sup>6</sup> As theorized, highly dysregulated irritable behaviors (e.g. tantrum till exhausted) are indicators of severity relative to normative misbehaviors elicited at times of transition and demand (e.g., tantrum during daily routine). In addition, severe irritability differs from normative misbehavior in that the former is very frequent (e.g., daily) or occurs in developmentally unexpected contexts (e.g. “out of the blue”).<sup>4</sup> These features are ascertainable from the MAP-DB, which assesses objective, rather than subjective, frequency of behavior, with concrete anchors such as “3 times in the past week,” rather than “sometimes” or “often.”<sup>13</sup>

However, investigations designed to generate clinical applications of early irritability measurement are lacking.<sup>14</sup> A number of developmentally sensitive methods for assessing clinically salient features of irritability in young children have been created (e.g., MAP-DB; Disruptive Behavior Diagnostic Observation Schedule, DB-DOS<sup>15, 16</sup>; and Preschool Age Psychiatric Assessment, PAPA<sup>17</sup>). However, to date, no study has identified which features are essential for sensitive and specific clinical identification. Generating optimized (i.e., brief, parsimonious) methods for determining when young children’s irritability is of clinical concern and requires intervention is essential to translate this science base to clinical use. As we have noted, differentiation of irritability and related disruptive behaviors in *DSM* is developmentally imprecise because it does not provide parameters for differentiating common misbehaviors from symptoms.<sup>18</sup> Although the *DSM-5* attempts to better account for developmental considerations in criteria for ODD (i.e., specifying a higher frequency threshold relative to older children), this was not empirically derived. Further, *DSM-5* criteria for DMDD automatically exclude children younger than 6 years, despite research suggesting abnormally high levels of irritability are identifiable in early childhood and associated with impairment.<sup>6, 19, 20</sup> Empirically deriving cutoffs for irritable behaviors, particularly based on developmentally sensitive data that span the full spectrum of normal to abnormal behavior, is advantageous for reducing false positives and false negatives at the

early phase of the clinical sequence, particularly because irritability presages myriad clinical syndromes.<sup>6</sup>

The overall goal of the present study is to generate an efficient, clinically informative approach that will parsimoniously identify “when to worry” about irritability in preschool-aged children. First, of the 22 behaviors on the MAP-DB Temper Loss scale reflecting a dimensional spectrum of irritability, we identify those individual behaviors uniquely informative for cross-domain impairment. We next determine a threshold for the summed score of the frequencies of those irritability items (i.e., the clinically optimized irritability score) that balances peak sensitivity and specificity in relation to key irritability-related *DSM* disorders, i.e., ODD, DMDD (ascertained using an algorithm validated in preschool samples that relaxed the age-graded exclusion criteria and increased the frequency threshold), and other non-DMDD depressive disorders (major depression, dysthymia, depressive disorder not otherwise specified).<sup>21–23</sup> Lastly, to establish predictive validity of this clinically optimized irritability score, we compare preschoolers with and without elevated irritability at preschool age on growth of impairment and irritability into school age and likelihood of early school-age *DSM* disorders.

## METHOD

### Participants

Data were from the Multidimensional Assessment of Preschoolers Study (MAPS<sup>4</sup>), comprised of a large sample of children, aged 3–5 at baseline, recruited from the waiting rooms of five pediatric clinics in the greater Chicago area (for details see <sup>6</sup>). For the present study, we focused on the intensive subsample of the MAPS cohort assessed longitudinally over 3 time points spanning preschool to early school age (T1: mean age= 4.66 years, SD=.85; T2: mean=5.46 years, SD=.91; T3: mean=7.08 years, SD=1.1). This intensive subsample was clinically enriched due to oversampling for disruptive behavior as well as domestic violence, which increases risk for irritability and other psychopathology in children. N=425 families participated in the baseline clinical assessment at T1 (preschool); of these, data on irritability and impairment were available for 95% (n=403) at T2 and 84% (n=355) at T3 (early school age). n=316 additionally participated in a visit in the lab at early school age for diagnostic reassessment; n=5 children were missing data on key measures and were excluded from analyses, thus the sample analyzed for follow-up diagnoses at early school age was n=311. Children who did not participate in one or more follow-ups had comparable overall irritability levels (MAP-DB Temper Loss scale, see below,  $t_{423}=.639$ ,  $p=.523$ ), race/ethnicity ( $\chi^2=.013$ ,  $df=1$ ,  $p=.909$ ), maternal education level ( $\chi^2=2.768$ ,  $df=5$ ,  $p=.736$ ) to children with data at all time points, although they were more likely to be boys ( $\chi^2=5.597$ ,  $df=1$ ,  $p=.018$ ). Participant characteristics are provided in Table 1.

### Measures

**Irritability**—The Temper Loss scale of the MAP-DB<sup>6</sup> was administered at all three time points (for detailed description see <sup>6</sup>). The 22 Temper Loss items (see Supplemental Methods, available online, and <sup>4</sup>) measure mood and tantrum features of irritability and capture variation in quality, intensity, and context. Items are rated on frequency in the past

month: 0=never in the past month; 1=rarely (less than weekly); 2=some days (1–3 days per week); 3=most days (4–6 days); 4=daily; 5=multiple times per day. The dimensional Temper Loss scale comprised of these items has excellent psychometric properties and validity.<sup>4, 13</sup> Here, we examined individual irritable behavior items from this scale in relation to impairment.

The anger modulation scores for parent and examiner contexts in the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS<sup>7</sup>; obtained at T1) and the Affective Reactivity Index (ARI<sup>1</sup>; T3) were used as additional indicators of irritability (Table 1). The DB-DOS is a standardized clinical observation measure coded by trained, reliable examiners and includes observed anger modulation with the parent<sup>7</sup>. The ARI is a brief, 6-item, parent-reported survey indexing children's trait irritability, with very good psychometric properties.<sup>1</sup>

**Impairment**—Impairment was assessed with the Family Life Impairment Scale (FLIS<sup>24</sup>), administered at all three time points to assess the extent to which children's emotions and behavior interfered with functioning in three domains: the child's functioning (Child Functional Impairment, e.g., “My child RARELY plays with other children his/her age [due to child's challenging behavior/feelings],” 5 items); interactions with their family (Family Impairment, e.g., “We RARELY take him/her to visit our family or friends,” 5 items); or childcare, for the 89% currently in childcare/school (Childcare Impairment, e.g., “...we OFTEN get complaints about his/her behavior from his/her teacher, babysitter, or school,” 9 items). Items were rated on a 3-point scale (0=“not true,” 1=“somewhat true,” 2=“very true”); mean Cronbach's  $\alpha$ =.75, range=.63–.81).<sup>6</sup> Cross-domain impairment was defined as present if one or more indicators of impairment were endorsed (score = 1) in two or more FLIS domains (i.e., Child Functional, Family, and/or Childcare Impairment).

**Clinical Diagnoses**—ODD, DMDD (excluding the >6 years of age requirement and using the higher frequency threshold (i.e., “every day,” as established by other early childhood cohorts to prevent over-identification,<sup>23</sup> vs. 3x/week as in *DSM-5*)<sup>21, 25</sup> and other depressive disorder diagnoses (major depressive disorder, dysthymic disorder, or depressive disorder not otherwise specified), were ascertained at baseline using the Preschool Age Psychiatric Assessment (PAPA<sup>17</sup>), a semi-structured interview with the parent. Diagnostic algorithms excluded the impairment requirement to avoid potential circularity with Aim 1. Although our focus was on ODD, DMDD, and other depression diagnoses, we additionally assessed for other *DSM* disorders: generalized anxiety, social anxiety, conduct, and attention-deficit/hyperactivity disorders. Inter-rater reliability was monitored on 20% of all interviews and ranged from  $\kappa$ =.83 to 1.00 on the diagnoses. We additionally re-assessed diagnoses at early school age for a subset of the children who participated in the lab visit at T3 (n=311) using the Kiddie Schedule for Affective Disorders and Schizophrenia – Present and Lifetime Version (K-SADS-PL).<sup>26</sup> In addition to children who did or did not meet all criteria for each disorder, a “subclinical” category captured children who met symptom but not duration or impairment criteria for ODD or DMDD.

## Analytic Overview

**Aim 1: Develop Optimized Model of Clinically Salient Irritable Behaviors at Preschool Age**—For our first aim, we ascertained the irritability items of the Temper Loss scale of the MAP-DB that related most strongly and uniquely to presence of FLIS cross-domain impairment. As tantrums and irritable mood are normative during this period, cross-domain impairment was considered a more rigorous indicator of concern to account for developmentally expectable expressions of irritability in this age period.

We used individual behaviors from the Temper Loss scale, rather than the underlying factor score, to create the most parsimonious model of irritability features necessary to discriminate impairment. Including the individual items in this first step of the model allows us to pinpoint the most salient aspects of quality, intensity, and context of mood and tantrum features of irritability, all of which have been shown to contribute to severity. This approach further allows us to determine the minimum number of items required, helping to reduce patient burden in clinical settings. We implemented stepwise logistic regression, using forward entry with likelihood ratio comparison of models, in IBM SPSS statistical software (Chicago, IL) to ascertain the predictive value of each individual irritability item from the Temper Loss scale using the FLIS cross-domain impairment indicator. Stepwise logistic regression tests all items individually and in combination and identifies the items with the strongest relationship with impairment; this method is advantageous for variable selection when predictors may be correlated<sup>27</sup> and to determine most parsimonious models.

**Aim 2: Derive Empirically-Based Cutoff for Clinically Optimized Irritability Score for *DSM-5* Disorders Featuring Irritability (i.e., ODD, DMDD, Other Depressive Disorders)**—For our second aim, we derived a threshold for the summed score of the derived irritability items from Aim 1 (i.e., clinically optimized irritability score), i.e., the frequency cutoff for the behaviors identified in Aim 1 that balances peak sensitivity and specificity in relation to ODD, DMDD, and other non-DMDD depressive disorders (major depression, dysthymia, depressive disorder not otherwise specified), using receiver-operator-characteristic (ROC) analysis. Though the primary focus was on irritability-related syndromes, we also used any *DSM* disorder (generalized anxiety, social anxiety, conduct, attention-deficit/hyperactivity disorders in addition to DMDD, ODD, or other depressive disorders) for comparison in an ROC analysis. T1 summed score of the irritability items from Aim 1 was the criterion variable and diagnosis (present vs. absent) was the classification variable. Youden's Index values were utilized to identify the optimal cutoff score, based on the highest levels of both specificity (i.e., true negative rate) and sensitivity (i.e., true positive rate). This analysis was implemented in R with package `OptimalCutpoints`.<sup>28</sup>

Sensitivity and specificity values are considered acceptable at rates  $\geq 70\%$ .<sup>29</sup> We additionally calculated positive predictive (PPV, i.e., probability that children who met the irritability cutoff also met diagnostic criteria) and negative predictive (NPV, i.e., probability that children who did not meet irritability cutoff also did not meet diagnostic criteria) values. We are unaware of published, recommended minimum PPV or NPV, although other high-stakes



tests deem PPV as low as 10% acceptable.<sup>30</sup> Last, we tested for sociodemographic differences between children who meet vs. do not meet the irritability cutoff.

### **Aim 3: Test Predictive Validity of the Clinically Optimized Irritability Score Cutoff for Growth of Impairment and Irritability From Preschool to Early School Age and Early School Age *DSM* Disorders**

We next compared growth curves for children above or below the baseline irritability cutoff from Aim 2. To examine changes in growth of impairment across these groups, we utilized growth mixture modeling with known classes (meeting vs. not meeting the irritability cutoff at baseline). Age, centered on the mean age at the first time point (4.66 years), was entered as the time score, loading each time point by the child's actual age, to more accurately estimate trajectories. We compared children meeting vs. not meeting the irritability cutoff at baseline on their intercepts (representing mean levels of impairment at average baseline age, 4.66 years) and slopes (representing degree of change in impairment over time) in models estimating separate trajectories for FLIS Child and Family Impairment subscales. We additionally estimated single class growth curve models for comparison,<sup>31</sup> considering multiple fit indices (Bayesian Information Criterion [BIC], sample size-adjusted BIC [SSABIC], and Akaike Information Criterion [AIC]) to confirm that the two-class models fit better than the single-class models; lower values on these indices indicate relatively better model fit. To avoid local maxima, all models were run with 5000 random starts. Analyses were implemented in Mplus 8 statistical software (Muthen and Muthen, Los Angeles, CA), because Mplus uses the expectation maximization algorithm to obtain maximum likelihood estimates with robust standard errors, the preferred method for data that are missing at random.<sup>32</sup> The trajectory analyses in Mplus identified group differences at T1 (i.e., the intercept); to further describe potential group differences in trajectories, we performed post hoc pair-wise comparisons in SPSS for the children who meet vs. do not meet the irritability threshold on impairment at the subsequent time points (T2, T3), false discovery rate (FDR) corrected for multiple comparisons. We examined continuity of the clinically optimized irritability score with a similar growth curve approach as well as correlations over time. Moreover, we examined whether meeting the irritability cutoff at one, two, or three time points significantly related to FLIS Child and Family Impairment using ANOVA with Tukey-corrected post hoc comparisons. Lastly, we examined whether meeting the irritability cutoff at preschool age predicted *DSM* disorders in early school age.

## **RESULTS**

### **Aim 1: Develop Optimized, Parsimonious Model of Clinically Salient Irritable Behaviors at Preschool Age**

The final stepwise logistic regression model identified 2 of 22 MAP-DB Temper Loss irritability items as uniquely predicting FLIS cross-domain impairment: “become frustrated easily” ( $p < .001$ ), and “break or destroy things during a temper tantrum, fall-out, or meltdown” ( $p = .002$ ). With every unit increase in score (increase in frequency) on “frustrated easily,” the odds of cross-domain impairment increased by 1.46; for “break/destroy,” odds increased by 1.56. (Of note, performance of each of these two items in predicting cross-domain impairment exceeded that of the full 22-item Temper Loss scale, odds ratio=1.04,

$p < .05$ .) This final model performed significantly better than the baseline model (i.e., modeling a constant only) ( $\chi^2 = 43.82$ ,  $df = 2$ ,  $p < .001$ ) as well as the step 1 model ( $\chi^2 = 10.39$ ,  $df = 1$ ,  $p = .001$ ), which added “frustrated easily” to the baseline model. The final model saw these two items (“frustrated easily,” “break/destroy”) explain 15.6% of the variance in cross-domain impairment and was a good fit to our data (Hosmer and Lemeshow  $\chi^2 = 1.59$ ,  $df = 5$ ,  $p = .902$ ). Results were similar when rerunning the model with probit-transformed irritability items to address potential skew in the predictors as well as with only two domains of impairment to account for the 11% of children not in childcare (Supplement 1, Results, available online).

## **Aim 2: Derive an Empirically-Based Cutoff for Clinically Optimized Irritability Score in Relation to Concurrent *DSM-5* Irritability-Related Syndromes (i.e., ODD, DMDD, Other Depressive Disorders)**

**Generating Cutoff**—Scores on the two items identified in Aim 1 (“frustrated easily” and “break/destroy”) were summed to create the clinically optimized irritability score used in Aims 2 and 3 (mean = 2.05, SD = 1.96, range = 0–10; Figure 1). ROC analyses showed that this optimized irritability score was a sensitive and specific indicator of irritability-related syndromes (ODD, DMDD, other non-DMDD depressive disorders [major depression or dysthymic disorder, or depressive disorder not otherwise specified]) (Table S1, available online). Areas under the curve for ODD, DMDD, and other depressive disorders (AUC = .80–.81, all  $p < .001$ , Figure 2) indicate good classification accuracy of the irritability score. Across all three disorder categories, the optimal cutoff was a score of 3 on the summed MAP-DB irritability items (peak Youden’s Indices = .46–.53), which balanced sensitivity (70–73%) and specificity (74–83%). That is, up to 73% of children identified as having an irritability-related diagnosis were correctly classified as meeting the irritability cutoff (sensitivity) and up to 83% of children who did not meet each of the diagnoses separately were correctly classified as not meeting the irritability cutoff (specificity). The categories (ODD, DMDD, other depressive disorders) did not differ on AUC, sensitivity, or specificity values ( $p > .05$ ). In addition, up to 58% of children meeting the irritability cutoff also meet diagnostic criteria for one of the three disorder categories, calculated separately (PPV), and 97% of children not meeting diagnostic criteria also do not meet the irritability cutoff (NPV) (Table S1, available online). Additional analyses with the “any *DSM* diagnosis classification” variable indicated significant AUC (.75,  $p < .001$ ) and identified the same cutoff (score of 3), but with lower sensitivity (51%) and slightly higher specificity (87%) than each of the irritability-related syndromes separately. PPV for *any* diagnosis vs. each of the irritability-related syndromes was higher (74%), though NPV was lower (71%) (Table S1, available online).

**Variation in Patterns Meeting Cutoff**—In this enriched sample, 30% scored above and 70% below the cutoff of 3. To achieve a score at the irritability cutoff value of 3, parents must have at least endorsed both irritable behaviors occurring over the course of the past month (e.g., one behavior occurring monthly [score of 1] and the other 1–3 days/week [score of 2] or, one irritable behavior occurring more frequently (most days of the week [score of 3 or higher]). Table S2 (available online) shows the frequencies of potential combinations of scores on the two items. Of children who meet the irritability cutoff, all were reported to



“become frustrated easily,” with most (96.1%) reporting the behavior as occurring at least weekly. This is consistent with the normative nature of this milder behavior. Thus, virtually all irritable young children who are impaired are easily frustrated. However, they vary more substantially in the frequency in which they exhibit the severely dysregulated behavior, i.e., “break or destroy things during a temper tantrum...” (23.6% never in the past month; 27.6% less than weekly; 30.7% weekly, 18.1% 4–6 days per week). That is, nearly a quarter never exhibits this behavior (over past month), and another quarter does so infrequently. In contrast, of children who do not meet the irritability threshold, about two-thirds (63.9%) endorse mild frequencies of “becom[ing] frustrated easily” (scored 1 or 2), but the large majority do not “break or destroy things during a temper tantrum...” (86.6% scored 0), which also rarely occurred in two independent MAPS community samples (<10%).<sup>4</sup> Children above and below the cut-off were sociodemographically comparable. Elevated irritability in children above cutoff was confirmed via higher irritability scores on independent irritability measures (ARI<sup>1</sup> and DB-DOS<sup>7</sup>; Table 1).

### **Aim 3: Test Predictive Validity of the Clinically Optimized Irritability Score Cutoff for Growth of Impairment and Irritability and Early School-Age *DSM* Disorders**

**Impairment Growth Patterns**—All fit indices (AIC, BIC, SSABIC) indicated that the two-group (meeting vs. not meeting the irritability cutoff) known class model fit better than single-group model for trajectories of the FLIS Child and Family Impairment subscales (Table 2). The two-group known class models identified distinct trajectories of impairment for high vs. low irritable children. In particular, highly irritable preschoolers had significantly greater impairment on the FLIS Child and Family Impairment subscales at baseline (group differences in intercepts,  $p < .05$ ), which was maintained throughout the preschool to school-age period (means at T2 and T3, both  $p < .05$ , corrected) (Figure 3). In addition, meeting the irritability cutoff at zero, one, two, or three time points significantly predicted impairment at early school age (FLIS Child Impairment,  $F_{3,351} = 19.840$ ,  $p < .001$ ; a trend for FLIS Family Impairment,  $F_{3,351} = 2.564$ ,  $p = .055$ ). That is, post hoc analyses revealed that children who scored 3 or above on the clinically optimized irritability score at all three time points had greater FLIS Child Impairment at T3 than children who met at just two or one time points ( $p < .05$ ).

**Irritability Growth Patterns**—All fit indices supported a two-group known class model over a single group model for irritability growth curves, indicating that high irritability preschoolers maintain higher levels of irritability from preschool to school age than low irritability children (difference in intercepts, T1:  $p < .05$ ; difference in subsequent irritability levels, T2 and T3,  $p < .05$ , corrected) (Figure 1C). Correlation analyses indicated that the clinically optimized irritability score is moderately stable over time ( $r_{T1, T2} = .55$ ,  $p < .001$ ;  $r_{T1, T3} = .54$ ,  $p < .001$ ). Examining overlap in these patterns longitudinally (children meeting cutoff at T1, T2, and T3; Table S3, available online) showed that almost half (48.82%) of children who met cutoff at baseline continued to do so at T2, and 22.83% had persistent irritability across all three time points. Of children who were not identified as high irritable at T1, 88.59% remain below the threshold at T2, and 79.53% continue to stay below the cutoff at both T2 and T3.

**Early School-Age *DSM* Disorders**—Finally, children who met the irritability cutoff as preschoolers (T1) continued to be more likely at early school age (T3) to meet clinical or subclinical criteria for ODD ( $\chi^2=30.78$ ,  $df=2$ ,  $p<.001$ ), DMDD ( $\chi^2=9.66$ ,  $df=2$ ,  $p=.008$ ), or any *DSM* disorder ( $\chi^2=33.25$ ,  $df=1$ ,  $p<.001$ ). However, likelihood of being diagnosed with other depressive disorders did not significantly differ between groups ( $\chi^2=2.26$ ,  $df=1$ ,  $p=.133$ ).

## DISCUSSION

The current work represents an important step toward translating basic developmental work into clinical application. The developmental psychopathology perspective emphasizes two principles in this translation process, i.e., dimensional assessment and developmental specification<sup>3, 33</sup>: first, modeling behaviors dimensionally allows for identification of clinical behaviors as they unfold from mild to severe and will enable earlier identification of at-risk patterns, rather than relying only on extreme expressions.<sup>34</sup> We incorporated this into the current work by characterizing irritability dimensionally and empirically identifying the degree to which irritability becomes “worrisome,” i.e., related to impairment, concurrent and future diagnoses with irritability as a central component, and clinical course. Interestingly, this is optimized when both high frequency (normative) and low frequency (atypical) expressions of irritability are considered. Second, taking into account developmental period can yield a classification system informed by empirical study of developmental processes from the bottom up, rather than simply downward extension of adult/older child categories.<sup>34</sup> We addressed this principle by employing a method developmentally specified for early childhood.<sup>3</sup> This is an example of how a developmental science base can critically inform nosologic systems such as the *DSM*, which does not currently use empirically derived parameters for identification of irritability or other behavioral patterns that vary across development.<sup>3</sup>

Along these lines, the present findings along with our prior work<sup>4, 6</sup> do not support the current *DSM* approach, i.e., exclusion of young children from DMDD diagnosis and requiring near daily occurrence of irritable behaviors in ODD. The clinically optimized irritability screener derived from the MAP-DB identified a subgroup of impaired young children who were at high probabilistic risk for persistence and future *DSM* disorders. Further, the frequency point at which irritable behaviors become clinically informative appears to vary depending on whether the expression of irritability is a commonly occurring normative misbehavior or atypical and rarely occurring severe behavior. That is, the mild misbehavior item, “frustrated easily,” is clinically concerning when displayed nearly every day, but the severe behavior item, “break/destroy,” is clinically concerning at lower frequencies. This suggests that a blanket frequency threshold as is used by the *DSM* is not maximally informative, as thresholds may be too stringent for some behaviors and not stringent enough for others.<sup>4</sup>

This two-item clinically optimized irritability score is brief, yet with good sensitivity and specificity in relation to *DSM* diagnoses, and holds promise as a quick, efficient, transdiagnostic pediatric screening method to flag preschoolers with clinically concerning irritability. Temper tantrums are common in early childhood. Here we show that children

who are very easily frustrated and who engage in highly dysregulated tantrums are significantly more likely to show pervasive and sustained impairment and clinically salient irritability as expressed in concurrent and future *DSM-5* disorders. The frequency cutoff we identified in preschoolers is notably low in comparison to the higher frequency of irritable behaviors (irritable or angry mood most of the day, nearly every day; severe, inappropriate temper outbursts three or more times per week) required by DMDD diagnostic criteria, which were developed for children older than 6 years, as well as to the *DSM-5* ODD criteria (nearly daily).<sup>35</sup> The low frequency threshold identified on our clinically optimized scale may reflect the capacity of this empirical approach to separate the wheat from the chaff, i.e., honing in on behaviors specifically related to impairment and *DSM* diagnosis that might otherwise be obscured by the noise of other early childhood behaviors. Indeed, validation of these items and threshold shows that they identify children with a distinctive developmental trajectory, marked by persistent irritability, impairment, and *DSM* disorders.

Interestingly, the clinically optimized irritability score threshold reflects one mild, normative misbehavior item, “frustrated easily,” occurring frequently in addition to one severe, pathognomonic, low-occurring behavior, “break/destroy.” That is, within the group of children who meet the irritability cutoff in our sample, 96% of children are reported to become frustrated easily (100%) 1–3 days per week, and the majority (76%) also exhibited destructive tantrums (break/destroy) over the past month. In contrast, 22% of children below the threshold became frustrated easily 1–3 day per week, and 13% exhibited destructive tantrums (break/destroy) over the past month. For the former behavior, clearly it is the frequency that differentiates clinical concern from normative misbehavior; thus, screening without frequency thresholds would over-identify. For dysregulated tantrums, on the other hand, low irritable children are unlikely to exhibit this atypical behavior. The normative misbehavior item, frustrated easily, likely underlies multiple other tantrum features. For example, a child who becomes frustrated easily across a range of contexts likely has tantrums triggered not only in expectable situations (e.g., times of transition and demand) but in developmentally unexpected situations (with non-parental adults) and may also have challenges restoring positive mood. In contrast, destructive tantrums (break/destroy) are worrisome by definition. The combination of these items suggests that a sole focus on extreme behaviors, e.g., for symptom counts in diagnostic criteria, will not be most discriminating at this age. These findings also suggest that such differentiation may assist in elucidating differential pathways from early irritability to clinical sequelae.

In terms of linkage to clinical disorders, we see similar patterns for both disruptive and mood outcomes at this age, suggesting a transdiagnostic pattern. Of note, as has been done previously,<sup>21, 23, 25</sup> we “made do” with an adapted approach to DMDD here despite the fact that it is technically not appropriate at preschool age within the current *DSM* approach. Our findings suggest that future research using an empirical approach to derive DMDD diagnostic parameters for this age group will be fruitful and will prevent delay in identification of the early phase of clinical irritability pathways via arbitrary age exclusions. Along these lines, unlike the *DSM-5* DMDD criteria designed for children older than 6 years, which require both temper tantrum/phasic and irritable mood/tonic behaviors,<sup>35</sup> the optimized clinical score we identified for preschoolers selected temper tantrum (break/destroy) and dysregulation/temper tantrum (frustrated easily) behaviors (although tantrum

and mood items from MAP-DB were included in the model at Step 1). This suggests the possibility that temper tantrums may be more clinically salient than irritable mood at this age, which is consistent with prior work and also has face validity since expressions of dysregulated behavior are likely to be prominent and disruptive across developmental contexts.<sup>36</sup> On the other hand, it remains possible that this is a methodologic artifact as tantrum items are disproportionately represented on the preschool version of the MAP-DB, and more extensive coverage of irritable mood may increase its salience for prediction. To this end, we have increased coverage of irritable mood in newer versions of the MAP-DB for infant/toddlers and older youths.<sup>11</sup> This will be an important area for future research.

It is important to note limitations of our study. First, our main measures are based on maternal report (surveys and clinician interview) with the potential for shared method variance or biases. However, we note that patterns differentiating normal vs. abnormal irritability on the MAP-DB have been robust to a range of individual differences (e.g. racial/ethnic and poverty status), patterns we have confirmed via qualitative interviews.<sup>4, 37</sup> Moreover, additional analyses with Anger Modulation from the DB-DOS, a performance-based, observational measure, corroborated our findings with the maternal report measures, as did the ARI (Table 1). Nevertheless, expanding this approach more directly incorporating cross-method and cross-informant approaches will be important as cross-contextual variation such as that captured by the DB-DOS has proven clinically informative.<sup>7, 38</sup> Second, it is necessary to replicate these findings and to validate clinically optimized irritability score approach as a clinical screener within referred populations to further establish its psychometric properties and clinical utility as well as within population-based samples, as PPV and NPV are affected by prevalence of clinical phenomena. Pending replication, we cannot definitively conclude that the two behaviors selected here are “the” optimal early identification indicators. Moreover, it should be noted that this cutoff specifically reflects preschool age. Because irritability levels normatively decrease from toddlerhood through childhood (although individual trajectories vary),<sup>39</sup> different cutoffs will be needed for other ages. Regardless, the present findings do suggest that differentiation of clinically salient irritability in young children can be done efficiently, sensitively, and with good predictive validity. Last, examination of age and sex differences in these pathways will be another important area for research.

To conclude, the present study uses an empirical approach to identify irritable behaviors that most sensitively and specifically characterize clinically significant irritability within the developmental context of early childhood. Notably, for this developmental period with high levels of normative temper outbursts, we have identified a brief (2 item) clinically optimized score that is sensitive and specific for differentiating preschoolers with irritability-related syndromes and high risk for persistent irritability-related impairment from those without. This is a foundational step towards screening methods that provide an empirical basis for “when to worry” about young children with clinically impairing irritability. This provides a critical building block for a crucial science supporting targeted early intervention in this initial phase of the clinical sequence when experience-dependent neuroplasticity is at its peak.<sup>3</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors wish to thank Kimberly McCarthy, BS, of the University of Connecticut Psychiatry Department, for assistance in data management and the families who participated.

*Funding:* This study was supported by NIMH R01MH082830, 2U01MH082830, U01MH090301, and 2U01MH082830.

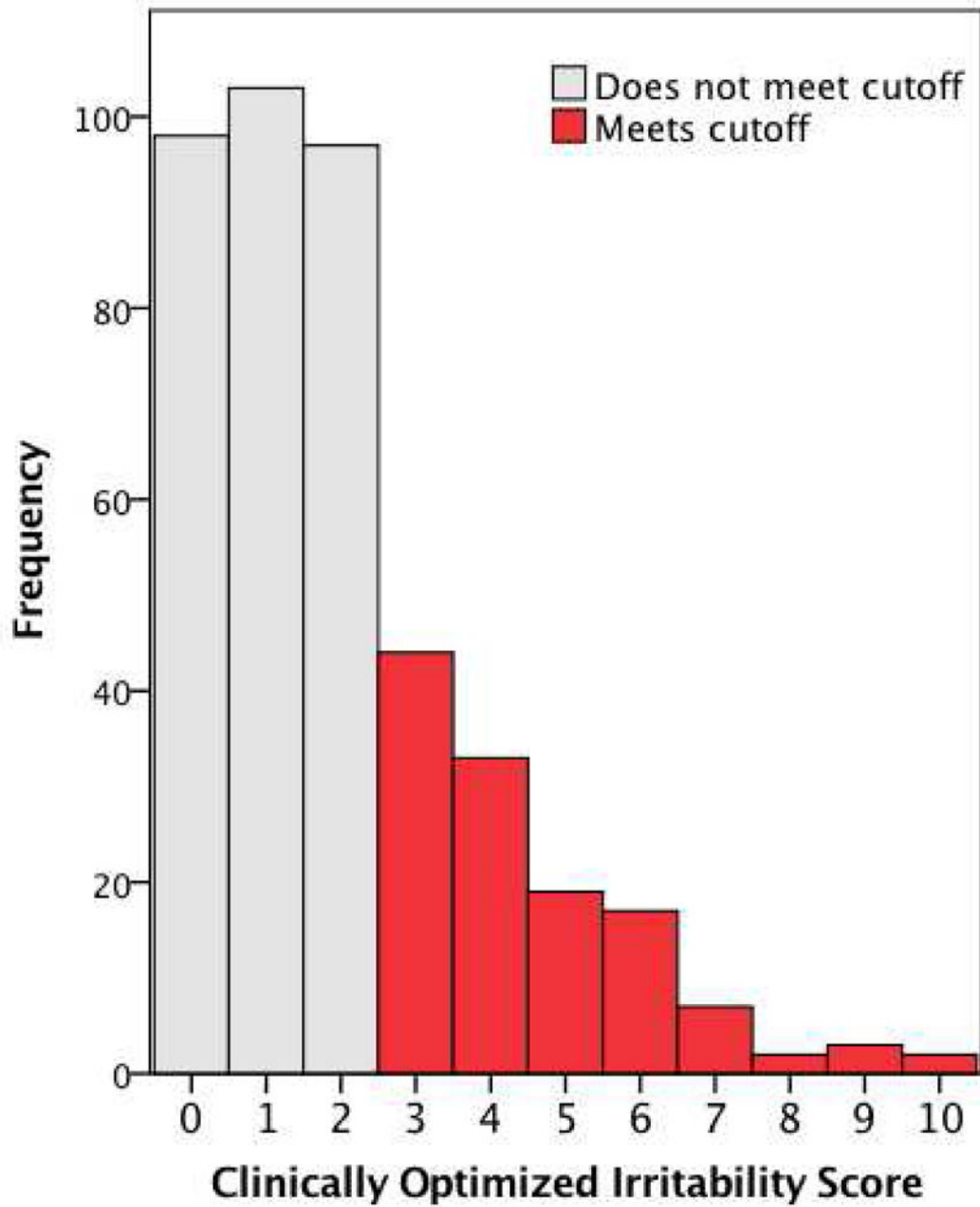
## References

1. Stringaris A, Goodman R, Ferdinando S, et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *J Child Psychol Psychiatry*. 2012; 53:1109–1117. [PubMed: 22574736]
2. Leibenluft E. Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *Am J Psychiatry*. 2011; 168:129–142. [PubMed: 21123313]
3. Wakschlag LS, Perlman SB, Blair RJ, Leibenluft E, Briggs-Gowan MJ, Pine DS. The neurodevelopmental basis of early childhood disruptive behavior: Irritable and callous phenotypes as exemplars. in press.
4. Wakschlag LS, Choi SW, Carter AS, et al. Defining the developmental parameters of temper loss in early childhood: implications for developmental psychopathology. *J Child Psychol Psychiatry*. 2012; 53:1099–1108. [PubMed: 22928674]
5. Perlman SB, Jones BM, Wakschlag LS, Axelson D, Birmaher B, Phillips ML. Neural substrates of child irritability in typically developing and psychiatric populations. *Dev Cogn Neurosci*. 2015; 14:71–80. [PubMed: 26218424]
6. Wakschlag LS, Estabrook R, Petittlerc A, et al. Clinical Implications of a Dimensional Approach: The Normal:Abnormal Spectrum of Early Irritability. *J Am Acad Child Adolesc Psychiatry*. 2015; 54:626–634. [PubMed: 26210331]
7. Petittlerc A, Briggs-Gowan MJ, Estabrook R, et al. Contextual variation in young children's observed disruptive behavior on the DB-DOS: implications for early identification. *J Child Psychol Psychiatry*. 2015; 56:1008–1016. [PubMed: 26095766]
8. Brotman MA, Schmajuk M, Rich BA, et al. Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. *Biol Psychiatry*. 2006; 60:991–997. [PubMed: 17056393]
9. Stringaris A, Cohen P, Pine DS, Leibenluft E. Adult outcomes of youth irritability: a 20-year prospective community-based study. *Am J Psychiatry*. 2009; 166:1048–1054. [PubMed: 19570932]
10. Copeland WE, Brotman MA, Costello EJ. Normative Irritability in Youth: Developmental Findings From the Great Smoky Mountains Study. *J Am Acad Child Adolesc Psychiatry*. 2015; 54:635–642. [PubMed: 26210332]
11. Biedzio, D., Wakschlag, LS. Developmental emergence of disruptive behaviors beginning in infancy: Delineating normal:abnormal boundaries to enhance early identification. In: Zeenah, C., editor. *Handbook of Infant Mental Health*. 4. New York: Guilford; in press
12. Belden AC, Thomson NR, Luby JL. Temper tantrums in healthy versus depressed and disruptive preschoolers: defining tantrum behaviors associated with clinical problems. *J Pediatr*. 2008; 152:117–122. [PubMed: 18154912]
13. Wakschlag LS, Briggs-Gowan MJ, Choi SW, et al. Advancing a multidimensional, developmental spectrum approach to preschool disruptive behavior. *J Am Acad Child Adolesc Psychiatry*. 2014; 53:82–96e83. [PubMed: 24342388]
14. Bufferd, SJ., Dyson, M., Hernandez, I., Wakschlag, LS. Explicating the “developmental” in preschool psychopathology. In: Cicchetti, D., editor. *Handbook of Developmental Psychopathology*. 3. Hoboken, NJ: Wiley; 2015. p. 152-186.

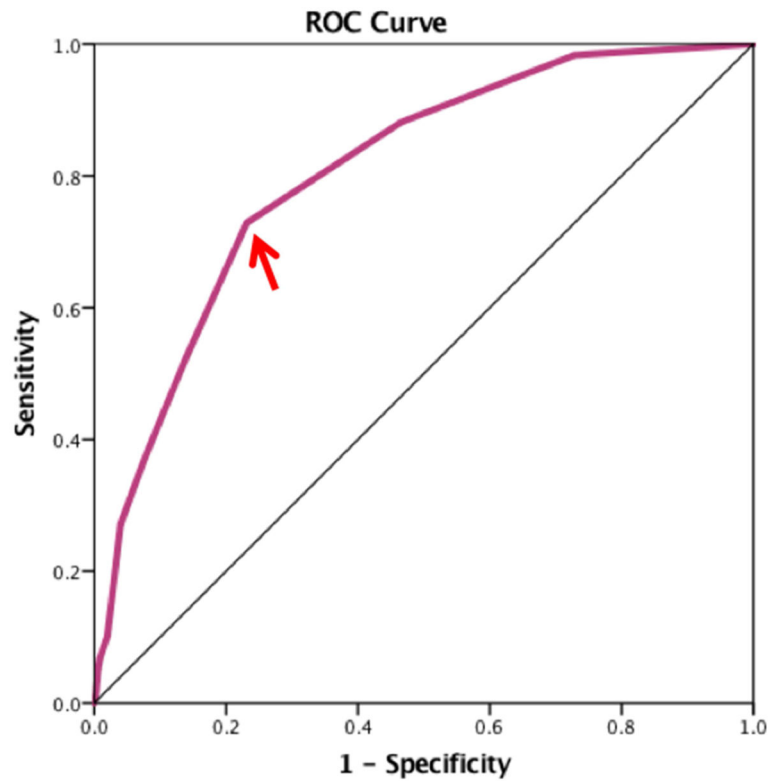
15. Wakschlag LS, Hill C, Carter AS, et al. Observational Assessment of Preschool Disruptive Behavior, Part I: reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *J Am Acad Child Adolesc Psychiatry*. 2008; 47:622–631. [PubMed: 18434926]
16. Wakschlag LS, Briggs-Gowan MJ, Hill C, et al. Observational Assessment of Preschool Disruptive Behavior, Part II: validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *J Am Acad Child Adolesc Psychiatry*. 2008; 47:632–641. [PubMed: 18434925]
17. Egger, HL., Angold, A. The preschool age psychiatric assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In: DelCarmen-Wiggins, R., Carter, AS., editors. *Handbook of Infant and Toddler Mental Health Assessment*. New York: Oxford University Press; 2004. p. 223-243.
18. Wakschlag, LS., Leventhal, B., Thomas, B. Disruptive behavior disorders & ADHD in preschool children: Characterizing heterotypic continuities for a developmentally informed nosology for DSM-5. In: Narrow, W.First, M.Sirovatka, P., Regier, D., editors. *Age and gender considerations in psychiatric diagnosis: A research agenda for DSM-5*. Arlington, VA: American Psychiatric Association; 2007. p. 243-258.
19. Dougherty LR, Smith VC, Bufferd SJ, Kessel E, Carlson GA, Klein DN. Preschool irritability predicts child psychopathology, functional impairment, and service use at age nine. *J Child Psychol Psychiatry*. 2015; 56:999–1007. [PubMed: 26259142]
20. Egger HL, Angold A. Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *J Child Psychol Psychiatry*. 2006; 47:313–337. [PubMed: 16492262]
21. Dougherty LR, Smith VC, Bufferd SJ, et al. DSM-5 disruptive mood dysregulation disorder: correlates and predictors in young children. *Psychol Med*. 2014; 44:2339–2350. [PubMed: 24443797]
22. Kessel EM, Dougherty LR, Kujawa A, Hajcak G, Carlson GA, Klein DN. Longitudinal Associations Between Preschool Disruptive Mood Dysregulation Disorder Symptoms and Neural Reactivity to Monetary Reward During Preadolescence. *J Child Adolesc Psychopharmacol*. 2016; 26:131–137. [PubMed: 26771832]
23. Copeland WE, Angold A, Costello EJ, Egger H. Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *Am J Psychiatry*. 2013; 170:173–179. [PubMed: 23377638]
24. Briggs-Gowan, MJ., Horowitz, S., Carter, AS. Unpublished Measure. University of Connecticut Health Center & University of Massachusetts; Boston: 2010. The Family Life Impairment Scale: Version 3.0.
25. Dougherty LR, Smith VC, Bufferd SJ, Kessel EM, Carlson GA, Klein DN. Disruptive mood dysregulation disorder at the age of 6 years and clinical and functional outcomes 3 years later. *Psychol Med*. 2016; 46:1103–1114. [PubMed: 26786551]
26. Kaufman J, Birmaher B, Brent D, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*. 1997; 36:980–988. [PubMed: 9204677]
27. Chong IG, Jun CH. Performance of some variable selection methods when multicollinearity is present. *Chemometr Intell Lab*. 2005; 78:103–112.
28. Lopez-Raton M, Cadarso-Suarez C, Rodriguez-Alvarez MX, Gude-Sampedro F. *OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests*. *J Stat Softw*. 2014; 61:1–36.
29. Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. *Pediatrics*. 2001; 108:192–196. [PubMed: 11433077]
30. Hensley ML, Spriggs DR. Cancer screening: how good is good enough? *J Clin Oncol*. 2004; 22:4037–4039. [PubMed: 15381685]
31. Bollen, K., Curran, PJ. *Latent Curve Models: A Structural Equations Approach*. Hoboken, NJ: Wiley; 2006.
32. McCartney K, Burchinal MR, Bub KL. Best practices in quantitative methods for developmentalists. *Monogr Soc Res Child Dev*. 2006; 71:1–145.



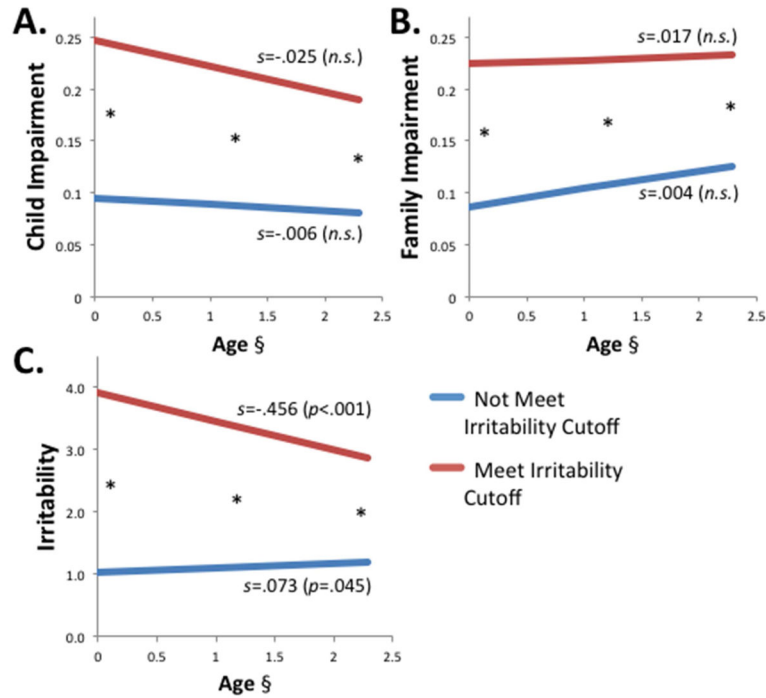
33. Wakschlag LS, Tolan PH, Leventhal BL. Research Review: ‘Ain’t misbehavin’: Towards a developmentally-specified nosology for preschool disruptive behavior. *J Child Psychol Psychiatry*. 2010; 51:3–22. [PubMed: 19874427]
34. Sroufe LA. The Concept of Development in Developmental Psychopathology. *Child Dev Perspect*. 2009; 3:178–183. [PubMed: 20161376]
35. American Psychiatric Association. *DSM-5*. Washington, DC: Author; 2013.
36. Carlson GA, Danzig AP, Dougherty LR, Bufferd SJ, Klein DN. Loss of Temper and Irritability: The Relationship to Tantrums in a Community and Clinical Sample. *J Child Adolesc Psychopharmacol*. 2016; 26:114–122. [PubMed: 26783943]
37. Shaunfield S, Petittlerc A, Kaiser K, et al. “In their own voices”: A mixed methods study of socio-cultural differences in maternal perceptions of young children’s disruptive behavior. in prep.
38. Gray SA, Carter AS, Briggs-Gowan MJ, et al. Preschool children’s observed disruptive behavior: variations across sex, interactional context, and disruptive psychopathology. *J Clin Child Adolesc Psychol*. 2012; 41:499–507. [PubMed: 22540388]
39. Wiggins JL, Mitchell C, Stringaris A, Leibenluft E. Developmental trajectories of irritability and bidirectional associations with maternal depression. *J Am Acad Child Adolesc Psychiatry*. 2014; 53:1191–1205. 1205e1191–1194. [PubMed: 25440309]



**Figure 1.** Distribution of scores on clinically optimized irritability score. Note: Irritability items (summed) identified through stepwise logistic regression predicting cross-domain impairment (Aim 1). Cutoff (score of 3) empirically derived through receiver operating characteristic analysis with *DSM-5* criteria (Aim 2).



**Figure 2.** Receiver operating characteristic (ROC) curve for clinically optimized irritability score (criterion variable) and *DSM* diagnosis (classification variable). Note: Arrow shows peak specificity and sensitivity. Disruptive mood dysregulation disorder (DMDD) pictured for illustration; similar results for oppositional defiant disorder and other non-DMDD depressive disorders.



**Figure 3.**

Subsequent impairment and irritability trajectories of children who meet vs. do not meet the irritability cutoff. Note: Predicted values calculated from estimates generated by growth mixture modeling with a known class (irritability cutoff status, see Aim 3), spanning mean age at T1 (4.66 years) to mean age at T3 (7.08 years). Slopes ( $s$ ) and significance values marked. Asterisks indicate that groups significantly differ at each time point ( $p < .05$ , corrected for multiple comparisons). Age values centered on mean age at baseline, 4.66 years.

Table 1

## Participant Characteristics

	All Participants	Meet Irritability Cutoff	Not Meet Irritability Cutoff	$\chi^2$	df	p
N	425	127	298			
Child gender [% female]	51.1	48.0	52.3	0.66	1	.42
Child race/ethnicity (%)				3.64	3	.31
African American	49.9	53.5	48.3			
Hispanic	29.9	28.3	31.5			
European American	18.6	18.1	18.8			
Other	1.6	0.0	2.3			
Mother education (%)				8.53	5	.13
Less than high school	3.5	4.8	3.1			
High school/GED	20.1	26.2	17.9			
Some college	35.0	34.1	34.4			
Associates degree	16.8	14.3	18.2			
Bachelor's degree	14.2	8.7	16.5			
Graduate degree	10.4	11.9	10.0			
Child age at T1 [y]	4.66 (.85)	4.51 (.80)	4.72 (.86)	2.42	423	.016
ARI [T3]	1.43 (1.69)	2.9 (1.7)	0.81 (1.2)	13.8	408	<.001
Disruptive Behavior Diagnostic Observation Schedule [DB-DOS, T1] – Anger Modulation						
Parent Context	-.012	.473	-.218	3.20	406	.002
Examiner Context	-.008	.170	-.083	1.94	406	.053

Note: Values in parentheses indicate standard deviations. Factor scores for Disruptive Behavior Diagnostic Observation Schedule (DB-DOS) Anger Modulation generated from model in <sup>7</sup>. Positive scores reflect Anger Modulation above the mean, and negative scores indicate scores below the mean. ARI=Affective Reactivity Index.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 2**

Known Class Growth Mixture Model Fit Information and Estimates for Trajectories of Children Who Meet vs. Do Not Meet Irritability Cutoff

<i>Model Type</i>	<b>Model Fit Indices</b>			<b>2-Group Model Estimates</b>				
		<b>1-Group Fit</b>	<b>2-Group Fit</b>	<b>Groups</b>	<b>Intercept (mean at age 4.66)</b>	<b>p</b>	<b>Slope (rate of change over time)</b>	<b>p</b>
<b>Child Impairment</b>	<i>AIC</i>	90.959	<b>-66.567</b>	Meet irritability cutoff	0.247	<.001	-0.025	.273
	<i>BIC</i>	123.223	<b>-2.037</b>	Not meet irritability cutoff	0.095	<.001	-0.006	.529
	<i>SSABIC</i>	97.837	<b>-52.810</b>					
<b>Family Impairment</b>	<i>AIC</i>	402.687	<b>281.007</b>	Meet irritability cutoff	0.224	.001	0.004	.905
	<i>BIC</i>	434.951	<b>345.537</b>	Not meet irritability cutoff	0.087	<.001	0.017	.157
	<i>SSABIC</i>	409.565	<b>294.764</b>					
<b>Irritability</b>	<i>AIC</i>	4437.089	<b>3845.824</b>	Meet irritability cutoff	3.908	<.001	-0.456	<.001
	<i>BIC</i>	4469.506	<b>3910.657</b>	Not meet irritability cutoff	1.029	<.001	0.073	.045
	<i>SSABIC</i>	4444.119	<b>3859.883</b>					

Note: For each of the types of models (Child Impairment, Family Impairment, and Irritability over time), model fit indices favored a two-group solution, with the known class of children meeting vs. not meeting the irritability cutoff at baseline, over a single group model. More favorable model fit indices (i.e., lower values) indicated in bold italics. *AIC* = Akaike Information Criterion; *BIC* = Bayesian Information Criterion; *SSABIC* = sample size-adjusted *BIC*.