



## ORIGINAL ARTICLE

# What I see, what you say: How cross-method variation sharpens characterization of irritability in early childhood

Alyssa J. Parker<sup>1</sup>  | Peyton Brock<sup>2</sup> | Maria Kryza-Lacombe<sup>3</sup> |  
Margaret Briggs-Gowan<sup>4</sup> | Lea R. Dougherty<sup>1</sup> | Lauren S. Wakschlag<sup>5</sup> |  
Jillian Lee Wiggins<sup>2,3</sup> 

<sup>1</sup>University of Maryland, College Park, Maryland, USA

<sup>2</sup>Department of Psychology, San Diego State University, San Diego, California, USA

<sup>3</sup>San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology, San Diego, California, USA

<sup>4</sup>Child and Adolescent Psychiatry, University of Connecticut Health System, Farmington, Connecticut, USA

<sup>5</sup>Feinberg School of Medicine and Institute for Innovations in Developmental Sciences, Northwestern University, Evanston, Illinois, USA

## Correspondence

Alyssa J. Parker, Office 3140, Biology and Psychology Building, 4094 Campus Dr, College Park, MD 20742, USA.  
Email: [ajparker@umd.edu](mailto:ajparker@umd.edu)

## Funding information

National Institute of Mental Health, Grant/Award Numbers: 1R01MH082830, 2U01MH082830

## Abstract

**Objectives:** Identification of clinically significant irritability in preschool age is important to implement effective interventions. However, varying informant and measurement methods display distinct patterns. These patterns are associated with concurrent and future mental health concerns. Patterns across multi-informant methods in early-childhood irritability may have clinical utility, identifying risk for impaired psychosocial functioning.

**Methods:** Using data from the Multidimensional Assessment of Preschoolers Study (N = 425), latent profile analysis identified irritability patterns through the parent-reported Multidimensional Assessment Profile Scales–Temper Loss (MAPS-TL), parent-reported interviewer-rated Preschool Age Psychiatric Assessment (PAPA), and observer-rated Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). These profiles were characterized on protective factors, global functioning, and mental health syndromes, concurrently and at early school age and preadolescent follow-up.

**Results:** Fit indices favored a five-class model: Low All, High Observation with Examiner (high DB-DOS Examiner Context), High All, High Parent Report (high MAPS-TL/PAPA), and Very High Parent Report (very high MAPS-TL/PAPA). Whereas Low All and High Observation with Examiner exhibited strong psychosocial functioning, remaining profiles showed impaired psychosocial functioning, with the Very High Parent Report group showing higher impairment at follow-ups,  $d_s = 0.37\text{--}1.25$ .

**Conclusions:** Multi-informant measurements of irritability may have utility for clinical prediction, and future studies should test utility for diagnostic precision.

## KEYWORDS

assessment, child psychiatry, emotion dysregulation, irritability, measurement

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. International Journal of Methods in Psychiatric Research published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Irritability is a dispositional tendency toward anger and frustration that exists on a normal:abnormal spectrum (Brotman et al., 2017; Wakschlag et al., 2012). Clinically salient irritability is frequent, dysregulated, and disproportionate to context, and includes both persistent sullen or grouchy mood and temper tantrum elements (Leibenluft, 2017; Wakschlag et al., 2012). Clinically significant irritability in children can have both short- and long-term negative impacts on socioemotional functioning (Evans et al., 2017; Klein et al., 2021; Leibenluft, 2017; Wiggins, Ureña Rosario, et al., 2023). These patterns are evident in toddlers and preschoolers, with predictive utility for internalizing and externalizing symptoms and disorders, poorer peer functioning, poorer physical health, and increased risk for antidepressant and educational service use in adolescence (Finlay-Jones et al., 2023; Sorcher et al., 2022; Wakschlag et al., 2018). Irritability is the one of the most common concerns of parents who seek psychological evaluations for their children (Collishaw et al., 2010; Evans et al., 2022; Wakschlag et al., 2023). Given the centrality and long-term impact of early childhood irritability and the fact that multi-method assessment is a gold standard, elucidating the import of their informant/method patterns is an important next step to advancing clinical applications which are nascent (Wiggins, Roy, et al., 2023).

Research on the assessment of pediatric irritability has advanced dramatically in the past 20 years, including the development of parent-reported surveys and interviews as well as observer-rated assessments that reliably measure irritability (Egger et al., 2006; Roy & Comer, 2020; Wakschlag et al., 2020). The Multidimensional Assessment Profile Scales-Temper Loss (MAPS-TL; Wiggins, Ureña Rosario, et al., 2023) (formerly known as the Multidimensional Assessment Profile Disruptive Behavior [MAP-DB; Wakschlag et al., 2015]), is a developmentally-specified parent report survey of young children's irritability that characterizes the normal:abnormal spectrum of irritability across mood, behavior, and context, and uses frequencies to increase reporting precision (Wakschlag et al., 2018). The Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006) is a symptom based structured Diagnostic and Statistical Manual (DSM)-oriented interview with the parent, which allows for interviewer probing. As irritability is a transdiagnostic indicator of both internalizing and externalizing syndromes, Dougherty et al. derived a clinical irritability symptom indicator from multiple syndromes assessed on the PAPA (Dougherty et al., 2020). This greater depth of the interview is intensive by cost and training requirements. Furthermore, the PAPA irritability score is a symptom count that considers more extreme behaviors indicative of clinical significance, and thus the PAPA irritability score may not provide as useful information on children with subclinical irritability (Dougherty et al., 2020; Wakschlag et al., 2012). In contrast, the Disruptive Behavior Diagnostic Observation Schedule (DBDOS; Wakschlag, Briggs-Gowan, et al., 2008) Anger Modulation factor is a standardized reviewer based set of ordinal codes derived from a developmentally-based diagnostic observation by trained coders that

uses "presses" to efficiently elicit variation in how young children regulate mood and behavior across motivational and interactional contexts (Wakschlag, Hill, et al., 2008). This approach enables direct observation and ratings of behaviors of interest during interactions with caregivers and an examiner. However, the DB-DOS requires trained personnel, sufficient lab space, and post observation rigorous coding (Petitclerc et al., 2015), although it has recently been validated for remote administration (Krogh-Jespersen et al., 2022). Further, though the DB-DOS is a well validated measure, it assesses irritability within a standardized, lab-based setting compared to the more naturalistic ratings provided by parents; this may introduce an additional source of measurement variation. Of note, all measures considered were specifically developed for the early childhood and are not merely "downward" extensions of measures first developed based on expressions at older ages. Since these methods all take different approaches and informants to measure irritability, it is expected that different components of child irritability would be observed across measures. However, it is unclear what the pattern among measures means for the prognosis, presentation, tailored intervention, and outcomes in irritable children. There is also no guidance as for whom and when these different measurement levels have utility, and there has not been systematic cross-method evaluation within the context of pragmatic considerations (Morris et al., 2020). The present paper is a first step toward addressing these gaps via examination of multi-method, multi-informant profiles and their clinical predictive utility and represents a step toward ultimately determining diagnostic utility within a clinical context, an important future goal.

Current recommendations are to assess child psychopathology from multiple points of view, including multiple adult informants (e.g., parents, clinicians, teachers) and various methods (e.g., interview, survey, observation; Achenbach, 2006; Briggs-Gowan et al., 2016); however, correlations between adult raters of children's symptoms are, on average, low-to-moderate (Achenbach et al., 1987). With regard to irritability, in older samples, parents of preadolescent and adolescent youths with anxiety reported lower irritability versus their child's reports (Stoddard et al., 2014), and oppositional defiant and disruptive mood dysregulation disorders in youths ages 6–21 were associated with higher parent-reported irritability (Mallidi et al., 2023). Researchers have shown that disagreement among measures and reporters likely reflect contextual variation (e.g., differences in behaviors at school or home) and differences in perceptions of the behaviors they are reporting (De Los Reyes, 2011; De Los Reyes & Epkins, 2023; Dirks et al., 2012). These varying informant-method patterns from child psychopathology assessments provide clinically-relevant information regarding concurrent functional impairment or additional symptoms (Augenstein et al., 2022; Lerner et al., 2017) and predict longitudinal youth outcomes (Makol et al., 2019). For example, examiner's observations of preschool disruptive behaviors in parent and examiner contexts aligned with parent- and teacher-reported irritability, respectively (De Los Reyes et al., 2009). Moreover, father-reported child problem behaviors at age 3 predicted internalizing symptoms at age 5, and mother-

father-, teacher-, and clinician-reported problem behaviors independently predicted problem behaviors at age 5, though clinician ratings did not contribute significantly above and beyond other raters (Kerr et al., 2007). Thus, differences in reporting are not only expected due to differences in perspectives, such as parents, teachers, and clinicians, but these patterns may also be informative in characterizing concurrent and future clinical outcomes.

In this emerging work on informant method patterns, in pediatric irritability, little work has been done to characterize informant/method patterns, particularly moving beyond agreement/disagreement, nor have such patterns been examined in relation to longitudinal outcomes. The first goal of the current study is to characterize informant method patterns in irritability measures during preschool age, a key period when tantrums are common yet clinically significant irritability is robustly predictive of later internalizing and externalizing symptoms (Stringaris et al., 2009; Vidal-Ribas et al., 2016). Our second goal is to examine the ways that such informant method patterns may be meaningful, evaluating the extent to which the patterns of irritability observed among informants/methods of assessment in preschool aged children relate to concurrent psychosocial functioning, including protective factors, global functioning, and mental health syndromes. We also examine whether such informant method patterns predict these psychosocial domains in follow-ups in early school age and preadolescence. Although investigation of the predictive value of informant method patterns is necessarily limited by the potential shared variance (e.g., parent report at one time point correlating more with parent report than, for example, clinician-report at a future timepoint), this is an important first step in understanding how consideration of multi-method, multi-informant measurement of irritability may improve prediction.

## 2 | METHODS

### 2.1 | Participants

The current study uses data from the Multidimensional Assessment of Preschoolers Study (MAPS; Wakschlag et al., 2015). Participants ( $N = 425$ ) were recruited from pediatric clinics across the greater Chicago area, oversampling for young children with disruptive behavior and a history of parental intimate partner violence (for a review of the MAPS study, see Wiggins, Roy, et al., 2023). Eligibility criteria included no significant developmental delays and an English-speaking mother. The children were assessed at multiple time points from preschool age through preadolescence (preschool age:  $M$  [SD] = 4.66 [0.85] years; early school age:  $M$  [SD] = 7.47 [0.88]; preadolescence age:  $M$  [SD] = 9.24 [0.77]). See Table 1 for full demographic information. The percentage of participants with complete data at early school age ranged from 65.53% to 72.47% across measures, and at preadolescence ranged from 69.41% to 71.29% across measures. Sociodemographic factors measured at baseline, such as gender, race/ethnicity, and poverty status (as measured by annual household income and household size using federal poverty

guidelines as well as receipt of Temporary Assistance for Needy Families services), did not differ after attrition,  $ps > 0.107$ .

All procedures were approved by Northwestern University Feinberg School of Medicine institutional review boards. Parents provided written permission, and children provided assent.

## 2.2 | Measures

### 2.2.1 | Measures of preschool age irritability

#### *Parent-reported survey: MAP-DB Temper Loss Scale*

Mothers rated their child's irritability on 22-items for the MAPS-TL (Temper Loss; Wakschlag et al., 2014; Wakschlag et al., 2015). Parents reported on the occurrence of child behavior (including tantrums and irritability mood) using actual frequencies rather than judgment of frequency (e.g., "never," sometimes) to improve accuracy of parent reporting across multiple motivational and interactional contexts. Items range from normative misbehaviors (e.g., having a tantrum during daily routines) to severe, uncommon behaviors (e.g., having a tantrum "out of the blue"; "staying angry a long time"). The Multidimensional Assessment Profile Scales-Temper Loss (MAPS-TL) has good reliability,  $\alpha = 0.97$ .

#### *Parent-reported interview: PAPA Irritability Index Score*

The Preschool Age Psychiatric Assessment (PAPA; Egger et al., 2006) is a parent-based structured interview for children between the ages of 2 and 6. The PAPA Irritability Index Score (Dougherty et al., 2013) is a 6-item scale which is derived from irritability symptoms across various sections of the PAPA and has good internal consistency ( $\alpha = 0.73$ ) and moderate stability from age 3 to age 6 (Dougherty et al., 2013). The six PAPA-derived irritability items include irritability-related symptoms, such as temper tantrums and being easily frustrated. Items cover the past 3 months of the child's life and are subsequently rated by the interviewer based on the intensity, frequency, and duration of each symptom.

#### *Observer-rated irritability: DB-DOS*

Trained raters coded children's regulation of irritability across the multiple contexts of the DB-DOS via the Anger Modulation codes (Wakschlag, Briggs-Gowan, et al., 2008). The DB-DOS is a standardized diagnostic observation specifically designed to distinguish normative variation from behaviors of clinical concern. The DB-DOS is a 50-min observation in a laboratory setting that consists of two different interactional contexts: one with a parent present and one with an examiner. These enable coding of child regulation of irritability with both a familiar and unfamiliar adult, across similar presses. Presses are designed tasks intended to efficiently elicit clinically salient behaviors, including compliance "do" and "don't," cleanup, withdrawal of attention, social play, and frustration tasks (Wakschlag et al., 2005). Behaviors observed during the DB-DOS are rated using a clinical continuum: normative behavior (0), normative misbehavior (1), behavior of concern (2), and atypical behavior (3). The DB-DOS

TABLE 1 Sample demographic information by profile group.

	Profile group					
	Total sample	Low all	High observation with examiner <sup>a</sup>	High all	High parent report <sup>b</sup>	Very high parent report <sup>b</sup>
N [%]	425 [100%]	182 [43%]	101 [24%]	33 [8%]	83 [19%]	26 [6%]
Sex (N [%])						
Male	208 [49%]	88 [48%]	49 [49%]	21 [64%]	35 [42%]	15 [58%]
Female	217 [51%]	94 [52%]	52 [51%]	12 [36%]	48 [58%]	11 [42%]
Race (N [%])						
Black/African American	212 [50%]	88 [49%]	53 [52%]	16 [49%]	41 [50%]	14 [54%]
Hispanic/Latino	127 [30%]	66 [36%]	22 [22%]	8 [24%]	26 [31%]	5 [19%]
White	79 [18%]	24 [13%]	25 [25%]	8 [24%]	15 [18%]	7 [27%]
Other	7 [2%]	4 [2%]	1 [1%]	1 [3%]	1 [1%]	0 [0%]
Poverty status (N [%])						
Poor	209 [49%]	91 [50%]	40 [40%]	17 [52%]	41 [49%]	20 [77%]
Non-poor	216 [51%]	91 [50%]	61 [60%]	16 [48%]	42 [51%]	6 [23%]
Age (M [SD])						
Preschool age	4.55 [0.81]	4.85 [0.86]	4.63 [0.80]	4.18 [0.74]	4.48 [0.80]	4.61 [0.84]
Early school age	7.40 [0.82]	7.60 [0.92]	7.39 [0.88]	7.39 [0.84]	7.32 [0.81]	7.30 [0.64]
Preadolescence	9.22 [0.69]	9.29 [0.74]	9.23 [0.75]	9.22 [0.80]	9.15 [0.63]	9.20 [0.55]

Abbreviations: DB-DOS, Disruptive Behavior Diagnostic Observation Schedule; MAPS-TL, Multidimensional Assessment Profile Scales–Temper Loss; PAPA, Preschool Age Psychiatric Assessment.

<sup>a</sup>Observation are parallel standardized ratings of the DB-DOS Anger Modulation Factor during interactions with parent and with examiner on the DB-DOS.

<sup>b</sup>Parent-reported PAPA-Derived Irritability Index, and the MAPS-TL parent survey.

has good internal consistency ( $\alpha = 0.87\text{--}0.92$ ) and interrater reliability (ICC = 0.91; Petittlerc et al., 2015), and is sensitive to intervention effects (Lind et al., 2014).

## 2.2.2 | Psychosocial domains

We examined potential differences in three domains, protective factors, global functioning, and mental health syndromes, based on informant method patterns (see Table 2). Detailed psychometric information on the measures is available in Supporting Information S1.

Internalizing symptoms (depressive, anxious symptoms) were examined at all timepoints. Externalizing symptoms, including oppositional defiant and conduct symptoms, were examined at the early school age and preadolescent timepoints, and attention-deficit hyperactivity (ADHD) symptoms were examined at the preadolescent timepoint. No concurrent analyses on externalizing symptoms were conducted at baseline due to overlap between irritability and externalizing symptom measurements. Moreover, we included trauma symptoms at all timepoints given the high overlap between irritability and trauma reactions (Durham et al., 2018). While we had

this specific focus for the mental health syndromes, we also included a global functioning measure at all timepoints to capture other ways irritability may have impaired children. In addition, we included protective factors during preschool age, as they have relevance for potential intervention points. We examined psychosocial domains both concurrently (preschool age) and as children aged (early school age and preadolescent) to examine whether such differences in these psychosocial functioning based on informant method patterns showed developmental patterning. To in part mitigate shared method variance, we included other reporters as available (e.g., parent report for most measures and child report for mental health, which was available at the early school age timepoint).

## 2.3 | Analysis

A latent profile analysis (LPA) was conducted in Mplus using the MAPS-TL, the PAPA Irritability Index Score, and the two DB-DOS Anger Modulation Scores (with a parent present and with an examiner present) to identify data driven informant method patterns in preschool irritability. A model was chosen based on the interpretability of the classes and comprehensive consideration of

relevant fit indices, including the Akaike Information Criterion (AIC; Akaike, 1998), the Bayesian Information Criterion (BIC; Schwartz, 1978), and the sample size adjusted Bayesian Information Criterion (SSABIC; Sclove, 1987), where smaller values indicate better fit. In addition, the Lo-Mendell-Rubin Likelihood Ratio Test (LMR-LRT; Lo et al., 2001), which measures the information gained with the addition of a class; entropy, which measure the stability of the classes; and class size, including minimum class size of 5% for robustness, were also considered. Maximum likelihood was used to address missing data in the LPA given its robustness of missing data up to 50% (Burchinal et al., 2006).

The irritability pattern groups from the LPA were then used to characterize the children's psychosocial functioning in the three domains (protective factors, global functioning, mental health syndromes) at preschool age and to predict psychosocial functioning (global functioning, mental health syndromes) of the children at early school age and preadolescence follow-ups using ANOVA tests in R (Version 4.3; R Core Team, 2023). All available data were used for these analyses. All *p*-values reflect false discovery rate correction, applied at each timepoint (i.e., preschool, early school age, and preadolescence).

### 3 | RESULTS

#### 3.1 | Irritability measures

As seen in Table 3, cross-method correlations were aligned with method variance. The most highly correlated indicators were those which relied on parent report, although these were two different methods—that is, survey and interview (MAPS-TL and the PAPA Irritability Index Score). The two contexts of the rated observed irritability (DB-DOS Anger Parent Context and DB-DOS Anger Examiner Context scores) were moderately correlated with each other. The parent report measures of irritability (MAPS-TL and the PAPA Irritability Index Score) were also correlated with the DB-DOS Anger Parent Context scores, though they were not correlated with the DB-DOS Anger Examiner Context scores, which also aligns with method variance related to child behavior with parent. This shows that, while within-informant and within-method measures are moderately-to-highly correlated with each other, there is lower agreement across different informants and methods. Furthermore, even among the most correlated measures, much variance among measures remains unexplained.

TABLE 2 Psychosocial outcome measures.<sup>a</sup>

Domain	Measure	Subscales	Reporter	Time points
Protective factors	The Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999)	Attachment, initiative skills	Parent	Preschool age
Global functioning	Children's Global Assessment Scale (CGAS; Shaffer et al., 1983)	Global functioning	Interviewer	Preschool age, early school age, preadolescence
Mental health syndromes	Trauma Symptom Checklist (TSCYC; Briere et al., 2001)	Trauma	Parent	Preschool age, early school age, preadolescence
	Infant-Toddler Social and Emotional Assessment (ITSEA; Carter et al., 2003)	Depressive, anxiety (generalized, separation)	Parent	Preschool age
	Berkeley Puppet Interview (BPI; Ablow et al., 1999; Measelle et al., 1998)	Depressive, anxiety (generalized, separation), conduct, oppositional defiant	Youth	Early school age
	The Schedule for Affective Disorders and Schizophrenia for School Age Children (K-SADS; Ambrosini, 2000)	Depressive, anxiety (generalized, separation), attention-deficit hyperactivity conduct, oppositional defiant	Parent	Preadolescence

<sup>a</sup>See Supporting Information S1 for details.

TABLE 3 Correlation among irritability measures.

	M	SD	1	2	3	4
1. MAPS-TL	21.11	19.27	-			
2. PAPA derived irritability	1.44	1.79	0.68***	-		
3. DB-DOS anger parent context	-0.11	2.02	0.17**	0.12*	-	
4. DB-DOS anger examiner context	-0.01	1.21	0.05	<0.01	0.34***	-

Abbreviations: DB-DOS, Disruptive Behavior Diagnostic Observation Schedule; MAPS-TL, Multidimensional Assessment Profile Scales-Temper Loss; PAPA, Preschool Age Psychiatric Assessment.

\**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001.

### 3.2 | Latent class analysis of irritability measures

Model fit indices are shown in Table 4. Entropy was acceptable for all models. While LMR-LRT supported a four-class model, the percentage of the smallest class was very low (3%). Because AIC, BIC, and SSABIC favored five classes, a five-class model was chosen.

As seen in Figure 1, the class comprising of the largest proportion of children ( $n = 182$ , 42.82% of participants) was characterized by being low on all measures of irritability (Low All). The second largest class ( $n = 101$ , 23.77% of participants) was characterized by high observed irritability in the DB-DOS Examiner context but low parent reported irritability on survey and interview measures of irritability (High Observation with Examiner). An additional class ( $n = 33$ , 7.77% of participants) scored high in both observed and parent-reported

survey and interview irritability (High All). Another class, High Parent Report ( $n = 83$ , 19.53% of participants), was characterized by high PAPA and MAPS-TL, but low DB-DOS observed irritability scores. Finally, the smallest class, Very High Parent Report ( $n = 26$ , 6.12% of participants), was characterized by very high MAPS-TL and PAPA scores, but low observed irritability in the DB-DOS.

Irritability pattern groups did not significantly differ by sex,  $\chi^2(4) = 5.21$ ,  $p = 0.267$ , nor race/ethnicity,  $\chi^2(12) = 14.40$ ,  $p = 0.276$ . However, at preschool age, the Low All group was significantly older than the High All,  $p < 0.001$ , by an average of 8 months, and the High Parent Report groups,  $p = 0.007$ , by approximately 4 months (omnibus:  $F(4, 420) = 6.20$ ,  $p < 0.001$ ). Furthermore, the Very High Parent Report group was more likely to be in poverty than the High Observation with Examiner,  $p = 0.001$

TABLE 4 Fit measures for latent class analysis.

Number of classes	AIC	BIC	SSABIC	Entropy	Lo-Mendell-Rubin adjusted LRT test		% of smallest class
					Value	$p$	
2	8072.440	8125.117	8083.864	0.863	297.38	<0.001	24%
3	7947.125	8020.063	7962.942	0.901	130.986	0.003	6%
4	7890.838	7984.036	7911.049	0.944	64.166	0.003	3%
5	<b>7839.517</b>	<b>7952.976</b>	<b>7864.122</b>	<b>0.849</b>	<b>59.359</b>	<b>0.357</b>	<b>6%</b>
6	7536.411	7670.130	7565.409	0.978	295.696	0.107	3%

Note; The bold indicates the chosen model.

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; LRT, Likelihood Ratio Test; SSABIC, sample size adjusted Bayesian Information Criterion.

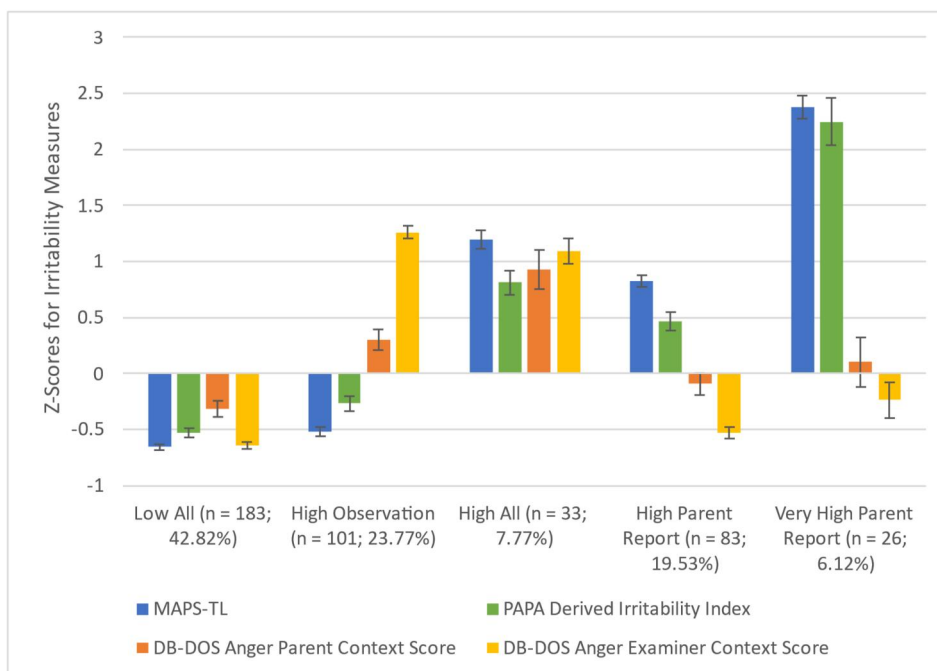


FIGURE 1 Irritability characterization of LPA groups. Irritability scores for each latent class. DB-DOS, Disruptive Behavior Diagnostic Observation Schedule; LPA, latent profile analysis; MAPS-TL, Multidimensional Assessment Scales–Temper Loss; PAPA, Preschool Age Psychiatric Assessment. Percentages refer to total percent of sample in each LPA group. Error bars represent standard errors.

(omnibus:  $X^2(4) = 8.34, p = 0.019$ ). Because of this, additional analyses were conducted adjusting for preschool age and poverty status to evaluate the impact on results. The results presented below remained significant after adjusting for age and poverty status. For full results, see Supporting Information S1. Other post-hoc comparisons of irritability pattern groups on age and poverty status were not significant.

### 3.3 | Predicting psychosocial domains using irritability pattern groups

Psychosocial domains were predicted using the irritability pattern groups through ANOVA. For any significant omnibus tests, follow-up

tests were performed to examine (1) the difference between the expected low irritability and elevated irritability pattern groups and (2) the differences among the elevated irritability groups. Only analyses with significant follow-up tests are reported in text, see Table 5 for full results.

#### 3.3.1 | Preschool age concurrent

##### *Omnibus results*

The irritability pattern groups were used to predict mental health at preschool age (Table 5). At preschool age, irritability group significantly distinguished children with differing global functioning, protective factors, and mental health symptoms.

TABLE 5 Differences between irritability pattern groups and global functioning, protective factors, and mental health measures.

Time point	Domain	Measure	Scale	F	df	p	$\eta^2$	95% CI of $\eta^2$		
								Low	High	
Preschool	Global functioning	CGAS	Global functioning	11.48	4, 388	<0.001	0.11	0.05	0.16	
	Protective factors	DECA	Attachment	7.72	4, 400	<0.001	0.07	0.03	0.12	
			Initiative skills	6.99	4, 400	<0.001	0.07	0.02	0.11	
	Mental health syndromes	ITSEA	Depressive	9.42	4, 418	<0.001	0.08	0.03	0.13	
			Anxiety							
			General	5.01	4, 416	0.002	0.05	0.01	0.08	
			Separation	13.44	4, 418	<0.001	0.11	0.06	0.17	
	TSC	Trauma	13.14	4, 271	<0.001	0.16	0.08	0.23		
School age	Global functioning	CGAS	Global functioning	6.40	4, 282	0.001	0.08	0.02	0.14	
	Mental health syndromes	BPI	Depressive	0.25	4, 295	0.999	0.00	0.00	0.01	
			Anxiety							
			General	0.78	4, 303	0.864	0.01	0.00	0.03	
			Separation	1.22	4, 295	0.672	0.02	0.00	0.04	
			Conduct	0.36	4, 293	0.999	0.01	0.00	0.02	
			Oppositional defiant	2.90	4, 296	0.443	0.04	0.00	0.08	
	TSC	Trauma	7.53	4, 289	<0.001	0.09	0.03	0.15		
Preadolescence	Global functioning	CGAS	Global functioning	5.54	4, 290	0.001	0.07	0.02	0.12	
	Mental health syndromes	K-SADS	Depressive	3.99	4, 293	0.010	0.05	0.01	0.10	
			Anxiety							
			General	6.26	4, 290	<0.001	0.08	0.02	0.13	
			Separation	17.07	4, 297	<0.001	0.19	0.11	0.26	
			Conduct	7.60	4, 298	0.031	0.09	0.03	0.15	
			Oppositional defiant	16.66	4, 298	<0.001	0.18	0.10	0.25	
			ADHD	6.00	4, 297	0.007	0.08	0.02	0.13	
	TSC	Trauma	7.22	4, 298	<0.001	0.09	0.03	0.14		

Note: The CGAS is a clinician-report measure. The TSC, K-SADS, DECA, and ITSEA are parent report measures. The BPI is a youth-report measure.  $p$ -values were corrected by time point including the post-hoc tests using the FDR correction.

Abbreviations: ADHD, attention-deficit hyperactivity; BPI, Berkeley Puppet Interview; CGAS, Children's Global Assessment Scale; DECA, Devereux Early Childhood Assessment; FDR, false discovery rate; ITSEA, Infant-Toddler Social and Emotional Assessment; K-SADS, Schedule for Affective Disorders and Schizophrenia for School Age Children; TSC, Trauma Symptom Checklist.



### Low versus elevated irritability at preschool age

As seen in Table 6, compared to the Low All group, the High All, High Parent Report, and Very High Parent Report groups had significantly higher concurrent separation anxiety ( $p = 0.011$ ,  $d = 0.50$ ;  $p < 0.001$ ,  $d = 0.57$ ;  $p < 0.001$ ,  $d = 1.25$ ; respectively), lower global functioning ( $p = 0.004$ ,  $d = -0.57$ ;  $p = 0.005$ ,  $d = -0.44$ ;  $p < 0.001$ ,  $d = -1.05$ ; respectively), and higher concurrent child trauma symptoms, ( $p = 0.004$ ,  $d = 0.75$ ;  $p = 0.038$ ,  $d = 0.51$ ;  $p < 0.001$ ,  $d = 0.99$ ; respectively). Furthermore, compared to the Low All group, the High All and Very High Parent Report groups had lower initiative skills ( $p = 0.009$ ,  $d = -0.54$ ;  $p < 0.001$ ,  $d = -0.83$ ; respectively). The Very High Parent Report group had higher concurrent generalized anxiety symptoms ( $p < 0.001$ ,  $d = 0.82$ ), higher depressive symptoms, ( $p < 0.001$ ,  $d = 0.85$ ), and lower attachment, ( $p < 0.001$ ,  $d = -0.76$ ), than the Low All group. There were no significant differences between the Low All and High Observation with Examiner groups.

### Comparing high parent-reported irritability groups at preschool age

Because the High Observation with Examiner group scored largely the same as the Low All group on these measures, for simplicity, the High Observation with Examiner group is not included within

the presented high irritability group comparisons, though they were included in analyses and  $p$ -value corrections; for the High Observation with Examiner group results, see Supporting Information S1.

As seen in Table 7, within the high irritability groups, compared to the High All and High Parent Report groups, the Very High Parent Report group had lower concurrent global functioning ( $p = 0.019$ ,  $d = -0.56$ ;  $p < 0.001$ ,  $d = -0.75$ ), higher child trauma symptoms ( $p = 0.003$ ,  $d = 0.57$ ;  $p < 0.001$ ,  $d = 0.80$ ), higher generalized anxiety ( $p = 0.039$ ,  $d = 0.52$ ;  $p = 0.003$ ,  $d = 0.68$ ), higher separation anxiety ( $p = 0.016$ ,  $d = 0.63$ ;  $p = 0.012$ ,  $d = 0.59$ ), higher depression symptoms ( $p = 0.015$ ,  $d = 0.48$ ;  $p = 0.002$ ,  $d = 0.56$ ), and lower attachment ( $p = 0.038$ ,  $d = -0.45$ ;  $p < 0.001$ ,  $d = -0.77$ ). Further, compared to the High Parent Report group, the Very High Parent had lower initiative ( $p = 0.003$ ,  $d = -0.66$ )

### 3.3.2 | Early school age follow up

#### Omnibus results

At the early school age follow up, irritability pattern group predicted trauma symptoms and global functioning (Table 5).

**TABLE 6** Characteristics of elevated irritability groups versus low all group.

Time point	Low all versus:	High all	High parent report	Very high parent report
Preschool		Higher levels of ...	Higher levels of ...	Higher levels of ...
		Anxiety	Anxiety	Anxiety
		Trauma	Trauma	Trauma
		Lower levels of...	Lower levels of...	Depression
		Initiative skills	Global functioning	Lower levels of...
		Global functioning		Attachment
Early school age		Lower levels of...	No differences	Higher levels of...
		Global functioning		Trauma
				Lower levels of...
Preadolescence		Lower levels of...	Higher levels of...	Higher levels of...
		Global functioning	Anxiety	Depression
			Oppositional defiant	Anxiety
			ADHD	Trauma
			Lower levels of...	Oppositional defiant
			Global functioning	ADHD
			Conduct	
			Lower global functioning	

Note:  $p$ -values were corrected by time point with omnibus tests using the FDR correction.

Abbreviations: ADHD, attention-deficit hyperactivity; FDR, false discovery rate.



TABLE 7 Comparison among high parent-reported irritability groups.

Time point	Very high parent report > high all	Very high parent report > high parent report	High all > high parent report
Preschool	Worse global functioning	Worse global functioning	No differences
	Lower attachment	Lower attachment	
	Trauma symptoms	Lower initiative skills	
	Depressive	Trauma symptoms	
	Anxiety	Depressive	
		Anxiety	
Early school age	No differences	Trauma symptoms	No differences
Preadolescence	Depressive	Depressive	No differences
	Anxiety	Anxiety	
	Oppositional defiant		
	Conduct	Oppositional defiant	
	Trauma symptoms	Conduct	
	ADHD	Trauma symptoms	
		Worse Global functioning	

Note:  $p$ -values were corrected by time point with omnibus tests using the FDR correction.

Abbreviations: ADHD, attention-deficit hyperactivity; FDR, false discovery rate.

#### Low versus elevated irritability at early school age

As seen in Table 6, compared to the Low All group, the Very High Parent Report group had higher trauma symptoms ( $p < 0.001$ ,  $d = 0.78$ ). In addition, compared to the Low All group, the High All and Very High Parent Report groups had lower global functioning ( $p = 0.022$ ,  $d = -0.61$ ;  $p = 0.003$ ,  $d = -0.86$ ). There were no significant differences between the Low All and High Observation with Examiner groups.

#### Comparing high parent-reported irritability groups at early school age

As seen in Table 7, within the high parent-reported irritability groups, the Very High Parent Report group had higher trauma symptoms than the High Parent Report group ( $p = 0.012$ ,  $d = 0.51$ ).

### 3.3.3 | Preadolescent follow up

#### Omnibus results

At the preadolescent follow up, irritability pattern group predicted global functioning and mental health syndromes (Table 5).

#### Low versus elevated irritability at preadolescence

As seen in Table 6, the High All, High Parent Report, and Very High Parent report groups had lower global functioning compared to the Low All group ( $p = 0.044$ ,  $d = -0.57$ ;  $p = 0.039$ ,  $d = -0.36$ ;  $p < 0.001$ ,  $d = -0.75$ ). Further, the High Parent Report and Very High Parent Report groups had significantly higher generalized anxiety symptoms ( $p = 0.008$ ,  $d = 0.44$ ;  $p < 0.001$ ,  $d = 1.00$ ), separation anxiety symptoms ( $p = 0.043$ ,  $d = 0.37$ ;  $p < 0.001$ ,  $d = 1.07$ ), oppositional defiant symptoms ( $p < 0.001$ ,  $d = 0.65$ ;  $p < 0.001$ ,  $d = 1.50$ )

and ADHD symptoms ( $p = 0.004$ ,  $d = -0.48$ ;  $p < 0.001$ ,  $d = 0.92$ ) than the Low All group. Compared to the Low All group, the Very High Parent Report group also had higher trauma symptoms ( $p < 0.001$ ,  $d = 0.87$ ), depression symptoms ( $p = 0.002$ ,  $d = 0.71$ ), and conduct symptoms ( $p < 0.001$ ,  $d = 0.63$ ). There were no significant differences between the Low All and High Observation with Examiner groups.

#### Comparing high parent-reported irritability groups at preadolescence

As seen in Table 7, within the high parent-reported irritability groups, the Very High Parent Report group had lower global functioning ( $p = 0.035$ ,  $d = -0.48$ ) than the High Parent Report group. In addition, the Very High Parent Report group had higher trauma symptoms ( $p < 0.001$ ,  $d = 0.78$ ;  $p < 0.001$ ,  $d = 0.61$ ), generalized anxiety symptoms ( $p = 0.010$ ,  $d = 0.86$ ;  $p = 0.035$ ,  $d = 0.51$ ), separation anxiety symptoms ( $p < 0.001$ ,  $d = 1.12$ ;  $p < 0.001$ ,  $d = 0.78$ ), depression symptoms ( $p = 0.002$ ,  $d = 1.00$ ;  $p = 0.039$ ,  $d = 0.44$ ), oppositional defiant symptoms ( $p < 0.001$ ,  $d = 1.04$ ;  $p < 0.001$ ,  $d = 0.78$ ), and conduct symptoms ( $p < 0.001$ ,  $d = 0.58$ ;  $p = 0.001$ ,  $d = 0.42$ ), than the High All and High Parent Report groups. Further, the Very High Parent Report group had higher ADHD symptoms than the High All group ( $p = 0.014$ ,  $d = 0.72$ ).

## 4 | DISCUSSION

Our results suggest that informant method patterns can be informative, providing evidence that considering information from multiple sources may be necessary to predict anticipated problems and prognoses for childhood irritability. To improve clinical utility through parsimoniousness, we reduced individual scores from the

four measures of preschool irritability to five common patterns among subjects. Consolidating the data into useable patterns is necessary so that providers can create individualized recommendations and effectively create treatment plans based on associated symptoms and prognoses.

The results of this study support a two-tiered approach for evaluating preschool children's irritability: parent report measures, then observation. If a parent very highly endorses their child's irritability on the parent-reported irritability measures (MAPS-TL and PAPA, as in the Very High Parent Report group), the data show that the child is at significant risk, regardless of their score on the DB-DOS, making such additional observation unnecessary. Furthermore, if a parent's report of child irritability is low (low score on MAPS-TL and PAPA in Low All and High Observation with Examiner groups), then the data do not suggest significant levels of future impairment, regardless of their DB-DOS Examiner Context score, and thus further evaluation through observation is not warranted. However, if parent-reported irritability is only moderately elevated (High Parent Report, High All), then the DB-DOS Examiner Context can provide useful information. In this case, if the DB-DOS is also elevated (High All), the data suggest that there may be concurrent co-occurring mental health problems to assess. However, if the DB-DOS Examiner Context score is not elevated (High Parent Report), the data suggest that children may develop psychosocial impairments later on, even if they are not showing other concurrent symptomatology in preschool age.

Through reducing the data to latent classes, our results identify a few common patterns of child irritability, distinguishing between low irritability (Low All), transient irritability (High Observation with Examiner), and persistent, clinically significant, irritability (High All, High Parent Report, Very High Parent Report). While the DSM does not have a developmental definition nor guidance around irritability at preschool age, recent work has shown that misbehavior that could comprise a chronic, severe irritability syndrome is distinguishable in preschool age (Wakschlag et al., 2012; Wiggins et al., 2021). Identifying which misbehavior is transient and what is indicative of clinically significant problems is necessary to form early interventions for irritability during a time of peak neuroplasticity (Wakschlag et al., 2018). This paper adds to the growing body of evidence that, unlike the common "belief" that the high prevalence of misbehavior in preschool age makes the normal:abnormal distinction impossible, it is indeed possible to distinguish between transient and clinically significant irritability in preschool (Wiggins et al., 2018, 2021). Here we show that combining information from different measures can add to the precision of this determination. This study suggests that if an examiner observes irritability through a snapshot, yet the parent does not report a long history of irritability, then the irritability is likely transient, and the child's prognosis is better, versus children whose parents reported a history of irritability.

There were several limitations of this study. First, while the results of this study support that the Very High Parent Report group had continued impairment, it is important to note that only parent report is being used to assess preadolescent mental health outcomes.

As previous research has shown that informants with similar positions are more highly correlated (e.g., parent and parent), informants may also be more correlated with themselves across time. That is, if observed, but not parent reported, irritability is elevated at baseline, it could be that observation-based measures of symptomatology would have been elevated at follow-ups as well. This potential for shared method variance is inherent in this type of research, as it is challenging to include assessments not influenced by the reporter. However, future research should characterize these profiles across a greater variety of informants and modes of assessment. Second, this study relied on parent reports and observers' ratings of preschool irritability, limiting the contexts of measured irritability. Future research incorporating teachers' ratings into profile creation may allow for a more comprehensive measurement of preschool irritability; however, at preschool age many children are not enrolled in outside care, and thus teacher report is not always available. Third, the oldest group in this study is preadolescent. As many mental health problems emerge during adolescence, future research should examine further into adolescence/young adulthood to examine whether preschool irritability patterns are predictive of adolescent or adult psychopathology. Finally, while the overall sample size was 425, small sizes of the High All ( $n = 33$ ) and Very High Parent Report ( $n = 28$ ) groups limit the power to identify group differences between youth with high levels of preschool age irritability. Focused oversampling of these subgroups should be implemented to provide additional insight into how informant method patterns of preschool age irritability may be used to assess the risk for development of internalizing and externalizing symptoms.

Overall, while observed and parent reported measures of irritability had low correlation, informants and methods, combined, contributed to create informative and predictive classes of child irritability. These irritability pattern groups predicted future mental health concerns and suggest that severity of symptoms and when problems emerge may be forecasted by differences among these informant-method patterns. The results suggest that parent reports of irritability, both through interviews and surveys, are reliable and important indicators of future problems, and observations may provide additional insight when parent reports are not clear. Future research should examine how these profiles of observation and parent informed reports can be used to facilitate early intervention and improve long term outcomes associated with elevated irritability.

#### AUTHOR CONTRIBUTIONS

*Conceptualization and design:* Alyssa J. Parker, Maria Kryza-Lacombe, Margaret Briggs-Gowan, Lauren S. Wakschlag, Jillian Lee Wiggins. *Data analysis:* Alyssa J. Parker, Peyton Brock, Maria Kryza-Lacombe. *Data interpretation:* Alyssa J. Parker, Maria Kryza-Lacombe, Lea R. Dougherty, Jillian Lee Wiggins. *Data collection:* Margaret Briggs-Gowan, Lauren S. Wakschlag. *Writing – original draft:* Alyssa J. Parker, Peyton Brock. *Writing – review and editing:* Alyssa J. Parker, Maria Kryza-Lacombe, Margaret Briggs-Gowan, Lea R. Dougherty, Lauren S. Wakschlag, Jillian Lee Wiggins.

## ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Mental Health under award number 1R01MH082830 and 2U01MH082830 (PI: Wakschlag). We would like to thank participants in the MAPS study and Dr. Erica Anderson for oversight of interview measures.

## CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to report.

## DATA AVAILABILITY STATEMENT

Deidentified data available upon reasonable written request to Dr. Wakschlag for non-commercial purposes.

## ETHICS STATEMENT

Study protocols were approved by Northwestern University Feinberg School of Medicine's Institutional Review Board. Informed consent was given by legal guardians and children provided assent to participate.

## ORCID

Alyssa J. Parker  <https://orcid.org/0000-0001-5094-7196>

Jillian Lee Wiggins  <https://orcid.org/0000-0002-4931-6589>

## REFERENCES

- Ablow, J. C., Measelle, J. R., Kraemer, H. C., Harrington, R., Luby, J., Smider, N., Dierker, L., Clark, V., Dubicka, B., Heffelfinger, A., Essex, M. J., & Kupfer, D. J. (1999). The MacArthur Three-City Outcome Study: Evaluating multi-informant measures of young children's symptomatology. *Journal of the American Academy of Child & Adolescent Psychiatry*, 38(12), 1580–1590. <https://doi.org/10.1097/00004583-199912000-00020>
- Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science*, 15(2), 94–98. <https://doi.org/10.1111/j.0963-7214.2006.00414.x>
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101(2), 213–232. <https://doi.org/10.1037//0033-2909.101.2.213>
- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199–213). Springer.
- Ambrosini, P. J. (2000). Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (K-SADS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 49–58. <https://doi.org/10.1097/00004583-200001000-00016>
- Augenstein, T. M., Visser, K. F., Gallagher, K., De Los Reyes, A., D'Angelo, E. J., & Nock, M. K. (2022). Multi-informant reports of depressive symptoms and suicidal ideation among adolescent inpatients. *Suicide and Life-Threatening Behavior*, 52(1), 99–109. <https://doi.org/10.1111/sltb.12803>
- Briere, J., Johnson, K., Bissada, A., Damon, L., Crouch, J., Gil, E., Hanson, R., & Ernst, V. (2001). The Trauma Symptom Checklist for Young Children (TSCYC): Reliability and association with abuse exposure in a multi-site study. *Child Abuse & Neglect*, 25(8), 1001–1014. [https://doi.org/10.1016/s0145-2134\(01\)00253-8](https://doi.org/10.1016/s0145-2134(01)00253-8)
- Briggs-Gowan, M. J., Godoy, L., Heberle, A., & Carter, A. S. (2016). Chapter 1: Assessment of psychopathology in young children. In D. Cicchetti (Ed.), *Developmental psychopathology: Theory and method* (3rd ed., Vol. 1, pp. 1–93). Wiley.
- Brotman, M. A., Kircanski, K., Stringaris, A., Pine, D. S., & Leibenluft, E. (2017). Irritability in youths: A translational model. *American Journal of Psychiatry*, 174(6), 520–532. <https://doi.org/10.1176/appi.ajp.2016.16070839>
- Burchinal, M., Nelson, L., & Poe, M. (2006). Best practices in quantitative methods for developmentalists: IV. Growth curve analysis: An introduction to various methods for analyzing longitudinal data. *Monographs of the Society for Research in Child Development*, 71(3), i–145.
- Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The infant-toddler social and emotional assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, 31(5), 495–514. <https://doi.org/10.1023/a:1025449031360>
- Collishaw, S., Maughan, B., Natarajan, L., & Pickles, A. (2010). Trends in adolescent emotional problems in England: A comparison of two national cohorts twenty years apart. *Journal of Child Psychology and Psychiatry*, 51(8), 885–894. <https://doi.org/10.1111/j.1469-7610.2010.02252.x>
- De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, 40(1), 1–9. <https://doi.org/10.1080/15374416.2011.533405>
- De Los Reyes, A., & Epkins, C. C. (2023). Introduction to the special Issue. A dozen years of demonstrating that informant discrepancies are more than measurement error: Toward guidelines for integrating data from multi-informant assessments of youth mental health. *Journal of Clinical Child and Adolescent Psychology*, 52(1), 1–18. <https://doi.org/10.1080/15374416.2022.2158843>
- De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, 37(5), 637–652. <https://doi.org/10.1007/s10802-009-9307-3>
- Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M., Cella, D., & Wakschlag, L. S. (2012). Annual Research Review: Embracing not erasing contextual variability in children's behavior – theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 53(5), 558–574. <https://doi.org/10.1111/j.1469-7610.2012.02537.x>
- Dougherty, L. R., Galano, M. M., Chad-Friedman, E., Olino, T. M., Bufferd, S. J., & Klein, D. N. (2020). Using item response theory to compare irritability measures in early adolescent and childhood samples. *Assessment*, 28(3), 918–927. <https://doi.org/10.1177/107319120936363>
- Dougherty, L. R., Smith, V. C., Bufferd, S. J., Stringaris, A., Leibenluft, E., Carlson, G. A., & Klein, D. N. (2013). Preschool irritability: Longitudinal associations with psychiatric disorders at age 6 and parental psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(12), 1304–1313. <https://doi.org/10.1016/j.jaac.2013.09.007>
- Durham, T. A., Byllesby, B. M., Lv, X., Elhai, J. D., & Wang, L. (2018). Anger as an underlying dimension of posttraumatic stress disorder. *Psychiatry Research*, 267, 535–540. <https://doi.org/10.1016/j.psychres.2018.06.011>
- Egger, H. L., Erkanli, A., Keeler, G., Potts, E., Walter, B. K., & Angold, A. (2006). Test-retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of the American Academy of Child & Adolescent Psychiatry*, 45(5), 538–549. <https://doi.org/10.1097/01.chi.0000205705.71194.b8>
- Evans, S. C., Burke, J. D., Roberts, M. C., Fite, P. J., Lochman, J. E., de la Peña, F. R., & Reed, G. M. (2017). Irritability in child and adolescent

- psychopathology: An integrative review for ICD-11. *Clinical Psychology Review*, 53, 29–45. <https://doi.org/10.1016/j.cpr.2017.01.004>
- Evans, S. C., Corteselli, K. A., Edelman, A., Scott, H., & Weisz, J. R. (2022). Is irritability a top problem in youth mental health care? A multi-informant, multi-method investigation. *Child Psychiatry and Human Development*, 54(4), 1027–1041. <https://doi.org/10.1007/s10578-021-01301-8>
- Finlay-Jones, A. L., Ang, J. E., Brook, J., Lucas, J. D., MacNeill, L. A., Mancini, V. O., Kottampally, K., Elliott, C., Smith, J. D., & Wakschlag, L. S. (2023). Systematic review and meta-analysis: Early irritability as a transdiagnostic neurodevelopmental vulnerability to later mental health problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, 63(2), 184–215. <https://doi.org/10.1016/j.jaac.2023.01.018>
- Kerr, D. C. R., Lunkenheimer, E. S., & Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: A longitudinal study from preschool to school entry. *Journal of Child Psychology and Psychiatry*, 48(10), 967–975. <https://doi.org/10.1111/j.1469-7610.2007.01776.x>
- Klein, D. N., Dougherty, L. R., Kessel, E. M., Silver, J., & Carlson, G. A. (2021). A transdiagnostic perspective on youth irritability. *Current Directions in Psychological Science*, 30(5), 437–443. <https://doi.org/10.1177/096372142111035101>
- Krogh-Jespersen, S., MacNeill, L. A., Anderson, E. L., Stroup, H. E., Harriott, E. M., Gut, E., Blum, A., Fareedi, E., Fredian, K. M., Wert, S. L., Wakschlag, L. S., & Norton, E. S. (2022). Disruption leads to methodological and analytic innovation in developmental sciences: Recommendations for remote administration and dealing with messy data [review]. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.732312>
- LeBuffe, P. A., & Naglieri, J. A. (1999). The Devereux Early Childhood Assessment (DECA): A measure of within-child protective factors in preschool children. *NHSA Dialog*, 3(1), 75–80. [https://doi.org/10.1207/s19309325nhsa0301\\_10](https://doi.org/10.1207/s19309325nhsa0301_10)
- Leibenluft, E. (2017). Irritability in children: What we know and what we need to learn. *World Psychiatry*, 16(1), 100–101. <https://doi.org/10.1002/wps.20397>
- Lerner, M. D., De Los Reyes, A., Drabick, D. A. G., Gerber, A. H., & Gadaw, K. D. (2017). Informant discrepancy defines discrete, clinically useful autism spectrum disorder subgroups. *Journal of Child Psychology and Psychiatry*, 58(7), 829–839. <https://doi.org/10.1111/jcpp.12730>
- Lind, T., Bernard, K., Ross, E., & Dozier, M. (2014). Intervention effects on negative affect of CPS-referred children: Results of a randomized clinical trial. *Child Abuse & Neglect*, 38(9), 1459–1467. <https://doi.org/10.1016/j.chiabu.2014.04.004>
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767–778. <https://doi.org/10.1093/biomet/88.3.767>
- Makol, B. A., De Los Reyes, A., Ostrander, R. S., & Reynolds, E. K. (2019). Parent-youth divergence (and convergence) in reports of youth internalizing problems in psychiatric inpatient care. *Journal of Abnormal Child Psychology*, 47(10), 1677–1689. <https://doi.org/10.1007/s10802-019-00540-7>
- Mallidi, A., Meza-Cervera, T., Kircanski, K., Stringaris, A., Brotman, M. A., Pine, D. S., Leibenluft, E., & Linke, J. O. (2023). Robust caregiver-youth discrepancies in irritability ratings on the affective reactivity index: An investigation of its origins. *Journal of Affective Disorders*, 332, 185–193. <https://doi.org/10.1016/j.jad.2023.03.091>
- Measelle, J. R., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (1998). Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Development*, 69(6), 1556–1576. <https://doi.org/10.1111/j.1467-8624.1998.tb06177.x>
- Morris, A. S., Wakschlag, L., Krogh-Jespersen, S., Fox, N., Planalp, B., Perlman, S. B., Shuffrey, L. C., Smith, B., Lorenzo, N. E., Amso, D., Coles, C. D., & Johnson, S. P. (2020). Principles for guiding the selection of early childhood neurodevelopmental risk and resilience measures: HEALthy Brain and Child Development Study as an Exemplar. *Adversity and Resilience Science*, 1(4), 247–267. <https://doi.org/10.1007/s42844-020-00025-3>
- Petitclerc, A., Briggs-Gowan, M. J., Estabrook, R., Burns, J. L., Anderson, E. L., McCarthy, K. J., & Wakschlag, L. S. (2015). Contextual variation in young children's observed disruptive behavior on the DB-DOS: Implications for early identification. *Journal of Child Psychology and Psychiatry*, 56(9), 1008–1016. <https://doi.org/10.1111/jcpp.12430>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Roy, A. K., & Comer, J. S. (2020). Advances in the conceptualization, assessment, and treatment of pediatric irritability. *Behavior Therapy*, 51(2), 207–210. <https://doi.org/10.1016/j.beth.2019.10.010>
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(MR0468014), 461–464. <https://doi.org/10.1214/aos/1176344136>
- Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52(3), 333–343. <https://doi.org/10.1007/bf02294360>
- Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry*, 40(11), 1228–1231. <https://doi.org/10.1001/archpsyc.1983.01790100074010>
- Sorcher, L. K., Goldstein, B. L., Finsaas, M. C., Carlson, G. A., Klein, D. N., & Dougherty, L. R. (2022). Preschool irritability predicts adolescent psychopathology and functional impairment: A 12-year prospective study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 61(4), 554–564.e551. <https://doi.org/10.1016/j.jaac.2021.08.016>
- Stoddard, J., Stringaris, A., Brotman, M. A., Montville, D., Pine, D. S., & Leibenluft, E. (2014). Irritability in child and adolescent anxiety disorders. *Depression and Anxiety*, 31(7), 566–573. <https://doi.org/10.1002/da.22151>
- Stringaris, A., Cohen, P., Pine, D. S., & Leibenluft, E. (2009). Adult outcomes of youth irritability: A 20-year prospective community-based study. *American Journal of Psychiatry*, 166(9), 1048–1054. <https://doi.org/10.1176/appi.ajp.2009.08121849>
- 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(7), 556–570. <https://doi.org/10.1016/j.jaac.2016.04.014>
- Wakschlag, L., Carroll, A., Friedland, S., Walkup, J., Wiggins, J., Mohanty, N., Papacheck, E., Bridi, S., Carrol R., Drelicharz, D., Hasan, Z., Kotagal, T., Davis, M., & Smith, J. (2023). Making it “EASI” for pediatric clinicians to determine when toddler tantrums are “more than the terrible twos”: Proof-of-concept for primary care screening with the Multidimensional Assessment Profiles–Early Assessment Scale for Irritability (MAPSEASI). *Submitted manuscript*.
- Wakschlag, L. S., Briggs-Gowan, M. J., Choi, S. W., Nichols, S. R., Kestler, J., Burns, J. L., Carter, A. S., & Henry, D. (2014). Advancing a multidimensional, developmental spectrum approach to preschool disruptive behavior. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(1), 82–96.e83. <https://doi.org/10.1016/j.jaac.2013.10.011>
- Wakschlag, L. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Leventhal, B. L., Keenan, K., Egger, H. L., Cicchetti, D., Burns, J., & Carter, A. S. (2008). Observational assessment of preschool disruptive behavior, Part II: Validity of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(6), 632–641. <https://doi.org/10.1097/CHI.0b013e31816c5c10>
- Wakschlag, L. S., Choi, S. W., Carter, A. S., Hullsiek, H., Burns, J., McCarthy, K., Leibenluft, E., & Briggs-Gowan, M. J. (2012). Defining the developmental parameters of temper loss in early childhood: Implications

- for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 53(11), 1099–1108. <https://doi.org/10.1111/j.1469-7610.2012.02595.x>
- Wakschlag, L. S., Estabrook, R., Petittclerc, A., Henry, D., Burns, J. L., Perlman, S. B., Voss, J. L., Pine, D. S., Leibenluft, E., & Briggs-Gowan, M. L. (2015). Clinical implications of a dimensional approach: The normal:abnormal spectrum of early irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(8), 626–634. <https://doi.org/10.1016/j.jaac.2015.05.016>
- Wakschlag, L. S., Hill, C., Carter, A. S., Danis, B., Egger, H. L., Keenan, K., Leventhal, B. L., Cicchetti, D., Maskowitz, K., Burns, J., & Briggs-Gowan, M. J. (2008). Observational assessment of preschool disruptive behavior, Part I: Reliability of the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(6), 622–631. <https://doi.org/10.1097/CHI.0b013e31816c5bdb>
- Wakschlag, L. S., Krogh-Jespersen, S., Estabrook, R., Hlutkowsky, C. O., Anderson, E. L., Burns, J., Briggs-Gowan, M. J., Petittclerc, A., & Perlman, S. B. (2020). The Early Childhood Irritability-Related Impairment Interview (E-CRI): A Novel method for assessing young children's developmentally impairing irritability. *Behavior Therapy*, 51(2), 294–309. <https://doi.org/10.1016/j.beth.2019.07.008>
- Wakschlag, L. S., Leventhal, B. L., Briggs-Gowan, M. J., Danis, B., Keenan, K., Hill, C., Egger, H. L., Cicchetti, D., & Carter, A. S. (2005). Defining the "disruptive" in preschool behavior: What diagnostic observation can teach us. *Clinical Child and Family Psychology Review*, 8(3), 183–201. <https://doi.org/10.1007/s10567-005-6664-5>
- Wakschlag, L. S., Perlman, S. B., Blair, R. J., Leibenluft, E., Briggs-Gowan, M. J., & Pine, D. S. (2018). The neurodevelopmental basis of early childhood disruptive behavior: Irritable and callous phenotypes as exemplars. *American Journal of Psychiatry*, 175(2), 114–130. <https://doi.org/10.1176/appi.ajp.2017.17010045>
- Wiggins, J. L., Briggs-Gowan, M. J., Brotman, M. A., Leibenluft, E., & Wakschlag, L. S. (2021). Toward a developmental nosology for disruptive mood dysregulation disorder in early childhood. *Journal of the American Academy of Child & Adolescent Psychiatry*, 60(3), 388–397. <https://doi.org/10.1016/j.jaac.2020.04.015>
- Wiggins, J. L., Briggs-Gowan, M. J., Estabrook, R., Brotman, M. A., Pine, D. S., Leibenluft, E., & Wakschlag, L. S. (2018). Identifying clinically significant irritability in early childhood. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(3), 191–199.e192. <https://doi.org/10.1016/j.jaac.2017.12.008>
- Wiggins, J. L., Roy, A. K., & Wakschlag, L. S. (2023). MAPping affective dimensions of behavior: Methodologic and pragmatic advancement of the Multidimensional Assessment Profiles scales. *International Journal of Methods in Psychiatric Research*, 32(S1). <https://doi.org/10.1002/mpr.1990>
- Wiggins, J. L., Ureña Rosario, A., Macneill, L. A., Krogh-Jespersen, S., Briggs-Gowan, M., Smith, J. D., & Wakschlag, L. S. (2023). Prevalence, stability, and predictive utility of the Multidimensional Assessment of Preschoolers Scales clinically optimized irritability score: Pragmatic early assessment of mental disorder risk. *International Journal of Methods in Psychiatric Research*, 32(S1). <https://doi.org/10.1002/mpr.1991>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Parker, A. J., Brock, P., Kryza-Lacombe, M., Briggs-Gowan, M., Dougherty, L. R., Wakschlag, L. S., & Wiggins, J. L. (2024). What I see, what you say: How cross-method variation sharpens characterization of irritability in early childhood. *International Journal of Methods in Psychiatric Research*, e2019. <https://doi.org/10.1002/mpr.2019>