UNIVERSITY OF CALIFORNIA

Los Angeles

Evidenced-based Assessment and Multivariate Risk Profiles of Pediatric Irritability

A dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy in Psychology

by

Felix Ka Kai So

2024

ABSTRACT OF THE DISSERTATION

Evidenced-based Assessment and Multivariate Risk Profiles of Pediatric Irritability

by

Felix Ka Kai So

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2024

Professor Steve Sung-Yul Lee, Chair

Adolescence is a period of rapid neurological, social, and developmental change which increases vulnerability to psychopathology onset (Bitsko et al., 2018; Ghandour et al., 2019). Dimensions of child psychopathology incur substantial public health, educational, and economic costs spanning high rates of suicide, increased risk for juvenile justice involvement, and physical health problems (Birmaher & Brent, 2007). To accelerate new knowledge for evidence-based assessment, there is an urgent need to identify modifiable childhood risk factors for adolescent socio-emotional development.

Characterized by excessive reactivity to negative emotional stimuli, childhood irritability is common and developmentally normative (e.g., preschool) (Copeland et al., 2015). Despite its transdiagnostic relevance, the knowledgebase of childhood irritability is underdeveloped, particularly surrounding multi-informant ratings of irritability as well as risk factors for irritability. Study I leveraged data from a well-characterized sample of children with (n = 120) and without (n = 110) ADHD followed prospectively from childhood through early adolescence to evaluate parent and teacher combinations of childhood irritability in their prediction of psychopathology

and functional impairment. Receiver operator characteristic (ROC) analyses assessed the association of informant combinations of childhood irritability with key outcomes to inform clinical decision making. Findings suggested that parent-rated and parent “or” teacher rated irritability significantly predicted CBCL Internalizing and Externalizing problems and functional impairments. ROC revealed that parent only ratings of irritability were significantly more accurate than parent “or” teacher report. Study II utilized the Adolescent Brain Cognitive Development (ABCD) study, a nationally representative sample followed prospectively across three years starting in childhood (ages 9-10) through early adolescence (age 12-13). We employed latent profile analysis to empirically identify risk profiles based dimensions of temperament (i.e., effortful control, surgency, and negative affect) and aspects of the family environment (i.e., family expression, cohesion, and conflict) and their association with childhood irritability. Findings suggested that a 3-profile model with an unrestricted/varying variance/covariance matrix was an appropriate fit for the data. Collectively, these findings accelerate the evidence-base assessment of childhood irritability, informing future research designs that can continue building an evidence base for the assessment of childhood irritability and its multivariate risk profiles.

The dissertation of Felix Ka Kai So is approved.

Han Du

Lauren Christina Ng

Tara S. Peris

Steve Sul-Yul Lee, Committee Chair

University of California, Los Angeles

2024

Table of Contents

List of Tables and Figures

Study I

Study II

Acknowledgments

Study I utilized prospective longitudinal data from the UCLA ADHD and Development Laboratory (PI, Lee), which was supported NIH and intramural funds. Study II used public use data from the Adolescent Brain and Cognitive Development (ABCD) Study.

I want to give special thanks to a few people who supported me throughout this journey. I am thankful to my mentor, Dr. Steve Lee for his assistance and guidance throughout my graduate career at UCLA. I would also like to thank my dissertation committee, Dr. Han Du, Dr. Lauren Ng, and Dr. Tara Peris, for their expertise and insight throughout this research project. I also appreciate all the support and advice from my UCLA cohort, CHLA cohort, and the Lee Lab members over the years. Finally, I want to express my utmost gratitude to my parents, sister, partner, and closest friends. I could not have done this without the continued support throughout the years.

Vita

EDUCATION

2020 **University of California, Los Angeles**
Master of Arts in Psychology

2017 **University of Illinois at Urbana-Champaign**
Bachelor of Science in Psychology with Honors (Clinical/Community Concentration)

FELLOWSHIPS AND AWARDS

2021 Summer Mentored Research Fellowship
2020 Graduate Summer Research Mentorship Fellowship
2019 Psychology Graduate Summer Research Mentorship

PUBLICATIONS

So, F. K*., Chavira, D., & Lee, S. S., (2021). ADHD and ODD dimensions time varying predictions of internalizing problems from childhood to adolescence. *Journal of Attention Disorders.* DOI: 10870547211050947.

Cohen, J.R., **So, F.K*.**, J.F., Hankin, B.L., & Lee, B., Young (2019). Youth depression screening with parents and self-reports: Assessing current and prospective depression risk. *Child Psychiatry & Human Development, 50(4),* 647-660.

Cohen, J.R., **So, F.K*.,** Hankin, B.L., & Young, J.F. (2017). Translating cognitive vulnerability theory into improved depression screening: A receiver operating characteristic approach. *Journal of Clinical Child & Adolescent Psychology*, *48(4)*, 582-595.

SELECTED CONFERENCE PRESENTATIONS

So, F*., Nguyen, H., & Lee, S. S., (2021, April) Dimensions of childhood irritability and predictions of adolescent psychopathology. Poster for the 2021 Society for Research on Child Development. Virtual.

So, F*., Zhang, J., Nguyen, H., & Lee, S. S., (2019, June) Predicting depression and anxiety in children with ADHD. Poster for the 2019 International Research on Child and Adolescent Psychopathology. Los Angeles, CA.

Study I: A Receiver Operator Characteristic Analysis of Optimal Informants for Childhood Irritability

The US Surgeon General's Office (2021) recently designated youth mental health problems as a "national state of emergency." Exacerbated by the COVID-19 pandemic, youth mental health problems are increasingly prevalent, and they uniquely and significantly burden families and society more generally. Youth psychopathology diversely predicts poor developmental and health outcomes spanning suicidal ideation and self-harm to substandard academic achievement and disrupted social relationships (Aganafors et al., 2020, Gloster et al., 2020; Madigan et al., 2023). These epidemiological trends must be contextualized by distressing trends with respect to the availability and use of evidence-based mental health resources; even gold standard interventions are often underutilized, variably effective, and are embedded in racial-ethnic and socioeconomic disparities (Costello et al., 2014). For example, treatment responses (e.g., cognitive-behavioral therapy; Hoffmann et al., 2012) are highly variable secondary to pre-treatment factors (e.g., hopelessness, readiness for change, and global functioning) (Lewis et al., 2012), with few studies evaluating long term treatment outcomes (van Dis et al., 2020). To disrupt escalations in the prevalence and burden of youth mental health problems, improved knowledge of its etiology, causal influences, underlying risk processes, as well as improved assessment strategies are a top priority (Hunsley & Mash, 2007; Kraemer et al., 1997).

Characterized by excessive reactivity to emotional stimuli (Copeland et al., 2015), irritability consists of correlated (yet separable) behavioral (e.g., aggression) and affective (e.g., anger) dimensions (Leibenluft & Stoddard, 2013). Diagnostically, irritability is a core symptom across internalizing (e.g., disruptive mood dysregulation disorder (DMDD) and generalized anxiety disorder (GAD)) and externalizing (e.g., oppositional defiant disorder (ODD)) disorders. Separate from its diagnostic relevance, irritability has clinical utility, including constituting a focus in treatment planning and clinical care more generally (Evans et al., 2017). Developmentally,

childhood irritability is normative, yet early elevations independently and uniquely predict clinical dysfunction and impairment, even with control of demographic covariates (e.g., age, sex) and baseline clinical problems (Copeland et al., 2015; Sorcher et al., 2021). Across youth and parent report in separate community and clinical samples, irritability is concurrently and prospectively associated with internalizing (e.g., anxiety and depression) and externalizing (e.g., disruptive-behavior disorders) psychopathology as well as poor functional outcomes (e.g., peer functioning and physical health) (Dougherty et al., 2015; Humphreys et al., 2019; Sorcher et al., 2021; Stoddard et al., 2014). Finally, further attesting to its clinical significance, irritability was rated as the most prevalent problem by parents and children among treatment-seeking families (Evans et al., 2022). Taken together, there is persuasive evidence that childhood irritability is clinically significant and transdiagnostically predicts vulnerability to the development of later psychopathology.

Despite its clinical utility and construct validity, important limitations must be addressed. For example, there is increasing recognition that item overlap likely inflates associations of temperament and personality with psychopathology (Lahey, 2009). Perhaps similarly, previous associations of irritability with internalizing and externalizing problems may also be inflated due to item overlap. To strengthen inferences about independent associations of irritability with psychopathology – a key criterion – reducing item overlap in tests of construct validity must be prioritized. Although irritability meta-analytically and prospectively predicted psychopathology, with meaningful effect sizes (e.g., odds ratios (OR) of 2.62 (CI [1.41, 4.85]) for oppositional defiant disorder (ODD) to 1.80 (CI [1.42, 2.27]) and 1.72 (CI [1.31, 2.26]) for depression and anxiety, respectively), irritability shares similar language with ODD including “temper outbursts” and “angry and resentful.” Thus, it is unclear if these putative effects are meaningful or if they are spurious secondary to shared item content (Vidal-Ribas et al., 2016). In a key exception, however, Dougherty et al. (2013) compared predictive models of irritability with “nonoverlapping” (i.e., scales created by specifically removing items related to irritability) dimensions of

depression, anxiety, ADHD, and ODD; interestingly, models with and without overlapping items significantly predicted depressive and ODD symptoms, even with control of baseline outcomes. These patterns suggest irritability uniquely predicts the development of psychopathology even with conservative adjustment for shared item/content.

Despite its clinical significance, problematically, there is relatively little empirical work to guide the evidence-based assessment of childhood irritability; this knowledge gap is particularly acute relative to the more well-developed evidence base on the assessment of other dimensions of child psychopathology (e.g., ADHD). A major obstacle to innovating the assessment of irritability is how to balance two competing lines of evidence: irritability is normative, especially early in development, yet elevations also predict clinically significant outcomes. Thus, assessment must thoughtfully balance the goals of early detection (and service provision) with the risk of falsely pathologizing developmentally transient expressions of irritability. Presently, childhood irritability is typically assessed using specific rating scales (e.g., Affective Reactivity Index – ARI; Stringaris et al., 2012), subscales derived from broader measures of child psychopathology (e.g., Child Behavior Checklist and Teacher Report Form – CBCL and TRF; Achenbach et al., 2001), and selected items from diagnostic interviews (e.g., Kiddie Schedule for Affective Disorders and Schizophrenia; Kaufman et al., 1997). Given that principles of evidence-based assessment – particularly in childhood – prioritize collection of multi-informant data, it is puzzling that there is minimal evidence to guide how multi-source (irritability) data should be integrated; instead, existing guidelines have mostly focused on which informant is optimal for youth at specific ages (Dougherty et al., 2021; Zik et al., 2021). To accelerate innovations in assessment (e.g., early detection), which will necessarily improve outcomes, there is an urgent need to evaluate different combinations of multi-informant childhood irritability data and their clinical utility with respect to meaningful external criteria.

Despite the centrality of multi-informant data collection to evidence-based assessment (Hunsley & Mash, 2007; Kraemer et al., 2003), there is a sizable gap in evidence on how to

reliably integrate multi-informant data. This gap is problematic given longstanding evidence on modest inter-rater reliability for common youth problems (de Los Reyes et al., 2015), which reflects situational specificity (i.e., behaviors vary according to situational context) as well as the amount and type of information available to informants (Alexander et al., 2017). Informant ratings are also developmentally-sensitive as evidenced by the psychometric superiority of parents and teachers with respect to disruptive behaviors (e.g., ADHD) in young children whereas adolescent self-report of internalizing problems significantly increases clinical utility (Dowdy & Kim, 2012). Crucially, the use of specific assessment methods (e.g., informant, instrumentation), including how multi-source data are combined, dramatically affects clinical inferences (e.g., referral for services) (de Los Reyes & Kazdin, 2005). For example, ADHD caseness varied dramatically based on use of the “and” (i.e., symptom is indicated if all the informants endorse it) vs. “or” (i.e., symptom is indicated if any the informants endorse it) rule algorithm with respect to parent and teacher report (Piacentini et al., 1992; Shemmassian & Lee, 2017). To date, little is known about the psychometric properties and clinical utility of informant combinations for pediatric irritability; in fact, many studies employ single informants (Wiggins et al., 2018). Even when multi-informant data are collected, they are often treated separately (Orri et al., 2019). Therefore, to maximize its potential clinical utility, combinations of multi-informant assessment data for childhood irritability (i.e., parents, teachers) must be evaluated using methods that detail their associations with independent criteria including predictions of psychopathology and functional outcomes.

Given that combining multi-informant data is associated with diagnostic errors (Swets et al., 2000; Youngstrom, 2013), evaluating specific informant combinations requires tests that are statistically and clinically informed. Algorithms must consider rates of false negatives (i.e., incorrect assessment that the trait is absent), false positives (i.e., inaccurate assessment that the trait is present), true negatives (i.e., accurate assessment that the trait is absent), and true positives (i.e., accurate assessment that the trait is present); collectively, their patterns

contribute to sensitivity (i.e., determining the presence of trait) and specificity (i.e., ability to determining the absence of a trait) (Saah & Hoover, 1997). Although the association of a risk factor with a discrete outcome yields an odds ratio (Orri et al., 2019; Stringaris et al., 2009), this does not inherently convey accuracy nor its differentiation of caseness (i.e., individuals with vs. without a condition; Pepe et al., 2004). However, receiver operating characteristic (ROC) analyses estimate classification accuracy and inform decision-making using clinical guidelines for significance (Swets, 1988; Youngstrom, 2014). For instance, ROC is routinely used to assess diagnostic accuracy for rating scales and interviews (Chen et al., 1994), to estimate algorithms of risk factors with respect to clinical diagnosis (Cohen et al., 2018), and to evaluate specific informant combinations (Lapalme et al., 2020; Shemmassian & Lee, 2017). ROC analyses graphically represent the true positive rates (sensitivity) as a function of the false positive rates (1 – specificity) which yield a ROC curve. The area under the curve (AUC) estimates diagnostic accuracy as well as clinical and statistical significance (Swets, 1988). To evaluate accuracy, we calculated AUCs and evaluated them for each informant combination; this approach has yet to be employed for informant algorithms for irritability and with respect to their association with relevant criteria. This study addresses these limitations directly.

Aims & Hypotheses

Aim 1a: To review, multiple-informant combinations have been evaluated for multiple dimensions of child psychopathology but not for pediatric irritability. To improve traction on the evidence-base assessment of childhood irritability, the current study prospectively followed 230 children (98% of youth was six to nine years old) with (n = 120) and without DSM-IV ADHD (n = 110) for six to seven years from childhood to early adolescence. Multilevel longitudinal logistic regression models separately evaluated four methods of combining parent and teacher ratings of childhood irritability: (a) parent only, (b), teacher only, (c) parent “and” teacher, and (d) parent “or” teacher. These methods were subjected to tests of predictive validity with respect to

independent measures of adolescent internalizing and externalizing problems as separately rated by parents, in addition to measures of functional impairment that were independently rated by parents and clinicians. Given the literature regarding optimal informants, individually, I hypothesized that parent report only would be positively associated with clinical elevations for internalizing and externalizing problems as well as functional impairment. On the other hand, I hypothesized that teacher report only would be positively associated with clinically elevated externalizing problems and functional impairment, given evidence that teachers may underreport youth internalizing symptoms (Tandon et al., 2009). I also hypothesized that combining parent and teacher ratings of childhood irritability – based on the "or" rule – would positively predict clinically elevated internalizing and externalizing problems and functional impairment.

Aim 1b. To minimize inflated associations, I reproduced analyses from Aim 1a with and without item overlap. That is, three items reflecting irritability in the outcome variables were removed and predictions from irritability were re-evaluated. I hypothesized that irritability would continue to significantly predict psychopathology (see Dougherty et al., 2013), although effects may show some attenuation.

Aim 2: Based on significant models Aim 1a, ROC curves were plotted to visualize the sensitivity of irritability algorithm as a function of false positive rates. AUC were estimated from ROC curves to quantify the accuracy of informant combinations and clinical benchmarks were used to determine both clinical and statistical significance (Swets et al., 2000). Finally, AUCs were tested against each other to determine the informant combinations with the highest accurate (Cleves, 2002). I hypothesized that parent and teacher ratings of childhood irritability integrated with the "or" rule would be the most accurate assessment approach.

Methods

Participants

230 ethnically diverse (50% Caucasian) youth aged five to ten (*M* =7.4, *SD* = 1.1, 68% male) with (*n* = 120) and without (*n* = 110) were recruited from a large metropolitan city in the Western US. Children were recruited through flyers posted in local elementary schools as well as referrals from local mental health centers and pediatric offices. Inclusion criteria consisted of English fluency, the child residing with at least one biological parent at least half the time, and full-time enrollment in school. Exclusion criteria consisted of youth IQ below 70, a neurological, pervasive developmental, or seizure disorder that prevented full participation in study activities. At baseline (i.e., Wave 1), there were no significant diagnostic group differences with respect to age, sex, race, or SES among youth with and without ADHD. Table 1.1 summarizes additional demographic and clinical characteristics of the participants separately at each wave.

Procedures

At Wave 1, initial study eligibility was determined through a telephone screener. Families who satisfied the inclusion criteria were invited to complete a laboratory-based assessment and were mailed rating scales to complete prior to the assessment. After obtaining parental consent and youth assent, parents and children were interviewed simultaneously in different rooms by separate interviewers, usually consisting of clinical psychology Ph.D. students and B.A.-level staff. All interviewers completed multi-day trainings in evidence-based clinical assessment with children and families led by the PI. Parents completed a fully structured diagnostic interview and rating scales related to parental psychopathology, family functioning, and children's socio-emotional development. Youth completed similar procedures (i.e., interviews, rating scales) probing psychopathology.

Two years after Wave 1, all families were asked to participate in a laboratory-based follow up visit (i.e., Wave 2). At Wave 2, 206 (89%) of the original sample participated in the follow up assessment (*M* = 9.7, *SD* = 1.3) and a third follow up assessment (i.e., Wave 3) was

conducted 2 to 3 years after Wave 2. At Wave 3, 183 of the original sample (79% retention) participated in the follow up assessment (*M* = 12.1, *SD* = 1.4). Wave 2 and 3 procedures consisted of parallel methods including laboratory-based assessment of parent and child psychopathology, family functioning, and peer relationships. Developmentally appropriate expansion into substance use and delinquency domains were included at Wave 2 and 3. Teachers were asked to complete teacher-rated forms at all waves of the study. All procedures were approved by the UCLA Institutional Review Board.

Measures

Childhood Irritability

The Disruptive Behavior Disorder (DBD; Pelham et al., 1992) rating scale is a 45-item rating scale of DSM-IV child ADHD, ODD, and CD symptoms completed by both parents and teachers. Ratings ranged from 0 to 3 (0 = *not at all*, 1 = *just a little*, 2 = *pretty much*, and 3 = *very much*). Symptoms endorsed as a 2 or 3 are considered present; for the "or" rule, the symptom was considered present if endorsed by either the parent or teacher whereas the "and" rule required that the symptom be endorsed as a 2 or 3 by both informants. Irritability will be estimated from the sum of three items: (1) is often angry and resentful, (2) is often touchy or easily annoyed by others, and (3) often loses temper (Stringaris & Goodman, 2009). Thus, each child had a total irritability score that ranged from 0 to 3 based on the four informant combinations described above.

Internalizing and Externalizing Problems

The Child Behavior Checklist (CBCL; Achenbach et al., 2001) is a 113-item parent rating scales that assesses clinical problems and adaptive functioning in youth over the previous six months. Items were scored on a 0 to 2 Likert scale (0 = *not true*, 1 = *somewhat true*, and 2 = *very or often true*) and raw scores were converted to T-scores based on the nationally representative normative sample. CBCL T-scores for Internalizing Problems and Externalizing Problems broadbands will be recoded into a dichotomous variable using a T-score of 64 to

determine clinical elevation (Nelson et al., 2016). Scores greater than or equal to 64 were coded as 1, which represents clinically significant problems, whereas scores below 64 were coded as 0, which represents normative levels. Full breakdown of distribution is located on Table 1.3.

Adjusted Outcomes Model

To combat item overlap and potential inflated associations, two separate outcome variables were calculated. First, the raw score sum of the CBCL Externalizing Problems scale was used. A second "adjusted" CBCL externalizing score was calculated with three overlapping items removed: "stubborn, sullen, or irritable," "sudden changes in mood or feelings," and "temper tantrums or hot temper" (Evans et al., 2019). The Adjusted Externalizing Problems thus reflects a raw score sum relatively free of contamination with irritability items. Because this adjustment prohibits inferences of clinical significance otherwise inferred from the normed T score, we instead calculated quartiles; specifically, we dichotomized adolescents in the highest quartile versus the remaining lowest three quartiles where 1 = elevated and 0 = subthreshold/normal (Lee & Hinshaw, 2004). Notably, the Adjusted Externalizing Problems and Unadjusted Externalizing Problems scale were only evaluated as part of Aim 1a.

Functional impairment

The Children's Global Assessment Scale (CGAS; (Shaffer et al., 1983) estimates a child's lowest level of overall functioning during the past 6 months completed by the clinician and parent. Scores ranged from 0 (most impaired) to 100 (no impairment). Parents and interviewers, following completion of a structured diagnostic interview, were asked to consider the child's emotional and behavioral functioning at home, school, with friends, and during leisure time. Parents and interviewers separately designed a score for each child; a cutoff score of 70 was used to estimate clinical significance given that this threshold previously discriminated normal functioning from severe problems where treatment may be warranted (Steinhausen, 1987). Scores greater than or equal to 70 suggested normal functioning (score = 0) whereas

scores less than 70 suggested functional impairment in at least one domain (e.g., home, school, social; score = 1). Full breakdown of distribution is located on Table 1.3.

Data Analytic Plan

Multilevel logistic regressions analyzed the prospective association of informant combinations – parent only, teacher only, parent "and" teacher, and parent "or" teacher – with respect to clinically elevated CBCL Internalizing and Externalizing Problems, CBCL Adjusted and Unadjusted Externalizing Problems (i.e., total after with 3 irritability-related items excluded), and CGAS functional impairment. Sex was included as a time-invariant covariate. Full multilevel model is as follows:

Level 1:

$$p_{it} = p(Y_{it} = 1|u_{it})$$
$$Y_{it}|u_{it} \sim Bernoulli(p_{it})$$
$$logit(p_{it}) = b_{0i} + b_{1i}Time_{it} + b_{2i}Irritability_{it} + b_{3i}Time_{it}(Irritability_{it})$$

Level 2:

$$b_{0i} = \beta_{00} + \beta_{01}Sex_i + \zeta_{0i}$$
$$b_{1i} = \beta_{10} + \beta_{11}Sex_i + \zeta_{1i}$$
$$b_{2i} = \beta_{20} + \beta_{21}Sex_i + \zeta_{2i}$$
$$b_{3i} = \beta_{30} + \beta_{31}Sex_i + \zeta_{3i}$$

Combined Equation:

$$\begin{aligned} logit(p_{it}) = {} & \beta_{00} + \beta_{01}Sex_i + \beta_{10}Time_{it} + \beta_{11}Time_{it}(Sex_i) + \beta_{20}Irritability_{it} \\ & + \beta_{21}Irritability_{it}(Sex_i) + \beta_{30}Time_{it}(Irritability_{it}) \\ & + \beta_{31}(Time_{it})(Irritability_{it})(Sex_i) + [\zeta_{0i} + \zeta_{1i} + \zeta_{2i} + \zeta_{3i} + \epsilon_{it}] \end{aligned}$$

We then employed follow up receiver Operating Characteristic (ROC) curves to explicitly model the accuracy of predictions for the CBCL Internalizing and Externalizing problems as well as CGAS parent and clinician-rated impairment. ROC curves graphically represent true positive versus false positive rates. For significant models from the longitudinal logistic regressions, we compared different assessment methods for childhood irritability (parent only, teacher only, combination of parent "and" teacher, or parent "or" teacher) separately with respect to their true (sensitivity) and false positive (1 – specificity) rates based on the areas between two curves

(Youngstrom, 2014). That is, when the overall model was unrelated with an outcome, the ROC curve was not graphed or compared. For ROC analyses, significance was estimated if the area under the curve (AUC) did not include 0.50 in the confidence interval; however, to enhance clinical utility, use of higher cutoffs is recommended. In the present study, we employed an AUC threshold greater than 0.56, which suggests a significant, small effect (Rice & Harris, 2005). An AUC of 0.70 (Swets, 1988) was prioritized if a specific informant combination was adequately associated with clinical (i.e., psychopathology, functional impairment) outcomes. Hanley & McNeil's (1983) method tested AUC differences to estimate which informant combination best predicted key outcomes, which represents a more accurate diagnostic test.

Results

Longitudinal Logistic Regressions

To review, we ran longitudinal logistic regressions for each of the informant combinations (i.e., parent only, teacher only, parent "and" teacher, and parent "or" teacher) separately with respect to outcomes (CBCL Internalizing and Externalizing Problems, CBCL Adjusted and Unadjusted Externalizing Problems (i.e., excluding irritability-related items), and CGAS functional impairment). Fixed effects estimated the associations of informant combinations with clinically elevated outcomes. For example, when a specific informant or informant combination significantly and positively predicted irritability, inferentially, this suggests that irritability was positively associated with increased log odds of the outcome, holding sex and time constant. In other words, as irritability increased, the likelihood of the outcome occurring also increased, after accounting for the effects of sex and time. Summary of regression models are on Tables 1.4 and 1.5.

Predicting Clinically Elevated Internalizing Problems

We used four methods (parent only, teacher only, parent "and" teacher, and parent "or" teacher) of estimating irritability to separately evaluate their predictions of clinically elevated CBCL Internalizing Problems. Across the separate models, parent only ($\text{Wald Chi}^2 = 29.19, p = 0.0001$)

and combined parent “or” teacher ($\text{Wald Chi}^2 = 25.30, p = 0.0007$) models significantly predicted elevated CBCL Internalizing Problems, but the teacher only and parent “and” teacher combination did not.

Parent only: Parent-rated irritability significantly ($\beta = 1.19, z = 2.07, p = 0.038$) predicted a higher probability of exhibiting clinically significant CBCL Internalizing Problems but additional effects of time, sex, and the associated cross level interactions (i.e., [time x sex], [irritability x sex], [irritability x time], [time x sex x irritability]) did not.

Parent “or” Teacher: The parent “or” teacher-rated irritability combination marginally predicted the probability of clinically elevated CBCL Internalizing Problems ($\beta = 0.91, z = 1.83, p = 0.067$). Time, sex, and the associated cross level interactions did not, however.

Predicting Clinically Elevated Externalizing Problems

We then tested the four informant combinations of irritability separately as predictors of clinically elevated CBCL Externalizing Problems. Parent only ($\text{Wald Chi}^2 = 34.87, p < 0.0001$), teacher only ($\text{Wald Chi}^2 = 14.93, p = 0.037$), parent “and” teacher ($\text{Wald Chi}^2 = 25.32, p < 0.0001$), and parent “or” teacher ($\text{Wald Chi}^2 = 39.07, p < 0.0001$) each significantly predicted CBCL Externalizing Problems.

Parent Only: Parent-rated irritability ($\beta = 1.94, z = 3.10, p = 0.002$) significantly predicted clinically elevated CBCL Externalizing Problems, as did time ($\beta = -1.76, z = -2.13, p = 0.033$). Sex and the associated cross level interactions did not.

Teacher Only: Fixed effects of teacher-rated irritability, time, sex, and their associated cross level interactions did not predict clinical elevations of CBCL Externalizing Problems.

Parent “and” Teacher: Similarly, parent “and” teacher-rated irritability, time, and sex did not predict clinically elevated CBCL Externalizing Problems. However, the cross-level sex x time interaction marginally ($\beta = -0.85, z = -1.72, p = 0.086$) predicted the outcome. This suggests

that the effect of time on the log odds of clinically elevated CBCL Externalizing Problems was weaker for boys than for girls.

Parent "or" Teacher: Combined parent "or" teacher-rated irritability ($\beta = 1.73, z = 3.14, p = 0.002$) significantly predicted CBCL Externalizing Problems. Effects of time, sex, and the associated cross level interactions did not, however.

Predicting Clinically Elevated Externalizing Problems: Adjusted vs Unadjusted Models

To combat potential inflated associations between irritability and outcomes secondary to potential item overlap, we removed three items related to irritability from the CBCL Externalizing Problems raw score to create an outcome score free of overlapping irritability items. We used identical procedures to those described above to separately test predictions of elevated externalizing problems from the four different informant combinations, but now adjusted for item overlap. The overall models for the parent only ($\text{Wald Chi}^2 = 45.42, p < 0.0001$), combined parent "and" teacher ($\text{Wald Chi}^2 = 17.78, p = 0.0068$), and combined parent "or" teacher ($\text{Wald Chi}^2 = 49.39, p < 0.0001$) significantly predicted Adjusted CBCL Externalizing Problems. Similarly, parent only ($\text{Wald Chi}^2 = 39.20, p < 0.0001$), combined parent "and" teacher ($\text{Wald Chi}^2 = 16.20, p = 0.013$), and combined parent "or" teacher ($\text{Wald Chi}^2 = 43.38, p < 0.0001$) significantly predicted Unadjusted CBCL Externalizing Problems.

Parent Only: Parent-rated irritability ($\beta = 1.98, z = 2.85, p = 0.004$) significantly predicted elevated Adjusted CBCL Externalizing Problems, however, time, sex, and their cross level interactions did not. Similarly, parent-rated irritability ($\beta = 2.11, z = 2.84, p = 0.005$) significantly predicted elevated Unadjusted CBCL Externalizing Problems as well.

Parent "and" Teacher: Sex significantly predicted elevated log odds in both Adjusted ($\beta = 2.12, z = 2.12, p = 0.034$) and Unadjusted ($\beta = 1.95, z = 2.28, p = 0.023$) CBCL Externalizing Problems.

Parent "or" Teacher: Parent "or" teacher-rated irritability ($\beta = 1.84, z = 2.97, p = 0.003$) significantly predicted clinical elevations of Adjusted CBCL Externalizing Problems but effects secondary to time, sex, and the associated cross level interactions did not. Parent "or" teacher-rated irritability ($\beta = 1.77, z = 2.71, p = 0.007$) also significantly predicted elevated Unadjusted CBCL Externalizing Problems.

Predicting Clinical Elevated Parent-Rated Functional Impairment

Similar tests evaluated whether the four methods of ascertaining irritability significantly predicted clinically elevated parent-rated functional/global impairment. The overall models for parent only ($\text{Wald Chi}^2 = 37.83, p < 0.0001$), parent "and" teacher ($\text{Wald Chi}^2 = 18.30, p = 0.0026$), and parent "or" teacher ($\text{Wald Chi}^2 = 32.56, p < 0.0001$), but not teacher only, significantly predicted parent-rated functional impairment.

Parent Only: Parent-rated irritability and time each significantly predicted parent-rated functional impairment ($\beta = 1.04, z = 2.44, p = 0.015$ and $\beta = -0.84, z = -2.43, p = 0.015$), respectively. Interpretively, this suggests that – with all other predictors held constant – a decrease in log odds over time was associated with more parent-rated functional impairment. Effects of sex and the associated cross level interactions did not, however.

Parent "and" Teacher: Parent "and" teacher-rated irritability significantly predicted parent-rated functional impairment ($\beta = 1.03, z = 2.38, p = 0.017$) whereas time, sex, and the associated cross level interactions did not.

Parent "or" Teacher: Parent "or" teacher-rated irritability significantly ($\beta = 0.81, z = 2.21, p = 0.027$) predicted parent-rated functional impairment. Other effects of time, sex, and the associated cross level interactions did not.

Predicting Clinically Elevated Clinician-Rated Functional Impairment

Finally, we used identical procedures to test predictions of clinician-related global impairment. Parent only ($\text{Wald Chi}^2 = 41.43, p < 0.0001$) and parent "or" teacher combination ($\text{Wald Chi}^2 =$

$42.33, p < 0.0001$) models significantly predicted clinician-rated functional impairment, but teacher only and combined parent "and" teacher did not.

Parent Only: Parent-rated irritability ($\beta = 1.74, z = 2.75, p = 0.006$) significantly predicted clinician-rated functional impairment, but time, sex, and the associated cross level interactions did not.

Parent "or" Teacher: Parent "or" teacher-rated irritability ($\beta = 1.65, z = 2.95, p = 0.003$) also significantly predicted clinician-rated functional impairment. However, secondary effects of time, sex, and the associated cross level interactions did not.

Receiver Operator Characteristics: Model Comparison

To review, when longitudinal logistic regression models significantly predicted psychopathology or functional impairment (i.e., CBCL Internalizing Problems, CBCL Externalizing Problems, Parent- and Clinician-rated Functional Impairment from the CGAS), we subsequently estimated ROC curves for each individual outcome. Statistical significance was indicated if the AUC confidence interval did not include 0.70. AUCs for models compared are summarized in Table 1.6.

Clinically Elevated CBCL Internalizing Problems

The ROC-derived comparison suggested that parent only and parent "or" teacher models were significantly associated with CBCL Internalizing Problems. With respect to AUCs values, parent-rated irritability was significantly more accurate than parent "or" teacher report ($chi^2 = 5.08, p = 0.0242$).

Clinically Elevated CBCL Externalizing Problems

Given that all informant combinations significantly predicted clinically elevated CBCL Externalizing Problems at the first step, we compared the accuracy of their ROC curves. Overall, based on their respective AUC values, the four informant combinations did not differ significantly with respect to their accuracy in predicting CBCL Externalizing Problems.

Clinically Elevated Parent- and Clinician-rated Functional Impairment

Parent-rated: Three methods of irritability – parent only, parent "and" teacher, and parent "or" teacher – were compared with respect to their prediction of parent-rated functional impairment. Similar to the pattern observed above with respect to CBCL Externalizing Problems, the three informant combinations did not differ significantly in accurately predicted parent-rated functional impairment.

Clinician-rated: We compared the parent only and combined parent "or" teacher irritability ratings; once again, the informant combinations did not differ significantly in detecting clinician ratings of clinically elevated functional impairment.

Discussion

Although childhood irritability is a major motivation for treatment-seeking families, it is surprising that there is modest evidence to guide optimal methods of assessment. The current study empirically prosecuted specific combinations of parent and teacher ratings of childhood irritability – the most common and clinically relevant informants -- with respect to their prediction of later clinical outcomes. We prospectively followed a diverse sample of 230 children (68% male, 50% non-White) with (*n*=120) and without (*n*=110) ADHD for 6 to 7 years from childhood into early adolescence. Multi-level longitudinal logistic regressions compared parent only, teacher only, parent "and" teacher, and parent "or" teacher ratings of childhood irritability as independent predictors of CBCL Internalizing and Externalizing Problems as well as parent- and clinician-rated functional impairment. Because shared item overlap likely inflates associations between irritability and psychopathology criteria, we tested parallel models with irritability items removed from the outcome. Next, when specific informants or informant combinations significantly predicted outcomes, we subsequently employed ROC to compare the relative accuracy of their predictions. Several key findings emerged: (a) overall, parent-rated and parent "or" teacher-rated irritability significantly predicted CBCL Internalizing and Externalizing Problems as well as functional impairment (b) although the overall model of teacher report significantly predicted CBCL Externalizing Problems, the fixed effect of teacher-rated irritability

did not; (c) when overlapping items between irritability and psychopathology outcomes were removed, parent only and parent “or” teacher-rated irritability significantly predicted Adjusted and Unadjusted Externalizing Problems; (d) follow-up tests revealed that parent only rated irritability was significantly more accurate than parent “or” teacher-rated irritability with respect to longitudinally predicted CBCL Internalizing Problems. None of the other informant combinations differed with respect to any other clinical or functional outcome.

In the present study, parent only models positively predicted internalizing and externalizing problems (e.g., Leibenluft et al., 2006, Stringaris et al., 2009). On the other hand, teacher only models specifically predicted externalizing outcomes, but not internalizing outcomes or functional impairment. This dissociation is somewhat surprising given that teachers are central to assessments of children’s socio-emotional functioning and their unique position to observe children intensively across multiple situations (e.g., social, academic) (Keiley et al., 2003). Previously, teacher assessment data were inconsistently associated with key outcomes; for example, some evidence suggested clinical utility for predicting internalizing and externalizing outcomes whereas other studies found that teacher data specifically predicted later externalizing outcomes (Dwyer et al., 2006; Tandon et al., 2009; Zee & Rudasill, 2011). Several complicating factors likely affected definitive inferences: first, previous studies were typically cross-sectional, limiting inferences to concurrent (versus predictive) validity. Second, teacher-rated data are also sensitive to factors ranging from normative changes in teachers across grades, variable relationship quality between students and teachers (Rucinski et al., 2018), and experience of teachers (i.e., years teaching and familiarity with youth mental health problems; Monducci et al., 2018), which weakens clinical inferences. Interestingly, teacher-rated data increased clinical utility when paired together with parent report using the “or” rule in the present study, perhaps capturing different patterns of child functioning across multiple settings, thereby increasing sensitivity (De Los Reyes & Kazdin, 2005, Rettew et al., 2009). Tests with high sensitivity align well with efforts to improve early detection and referral to interventions that

significantly reduce later burden (Chorpita & Daeiden, 2014; Patel et al., 2007). However, increasing sensitivity is accompanied by the risk of false positives (thus decreasing specificity). Thus, optimal informants and informant combinations must balance sensitivity and specificity and simultaneously attend to the goals of the assessment. For example, Youngstrom & Van Meter (2016)'s clinical management model contends that tests intended to identify high risk youth or high acuity conditions may require psychometric profiles of sensitivity and specificity that are meaningfully different than tests designed to identify youth who would benefit from ongoing monitoring or preventative strategies.

More than psychopathology, functional impairment is the principle motivation for treatment seeking (Barkley, 2015). To strengthen the clinical significance of this study, informants for irritability were evaluated with respect to their independent prediction of adolescent functional impairment. Parent-rated and parent "or" teacher-rated irritability positively predicted parent- and clinician-rated functional impairment whereas teacher ratings alone did not. Consistent with previous research, parents observe children more diversely and thus may be better positioned to assess functional/global impairment relative to teachers (De Los Reyes et al., 2009). It may also be the case that determining "optimal informants" is elusive; rather, focusing on specific contexts and specific clinical questions should guide decisions about clinical assessment strategies. In other words, ascertaining optimal informants must attend to specific situational contexts, an area that warrants further research (De Le Reyes et al., 2015).

Given its clinical and transdiagnostic utility (Beauchaine & Tackett, 2020), evaluating which informants – independently and in combination – optimally assess childhood irritability is urgently needed. Like other individual difference behaviors that interact with contextual factors (e.g., emotion regulation – Eisenberg et al., 2010), irritability is observable across settings and situations. Theoretically, context-dependent irritability is hypothesized to explain how irritability manifests across settings, including in response to situational influences (e.g., family stress versus peer stress at school), but that has yet to be directly prosecuted for childhood irritability

(Stringaris et al., 2017). For instance, whereas ODD symptoms observed at home were associated with child effortful control and parental hostility, teacher-rated ODD symptoms – but not parent-rated ODD symptoms – were positively associated with youth social problems (Drabick et al., 2007; Lavigne et al., 2014). Similarly, parent and teacher ratings of peer relationships showed similar divergence with socio-emotional adjustment (Cullerton-Sen & Crick, 2005). Therefore, in addition to elucidating limits of different informants (and their combinations), a parallel line of work should evaluate parent and teacher ratings of behaviors specific to different settings the potential that they may show unique patterns of associations with external criteria.

Collecting multi-informant clinical assessment data is a central tenet of developmental psychopathology, but widespread clinical implementation is disadvantaged by the absence of strong evidence on how to combine the data. Thus, psychometric evaluations of assessment methods and informants must report their performance with respect to balancing true-positives against false positives. Detrimentally, false-positive assessment data perpetuate systematic inequities, contribute to stigma, and create additional burden for acute mental health resources (Isaac et al., 2024; Millman & Schiffman, 2017). Overall, parent-rated irritability and parent “or” teacher informants consistently predicted outcomes with high AUC values (i.e., >0.90), suggesting strong accuracy and clinical utility. Compared to each other, parent only ratings of irritability were significantly more accurate than the “or” rule in detecting clinically elevated internalizing problems; these two approaches were comparable with respect to predictions of externalizing problems and functional impairment, however. This aligns with previous evidence where parents reliably predicted internalizing problems whereas both parent and teacher report usefully predicted externalizing problems (Sourander & Helstela, 2005). It may be the case that internalizing and externalizing problems are separable with respect to parent vs. teacher ratings. As noted above, the clinical purpose of clinical assessment must also be considered: for example, the school-based universal mental health screening (UMHS) is an evidence-based approach to identify children at risk of developing mental health concerns by utilizing teacher, parent, and self-report

(Weist et al., 2003). Although UMHS reduces biases by incorporating multiple informants to detect mental health challenges, not all contexts and settings (e.g., primary care, emergency rooms, schools) have the resources or training to assess multiple informants and provide comprehensive services. (Baker-Ericzen et al., 2013). Primary care providers assess mental health challenges, but limited face-to-face time with families challenges reliable use of evidence-based assessment methods (Baxter et al., 2013). The current findings suggest screening tools may prioritize specific informants for internalizing and externalizing outcomes, which may reduce opportunity costs and improve clinical uptake. This promotes triaging and informs evidence-based clinical decision making (e.g., need for services; Higa-McMillan et al., 2016).

Important strengths in the current study include a prospective longitudinal design from childhood to adolescence and multi-informant/method assessments of key constructs. Nevertheless, we readily acknowledge key limitations. For example, youth self-reported irritability – in childhood and adolescence – data were not collected, despite evidence of incremental validity across multiple dimensions of psychopathology (Aebi et al., 2017; De Los Reyes & Kazdin, 2005). Second, there is increasing evidence that irritability consists of facets that differentially predict outcomes (e.g., tonic versus phasic irritability; Silver et al., 2021; high anger, low anger, distress three factor model; Hirsch et al., 2022). To accelerate understanding of irritability as a transdiagnostic risk factor, future research should evaluate a broader range of assessment methods including task-based strategies and variations in situational and contextual aspects of irritability (Dougherty et al., 2020). Finally, although we tested the accuracy of different informant combinations across parent and teacher report, research also suggests the utility of discrepant reports in identifying children at risk for later challenges. For example, the *magnitude* of parent-child discrepancies with respect to three domains of parental monitoring (i.e., child disclosure, parental knowledge, and parental solicitation) were uniquely associated with childhood depressive symptoms, suggesting that patterns of discrepancies may aid in detecting later depression (De Los Reyes & Goodman et al., 2008). Additionally, there is

increasing recognition that informant discrepancies do not simply reflect measurement error but rather may be sensitive to unique and important clinical phenomena. There is considerable divergence between youth and parenting ratings on the Affective Reactivity Scale, suggesting potentially different interpretations of scale items (Mallidi et al., 2023). However, these discrepancies may be further affected by age and diagnosis (i.e., great discrepancy with young children or having a diagnosis of DMDD) (Zik et al., 2022). Thus, future studies should not only investigate how informant data can be *combined* but also to understand if informant *differences* are clinically meaningful (e.g., differentially predict clinical outcomes).

Study I Tables and Figures

Table 1.1: Descriptive Information for Demographics and Study Variables

Variable	Wave 1	Wave 2	Wave 3
Age (*SD*)	7.4 (1.1)	9.7 (1.3)	12.1 (1.4)
Sex (% Male)	68%	-	-
Race (% Caucasian)	51%	-	-
N ADHD diagnosis (% sample)	121 (52%)	88 (38%)	52 (22%)
Mean number Parent only irritability symptoms (SD)	0.53 (.94)	0.34 (.71)	0.44 (.84)
Mean number Teacher only irritability symptoms (SD)	0.41 (.90)	0.24 (.64)	0.24 (.65)
Mean number Parent AND Teacher irritability symptoms (SD)	0.16 (.56)	0.036 (.26)	0.053 (.31)
Mean number Parent OR Teacher irritability symptoms (SD)	0.66 (1.04)	0.42 (.78)	0.47 (.86)

Table 1.2: Summary of Study Measures

Construct	**Measure**	**Informant/Method**	**Measure Information**
Childhood Irritability	Disruptive Behavior Disorder Rating Scale	Parent and teacher report rating scale	Symptom categorized as 'present' if indicated a 2 or 3
Internalizing Problems	Child Behavior Checklist	Parent report rating scale	T scores converted to clinical elevations/normal limits
Externalizing Problems	Child Behavior Checklist – T Scores	Parent report rating scale	T scores converted to clinical elevations/normal limits
Adjusted & Unadjusted Externalizing Problems	Child Behavior Checklist – Externalizing Problems Raw Scores	Parent report rating scale	Raw scores dichotomized and raw scores that removed irritability-related items
Functional Impairment	Children's Global Assessment Scale	Parent-rated and clinician-rated interview	Clinical cutoff scores used to determine clinical elevations/normal limits

Table 1.3: Number of Participants with Clinical Elevations of Key Outcomes

Parent Report	Wave 1	Wave 2	Wave 3
Internalizing Problems	n = 52 (22%)	n = 42 (22%)	n = 30 (13%)
Externalizing Problems	n = 60 (26%)	n = 25 (11%)	n = 18 (8%)

Teacher Report	Wave 1	Wave 2	Wave 3
Internalizing Problems	n = 37 (16%)	n = 14 (6%)	n = 14 (6%)
Externalizing Problems	n = 35 (15%)	n = 13 (6%)	n = 12 (5%)
Parent-rated	Wave 1	Wave 2	Wave 3
Functional Impairment	n = 98 (42%)	n = 58 (25%)	n = 36 (16%)

Clinician Rated	Wave 1	Wave 2	Wave 3
Functional Impairment	n = 121 (52%)	n = 96 (41%)	n = 74 (32%)
Adjusted Models			
	Wave 1	Wave 2	Wave 3
Adjusted Externalizing Problems	n = 57 (25%)	n = 44 (22%)	n = 36 (22%)
Unadjusted Externalizing Problems	n = 50 (22%)	n = 49 (25%)	n = 36 (22%)

Table 1.4: Multilevel Longitudinal Logistic Regression

1.4a: *Informant Combinations of Irritability Predicting Internalizing Problems*

Model	Predictors	*β*	*z*	Wald	*p*
Parent Report Only					
	Constant	-4.02	-4.89		<.0001*
	Time	0.36	0.81		0.42
	Irritability	1.19	1.28		0.038*
	Sex	1.02	0.82		0.20
	Time x Irritability	0.36	-1.65		0.41
	Time x Sex	-0.88	0.13		0.099
	Irritability x Sex	0.084	0.82		0.90
	Time x Irritability x Sex	-0.51	-1.00		0..32
	Overall Model			29.19	<.0001*
Teacher Only					
	Constant	-5.95	-7.32		<.0001*
	Time	-0.26	-0.40		0.69
	Irritability	0.63	0.61		0.54
	Sex	0.94	0.95		0.34
	Time x Irritability	0.27	0.22		0.83
	Time x Sex	-0.69	-0.86		0.39
	Irritability x Sex	0.22	0.19		0.85
	Time x Irritability x Sex	-1.17	-0.78		0.44
	Overall Model			11.44	0.121
Parent “and” Teacher					
	Constant	-3.16	-5.07		<.0001*
	Time	0.44	1.35		0.18
	Irritability	0.63	0.62		0.53
	Sex	0.91	1.46		0.15
	Time x Irritability	0.32	0.16		0.88
	Time x Sex	-0.96	-2.39		0.017
	Irritability x Sex	0.34	0.30		0.76
	Time x Irritability x Sex	-0.59	-0.28		0.78
	Overall Model			13.11	0.0695
Parent “or” Teacher					
	Constant	-3.66	-4.90		<.0001*
	Time	0.15	0.36		0.72
	Irritability	0.91	1.83		0.067
	Sex	0.91	1.21		0.23
	Time x Irritability	0.40	1.02		0.31
	Time x Sex	-0.53	-1.06		0.29
	Irritability x Sex	-0.11	-0.19		0.85
	Time x Irritability x Sex	-0.56	-1.23		0.22
	Overall Model			25.30	0.0007*

1.4b: *Informant Combinations of Irritability Predicting Externalizing Problems*

Model	Predictors	*β*	*z*	Wald	*p*
Parent Report Only					
	Constant	-3.67	-4.40		<.0001*
	Time	-1.76	-2.13		0.033*
	Irritability	1.94	3.10		0.002*
	Sex	0.12	0.15		0.88
	Time x Irritability	0.56	1.09		0.28
	Time x Sex	0.41	0.47		0.64
	Irritability x Sex	0.64	0.86		0.39
	Time x Irritability x Sex	-0.78	-1.23		0.22
	Overall Model			34.87	<.0001*
Teacher Only					
	Constant	-2.55	-3.23		0.001*
	Time	-0.52	-0.96		0.34
	Irritability	1.25	1.58		0.11
	Sex	0.11	0.16		0.87
	Time x Irritability	-0.55	-0.57		0.57
	Time x Sex	-0.79	-1.13		0.26
	Irritability x Sex	-0.17	-0.20		0.84
	Time x Irritability x Sex	0.90	0.86		0.39
	Overall Model			14.93	0.037*
Parent “and” Teacher					
	Constant	-3.09	-4.48		<.0001*
	Time	-0.54	-1.38		0.17
	Irritability	0.68	0.59		0.89
	Sex	0.62	0.96		0.34
	Time x Irritability	-0.24	-0.14		0.89
	Time x Sex	-0.85	-1.72		0.086
	Irritability x Sex	2.12	1.37		0.17
	Time x Irritability x Sex	0.11	0.06		0.95
	Overall Model			25.32	0.007*
Parent “or” Teacher					
	Constant	-3.69	-4.56		<.0001
	Time	-1.49	-1.93		0.053
	Irritability	1.73	3.14		0.002*
	Sex	0.42	0.53		0.60
	Time x Irritability	0.47	0.99		0.32
	Time x Sex	0.21	0.25		0.80
	Irritability x Sex	0.033	0.05		0.96
	Time x Irritability x Sex	-0.36	-0.65		0.51
	Overall Model			39.07	<.0001*

1.4c: *Informant Combinations of Irritability Predicting Parent Rated Functional Impairment*

Model	Predictors	*β*	*z*	Wald	*p*
Parent Report Only					
	Constant	-1.21	-2.97		0.003*
	Time	-0.84	-2.43		0.015*
	Irritability	1.04	2.44		0.015*
	Sex	0.38	0.78		0.43
	Time x Irritability	0.25	0.79		0.43
	Time x Sex	0.56	1.43		0.15
	Irritability x Sex	-0.11	-0.22		0.83
	Time x Irritability x Sex	-0.49	-1.30		0.19
	Overall Model			37.83	<.0001*
Teacher Only					
	Constant	-0.90	-2.39		0.017
	Time	-0.052	-0.16		0.87
	Irritability	1.47	0.82		0.073
	Sex	0.39	0.84		0.40
	Time x Irritability	-0.055	-0.21		0.83
	Time x Sex	-0.14	-0.36		0.72
	Irritability x Sex	-1.04	-1.22		0.22
	Overall Model			9.06	0.17
Parent “and” Teacher					
	Constant	-0.93	-2.67		0.008*
	Time	-0.36	-1.44		0.15
	Irritability	1.02	2.38		0.017*
	Sex	0.44	1.06		0.29
	Time x Irritability	0.20	0.50		0.62
	Time x Sex	0.00096	<0.01		0.997
	Overall Model			18.30	0.0026*
Parent “or” Teacher					
	Constant	-1.15	-2.98		0.003
	Time	-0.60	-1.91		0.057
	Irritability	0.81	2.21		0.027*
	Sex	0.30	0.65		0.51
	Time x Irritability	0.12	0.45		0.66
	Time x Sex	0.43	1.17		0.24
	Irritability x Sex	-0.097	-0.23		0.82
	Time x Irritability x Sex	-0.35	-1.05		0.29
	Overall Model			32.56	<.0001*

1.4d: *Informant Combinations of Irritability Predicting Clinician Rated Functional Impairment*

Model	Predictors	*β*	*z*	Wald	*p*
Parent Report Only					
	Constant	-1.03	-2.64		0.008
	Time	-0.10	-0.36		0.72
	Irritability	1.74	2.75		0.006
	Sex	0.70	1.51		0.132
	Time x Irritability	-0.010	-0.02		0.98
	Time x Sex	0.15	0.45		0.66
	Irritability x Sex	-0.47	-0.67		0.51
	Time x Irritability x Sex	-0.26	-0.52		0.60
	Overall Model			41.43	<.0001
Teacher Only – Model did not converge					
Parent “and” Teacher – Model did not converge					
Parent “or” Teacher					
	Constant	-1.10	-2.83		0.005
	Time	0.0222	0.08		0.94
	Irritability	1.65	2.95		0.003
	Sex	0.76	1.61		0.107
	Time x Irritability	-0.21	-0.57		0.57
	Time x Sex	0.038	0.11		0.91
	Irritability x Sex	-0.64	-1.04		0.30
	Time x Irritability x Sex	0.057	0.13		0.84
	Overall Model			42.33	<.0001

Table 1.5: Multilevel Longitudinal Logistic Regression – Adjusted Models for Item Overlap

	Adjusted				*Unadjusted*			
Predictors	*β*	*z*	Wald	*p*	*β*	*z*	Wald	*p*
Parent Report Only								
Constant	-4.40	-4.96		<.0001*	-5.04	-4.78		<.0001*
Time	-0.24	-0.41		0.68	0.27	0.45		0.66
Irritability	1.98	2.85		0.004	2.11	2.84		0.005
Sex	0.93	1.06		0.29	1.31	1.33		0.18
Time x Irritability	0.090	0.19		0.85	0.17	0.32		0.75
Time x Sex	0.26	0.40		0.69	-0.33	-0.48		0.63
Irritability x Sex	0.39	0.50		0.62	0.19	0.23		0.82
Time x Irritability x Sex	-0.083	-0.14		0.89	0.30	0.46		0.64
Overall Model			45.42	<.0001			39.20	<.0001
Teacher Only								
Constant	-4.68	-2.98		0.003	-5.46	-2.75		0.006*
Time	0.64	0.91		0.37	0.88	1.13		0.26
Irritability	2.34	1.86		0.064	2.38	1.62		0.11
Sex	1.45	1.42		0.16	1.47	1.24		0.22
Time x Irritability	-1.2	-0.93		0.35	-0.86	-0.66		0.51
Time x Sex	-0.70	-0.89		0.38	-0.69	-0.79		0.43
Irritability x Sex	-1.43	-1.13		0.26	-1.05	-0.73		0.47
Time x Irritability x Sex	1.85	1.32		0.19	1.62	1.07		0.28
Overall Model			10.10	0.183			9.92	0.193
Parent “and” Teacher								
Constant	-4.01	-5.08		<.0001*	-5.01	-4.93		<.0001*
Time	0.30	0.76		0.45	0.85	1.93		0.054
Irritability	1.07	0.86		0.39	1.41	1.01		0.31
Sex	1.55	2.12		0.034	1.95	2.28		0.023
Time x Irritability	2.75	1.94		0.052	2.55	1.63		0.10
Time x Sex	-0.58	-1.27		0.21	-1.00	-1.99		0.047*
Time x Irritability	0.58	0.42		0.68	0.59	0.37		0.71
Overall Model			17.78	0.0068			16.20	0.0127
Parent “or” Teacher								
Constant	-4.50	-5.03		<.0001*	-5.05	-4.90		<.0001*
Time	-0.044	-0.07		0.94	0.35	0.58		0.56
Irritability	1.84	2.97		0.003*	1.77	2.71		0.007*
Sex	1.39	1.57		0.12	1.48	1.52		0.13
Time x Irritability	0.0009	<.00 1		0.99	0.16	0.32		0.75
Time x Sex	-0.12	-0.18		0.86	-0.58	-0.85		0.40
Irritability x Sex	-0.36	-0.53		0.59	-0.19	-0.26		0.79
Time x Irritability x Sex	0.42	0.78		0.43	0.66	1.09		0.27
Overall Model			49.39	<.0001			43.38	<.0001

Table 1.6: ROC Analyses – Model AUC Values and Model Comparisons

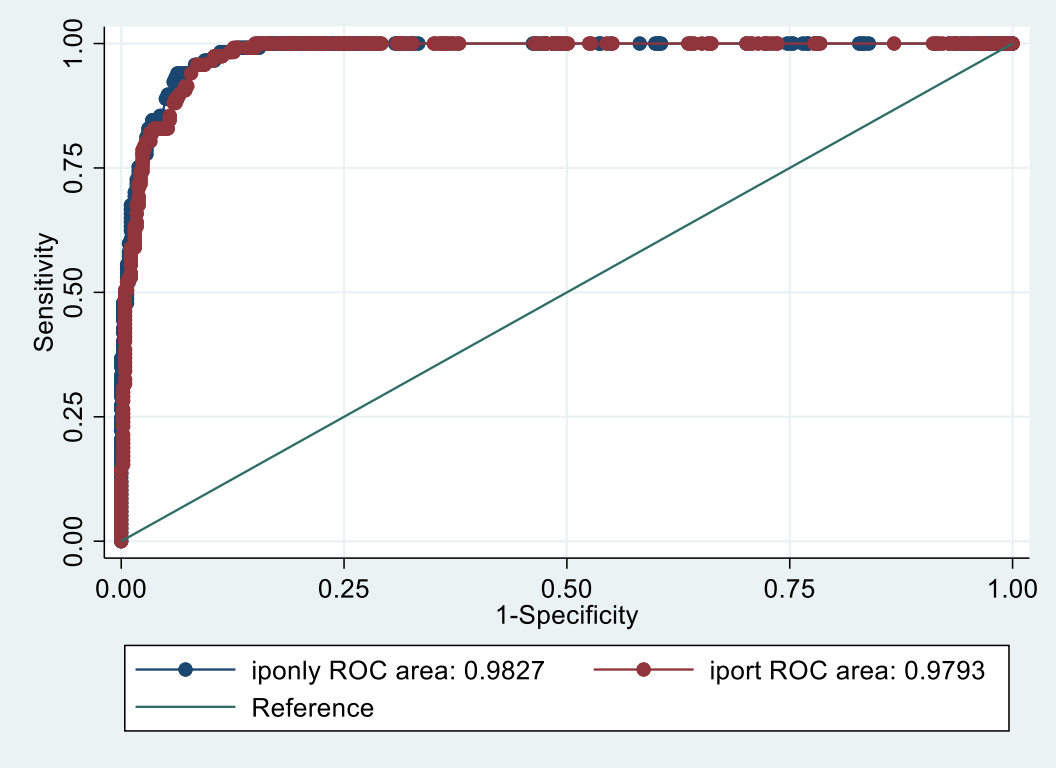
	Model AUC and Comparisons
Internalizing Problems	*Parent vs Parent "OR" Teacher*
AUC Values	Parent: 0.983 Parent "OR" Teacher: 0.979
Model Comparison	$\chi^2 = 5.08, p = 0.024$

Externalizing Problems	*Parent vs Teacher vs Parent "AND" Teacher vs Parent "OR" Teacher*
AUC Values	Parent: 0.991 Teacher: 0.984 Parent "AND" Teacher: 0.988 Parent "OR" Teacher: 0.987
Model Comparison	$\chi^2 = 4.56, p = 0.21$

Parent rated Functional Impairment	*Parent vs Parent "AND" Teacher vs Parent "OR" Teacher*
AUC Values	Parent: 0.95 Parent "AND" Teacher: 0.94 Parent "OR" Teacher: 0.94
Model Comparison	$\chi^2 = 2.82, p = 0.24$

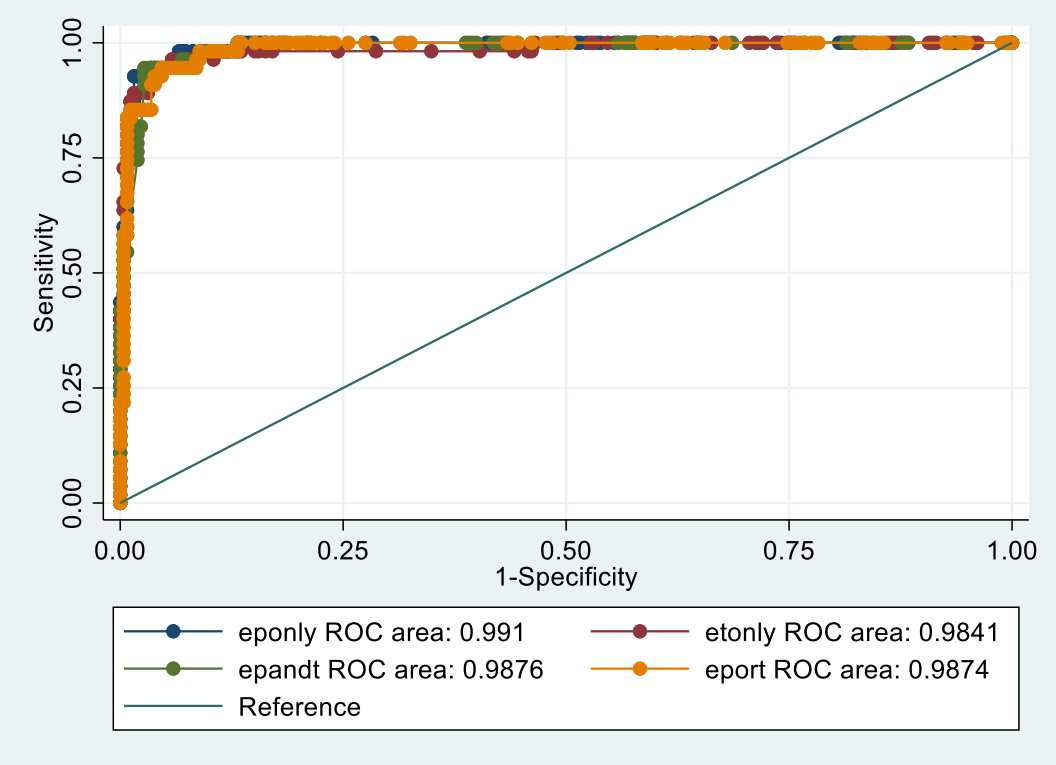
Clinician Rated Functional Impairment	*Parent vs Parent "OR" Teacher*
AUC Values	Parent: 0.938 Parent "OR" Teacher: 0.936
Model Comparison	$\chi^2 = 0.35, p = 0.56$

Figure 1.1. ROC Analyses Model Comparison – Internalizing Problems



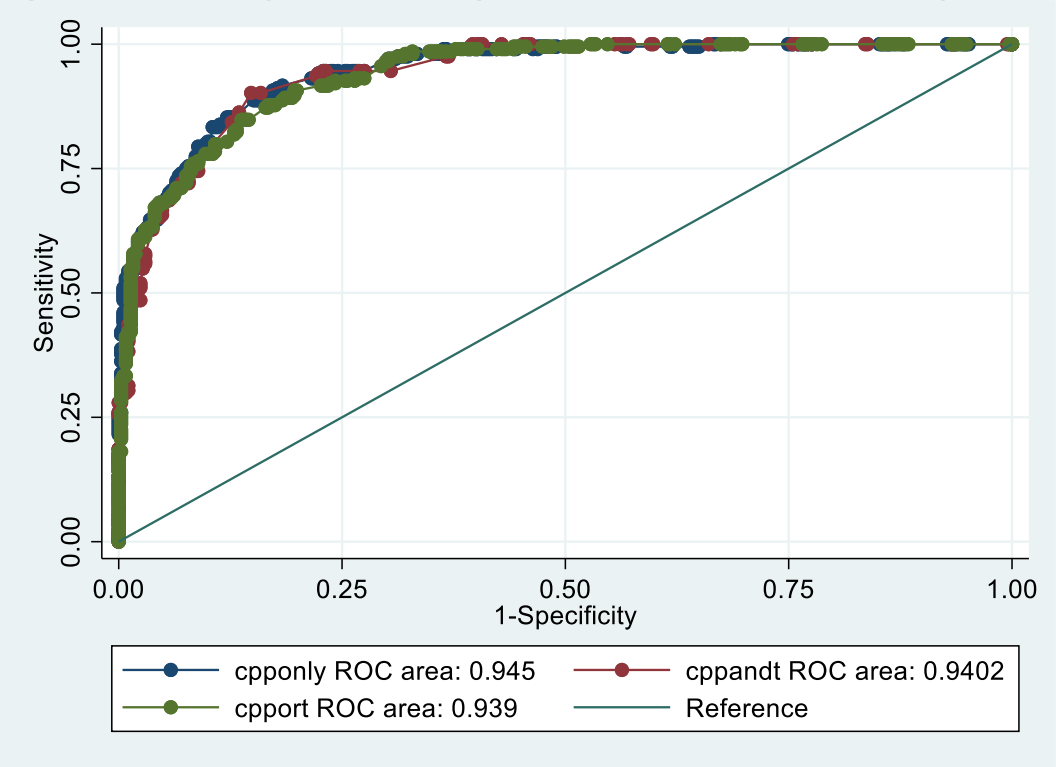
IP only = Parent only, IPorT = Parent “or” teacher

Figure 1.2. ROC Analyses Model Comparison – Externalizing Problem



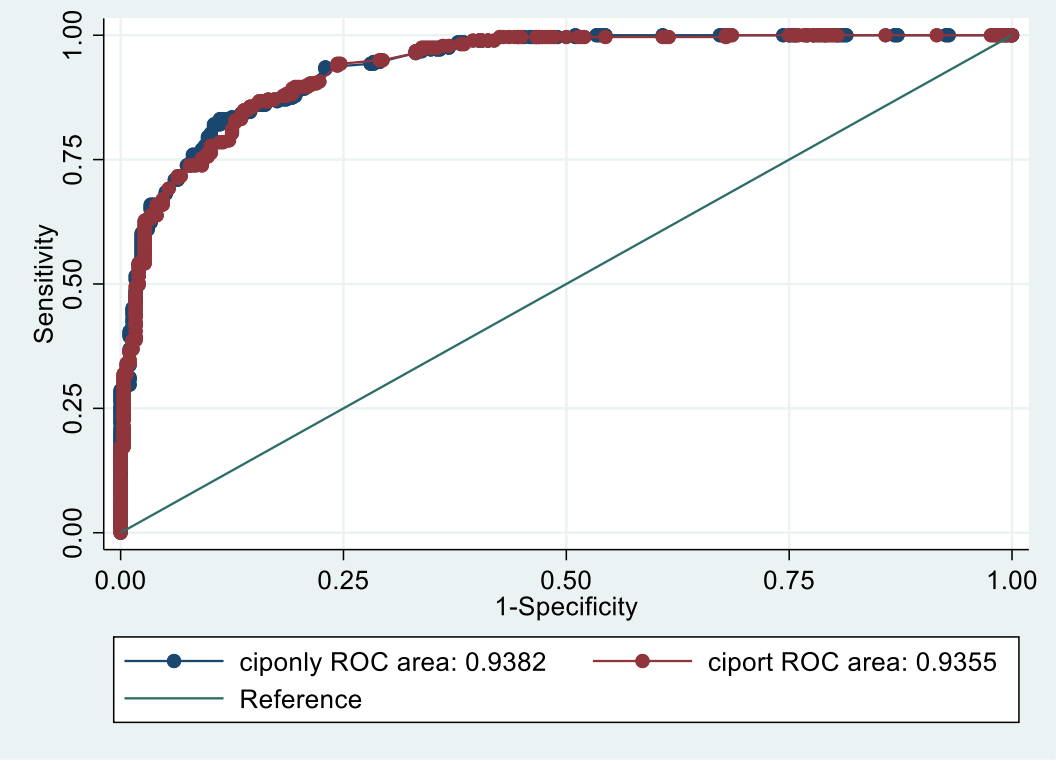
EPonly = Parent only, ETOnly = Teacher only, EPandT = Parent “and” teacher, EPorT = Parent “or” teacher

Figure 1.3. ROC Analyses Model Comparison – Parent-rated Functional Impairment



CPPonly = Parent only, CPPandT = Parent “and” teacher, CPPorT = Parent “or Teacher

Figure 1.4 ROC Analyses Model Comparison – Clinician-rated Functional Impairment



CIPonly = Parent only, CIPorT = Parent “or” teacher

Study II: Latent Profile Analysis of Family Conflict and Temperament Dimensions: Associations with Childhood Irritability

Defined by a propensity to experience anger in response to perceived frustration or threat, childhood irritability is normative (especially early in development) and typically declines during adolescence (Copeland et al., 2015). However, when frequent, developmentally-aberrant, and persistent, irritability transdiagnostically predicts externalizing (e.g., aggression) and internalizing (e.g., mood, anxiety) problems (Beauchaine & Tackett, 2020; Dougherty et al., 2015; Humphreys et al., 2019). Irritability also positively predicts worsening trajectories of functional outcomes (e.g., peer functioning, physical health, and service usage), even with control of baseline psychopathology (Sorcher et al., 2021). Its clinical significance is further suggested by recent evidence that irritability – rated independently by youth and parents – was the top problem among treatment seeking families, even more than other common dimensions of psychopathology (e.g., depression, anxiety, ADHD) (Evans et al., 2022). Taken together, there is increasing evidence that childhood irritability is prevalent, impairing, and significantly motivates treatment seeking, Thus, there is little question that irritability constitutes an important problem requiring innovations in theory, causal models, and intervention development.

The considerable heterogeneity of childhood irritability represents a key obstacle preventing accelerations in new knowledge (Brotman et al., 2017): factorially, irritability consists of separable tonic (i.e., irritable mood) and phasic (i.e., temper outbursts) components that differentially predict outcomes (Silver et al., 2021). Clinically, management of childhood irritability must differentiate whether irritability itself, its (frequent) co-occurring problems, and/or whether the entire constellation of problems should be intervention targets (Stringaris et al., 2017). Evidence-based interventions for depression, anxiety, and ODD, for example, may monitor irritability, but it rarely constitutes a unique intervention target (Stringaris et al., 2017; Vidal-Ribas et al., 2016). To date, there is inconsistent evidence on the efficacy of emerging interventions for irritability (Penton-Voak et al., 2013), including null results (Haller et al., 2022).

Alternatively, Forbes et al., (2019) proposed that targeting the general psychopathology risk factor (i.e., shared vulnerability factor across multiple dimensions of psychopathology; Caspi et al., 2014) may effectively prevent subsequent psychopathology. Based on this framework, it is plausible that targeting risk factors for irritability when designing interventions can help alleviate long term challenges as a result (Stringaris & Goodman, 2019). Thus, we contend that innovations in treatment development for irritability require careful attention to the co-occurring problems that accompany it as well as improved knowledge about unique and shared risk factors.

To improve early detection and to facilitate treatment delivery, characterizing the developmental origins of irritability should be a priority (Wang et al., 2004). For example, multiple epidemiological and longitudinal studies observed elevated mental health concerns (e.g., ADHD, depression; Biederman et al., 2006; Brotman et al., 2017; Kagan et al., 2019; Stoddard et al., 2014), among individuals with significant irritability. A recent review also identified shared risk factors for depression and irritability including genetic liability, family history of depression, and dimensions of temperament (i.e., emotionality, sociability, and activity) (Vidal-Ribas & Stringaris, 2021). However, prosecuting irritability in the context of co-occurring conditions complicates identification of unique risk factors, thus preventing innovations on targeted interventions for childhood irritability. Given that liability for irritability is significantly correlated with liability for other dimensions of psychopathology, risk factors for irritability must not be characterized in isolation. For example, in a population-based cohort of low-income families, maternal depression, harsh parenting, and parental violence exposure were independently associated with worse trajectories of irritability; however, given that low SES itself is correlated with many risk factors, it remains which factors are uniquely associated with irritability (Corcoran & Franklin, 2008; Wiggins et al., 2014). Next, clinic-referred samples yield higher comorbidity with unique demographic and clinical correlates relative to population- or school-based samples (Goodman et al., 1997) including earlier age of onset and greater clinical

severity (Tripepi et al., 2010). For instance, in a recent study of high-irritability youth, treatment seeking families often reported high rates of family dysfunction; surprisingly, however, self-reported irritability was *inversely* associated with family dysfunction (Zendarski et al., 2023), contrary to previous findings (van As & Janssens, 2002). This suggests methodological differences – spanning sampling strategies and variable methods of assessment to the continued myopic focus on individual risk factors without regard to correlated risk factors– prevent identification of novel risk factors for youth irritability. To specifically address these limitations, this study employed a nationally representative sample of youth and will rigorously prosecute unique risk factors for childhood irritability.

Reflecting fundamental developmental psychopathology principles (e.g., equifinality), multivariate models of risk factors and processes should be prioritized (Drabick & Kendall, 2010). Because of its clinical complexity and the centrality of multi-method/informant approaches to evidence-based assessment (Youngstrom & Van Meter, 2016), risk *profiles* for irritability across multiple indicators must be prioritized. Given that the assessment of environmental factors (e.g., SES, peer relationships, family environment) is necessary in case formulation, detection, and treatment of high-risk youth (Parra et al., 2010), identifying empirically-derived groups reflecting multiple domains associated with distinct long-term outcomes is a priority for monitoring and treatment triaging. For instance, in a key example of informing clinical practice, latent profiles consisting variability of irritability, self-regulation, and positive well-being characteristics were divergently associated with self-regulation, behavioral control, and prosocial behaviors. Specifically, youth characterized by high irritability/low self-regulation profile exhibited the lowest levels prosocial behaviors (Elvin et al., 2021). In a treatment seeking sample, latent profiles derived from baseline irritability and emotion regulation (i.e., high irritability/high emotion dysregulation and low irritability/low emotion dysregulation groups) were identified post treatment and individuals in the high irritability/ high emotion dysregulation group had more severe baseline problems but exhibited a steeper reduction in

symptoms over time relative to the low irritability/low emotion dysregulation group (Evans et al., 2021). Clinically, this suggests that specific profiles of risk may identify individuals who would benefit from additional surveillance. Thus, empirically-derived profiles have the potential to inform clinical practice. Put together, identifying and leveraging risk profiles, based partially on naturally-occurring variations in irritability, are hypothesized to improve early identification, intervention, and treatment planning.

Although the family environment consists broadly of facets spanning parenting behavior and parent-child synchrony to expressed emotion and marital conflict/satisfaction, family risk factors for childhood irritability have focused narrowly on parenting behaviors (e.g., permissive; Kessel et al., 2021) and dimensions of parent psychopathology (e.g., recurrent maternal depression; Wiggins et al., 2014). However, we contend that additional aspects of the family environment are strong candidates as risk factors associated with childhood irritability, including family expressiveness, cohesion, and conflict. Whereas family expressiveness is the extent to which family members are willing to express feelings directly to each other, family cohesion represents the commitment and support family members provide for each other; family conflict reflects the frequency and intensity of openly expressed anger and conflict towards each other (Moos, 1994). These family environment factors may critically contextualize childhood irritability: for example, the association of negative family expressiveness with youth aggressive behaviors was mediated by emotion regulation (Ramsden & Hubbard, 2002) and negative expressiveness was associated with increased internalizing and externalizing symptoms (Stocker et al., 2007). Next, family conflict was associated with substance use, ODD, depression, and anxiety (Biglan et al., 2015; Langley et al., 2010; Lavigne et al., 2015), all of which are themselves associated with irritability. Finally, family dysfunction (conceptualized as the lack of cohesion and structure) was associated with increased behavioral problems and poor functioning across childhood (Kroneman et al., 2009; van As & Janssens, 2002). However, these results did not consistently consider key covariates (e.g., race/ethnicity, SES). To improve precision and specificity,

dimensions of the family environment should be evaluated simultaneously to determine if naturally-occurring variations in conflict, cohesion, and expressiveness are independently associated with childhood irritability.

Although the precise architecture of temperament lacks consensus, Rothbart & Bates (1998) hypothesized that effortful control (e.g., attention, inhibitory control), surgency (e.g., fear, shyness), and negative affect (e.g., aggression, frustration) were fundamental. Individual differences in temperament are evident in infancy and they are moderately stable through adolescence (Kopala-Sibley et al., 2018). Hypothesized to affect reactivity to the environment (Ismatullina & Voronin, 2017) and coupled with evidence of its causal influences on youth and adult psychopathology (Rettew & McKee, 2005), dimensions of temperament constitute compelling risk factors for childhood irritability. Laboratory observed age 3 positive emotionality (e.g., positive affect, sociability) uniquely predicted escalating adolescent irritability whereas induced anger only marginally predicted adolescent irritability (Kessel et al., 2021). However, the ecological validity of laboratory-based observational strategies remains questionable (Goldsmith & Gagne, 2012), including only modest correlations with parent-rated temperament (Majdandžić & van den Boom, 2007); moreover, intensive administration requirements delimit widespread implementation. Parents commonly rate youth temperament given their intimate familiarity with children's emotional and behavioral response across settings and situations (Rothbart et al., 2000), which is central to contemporary perspectives on temperament. Thus, lab-based vs. parenting ratings of youth temperament may provide different/unique information regarding the structure and stability of temperament traits (Gagne et al., 2011). Next, behavioral and affective indicators of negative affect may overlap with irritability, thus potentially inflating associations of irritability with psychopathology. However, low perceived adaptability and high perceived intensity during toddlerhood were sequentially associated with the development of irritability and later depressive disorders; these associations even survived conservative adjustment for co-occurring anxiety/depressive symptoms and conduct disorder (Whelan et al.,

2015). Thus, dimensions of temperament may portend later irritability in ways that are unlikely to purely reflect methodological artifacts. This is partially supported by evidence that early temperamental dysregulation (i.e., emotionality and activity) positively predicted different patterns of ODD comorbidities such that emotionality predicted comorbidity of ODD and internalizing disorders whereas activity predicted comorbidity of ODD and ADHD, suggesting lawful divergence even with control of key demographic variables (Stringaris et al., 2010).

Although dimensions of the family environment and temperament constitute promising risk factors, their association with childhood irritability is typically prosecuted narrowly (e.g., covaried for the other). That is, despite their plausibility, putative joint effects are infrequently considered and typically evaluated using variable-centered approaches (Kessel et al., 2021). Crucially, whereas variable-centered approaches are agnostic about within-individual heterogeneity of constructs, person-centered approaches (e.g., latent profile analyses) empirically discern groups of individuals that vary significantly across multiple indicators (Howard & Hoffman, 2017; Lanza & Cooper, 2016). As a result, person-centered approaches are harmonized with developmental psychopathology by identifying groups of individuals across hypothesized risk factors, which reflect the complex and multifactorial nature of irritability. Kishida et al. (2022) discerned five latent profiles based on configurations anxiety, depression, oppositionality, and irritability; the profiles showed unique nomological associations with external criteria, including evidence that the “comorbid” profile was most impaired (e.g., compromised peer relationships). However, latent profile analyses with multiple risk factors for childhood irritability have rarely been employed. To improve predictive models and to guide evidence-based interventions for irritability, hypothesized effects simultaneously reflecting family environment (i.e., expression, cohesion, and conflict) and dimensions of temperament (i.e., effortful control, surgency, and negative affect) for childhood irritability should be employed; specifically, person-centered approaches are heuristic given the correlated and bidirectional effects central to risk factors for psychopathology (Kiff et al., 2011) and related dimensions.

Aims & Hypotheses

To advance understanding of risk factors specific to irritability, we utilized the Adolescent Brain Cognitive Development (ABCD) study, a nationally representative sample of nine- to 10-year-old children (n = 11,875) (Garavan et al., 2018). For the most recent data release, children were followed prospectively for three years into early adolescence (ages 12-13). Data used from the present study was drawn from wave three of their data collection. The ABCD study features standardized assessments of neurocognition, physical and mental health, social and emotional functioning, and culture and environment across multiple informants (i.e., parent and youth). We employed latent profile analysis to empirically identify risk profiles based concurrently on dimensions of temperament (i.e., effortful control, surgency, and negative affect) and aspects of the family environment (i.e., family expression, cohesion, and conflict) and their association with childhood irritability. We hypothesized detection of an elevated profile consisting of low effortful control, high surgency, and high negative affect with greater family conflict, low family cohesion, and low family expression will be uniquely associated with parent ratings of childhood irritability.

Methods

Procedure

All data will be derived from the Adolescent Brain Cognitive Development (ABCD) study, a longitudinal study of 11,875 youth (aged nine at baseline) recruited from 21 research sites across the United States with planned prospective follow-up for 10 years. Schools were selected based on demographic makeup. Baseline data collection was completed October 2018. Data used in the present study will come from the ABCD Release 4.01. The ABCD study consists of both parent and child report of behavioral and psychological indicators, physical well-being, cognitive functioning, and environmental factors. Parent report was gathered on each yearly visit and child report was gathered at each 6-month follow-up. Institutional Review Board

approval was obtained for each site before data collection. All parents provided written informed consent, and all children provided assent.

Measures

Childhood Irritability

The Child Behavior Checklist (CBCL; Achenbach et al., 2001) is a 113-item parent rating scales that assesses clinical problems and adaptive functioning in youth over the previous six months. Items were scored on a 0 to 2 Likert scale (0 = *not true*, 1 = *somewhat true*, and 2 = *very or often true*). Raw scores will be used to calculate a total irritability score ranging from 0 to 6 using the following items: “stubborn, sullen, or irritable”, “sudden changes in mood or feelings”, and “temper tantrums or hot temper” (Evans et al., 2019).

Dimensions of Temperament: Effortful Control, Surgency, and Negative Affect

The Early Adolescent Temperament Questionnaire (EATQ-R; Capaldi & Rothbart, 2016) is a rating scales that assesses 8 basic dimensions of temperament (activation, affiliation, attention, fear, frustration, surgency, inhibitory control, and shyness). Three higher order scales (effortful control, surgency, and negative affect) based on responses to the primary scales will be calculated for the present study (Muris et al., 2007). Effortful control was estimated from the sum of the attention, inhibitory control, and activation control subscales whereas Surgency was calculated from the sum of high intensity pleasure, fear, and shyness subscales. Negative affect was estimated from the sum of frustration, depressive mood, and aggression. Higher values represent higher effortful control, surgency, or negative affect. Dimensions of temperament were Z-score transformed to facilitate analyses and interpretation of observed effects.

Dimensions of Family Environment: Expression, Cohesion, Conflict

The Family Environment Scale (FES) (Moos, 1994; Zucker et al., 2018) is a 90-item rating scale that assesses family relationships, personal growth, and system maintenance. We utilized the family relationships domain which consists of three subscales; cohesion, expressiveness, and conflict. Each true/false item was scores as 0 or 1. Higher values represent more cohesion,

expressiveness, or conflict. Dimensions of family environment were Z-score transformed to facilitate interpretation of observed patterns.

Data Analytic Plan

To review, our goal was to discern potentially unique subgroups of the sample based on combinations of the family environment and temperament dimensions and their associations with childhood irritability. We employed latent profile analysis (LPA), which empirically derives continuous indices to discern configurations across multiple dimensions of temperament (i.e., effortful control, surgency, and negative affect) and facets of family environment (i.e., cohesion, conflict, and expression). LPA follows a “person-centered” approach rather than a “variable-centered” approach by identifying homogeneous groups within a heterogeneous population based on similar patterns of responses to observed indicator variables (Masyn, 2013). We subjected six continuous variables – family conflict, family cohesion, family expression, effortful control, surgency, and negative affect – to LPA. Four covariance structures (i.e., class-varying, unrestricted; class-invariant, unrestricted; class-varying, diagonal; class-invariant, diagonal) were evaluated during model selection (Masyn, 2013).

Latent Profile Analysis

Model Selection

LPA models with all four variance-covariance structures were tested using Mplus Version. Within each group, increasing number of classes were tested until models failed to converge or classes showed no meaningful differences. In selecting the optimal number of latent profiles, Akaike information criterion (AIC), Bayesian information criterion (BIC), and sample size-adjusted BIC will be evaluated where lower values indicate better fit (Ferguson et al., 2020). The Lo, Mendell, and Rubin (LMR) test was used to test the likelihood ratio of one model to the model with k - 1 profiles (Lo et al., 2001). The LMR test assesses significance across differences in degrees of freedom, including detailing whether the model with fewer profiles fits the data significantly worse than the model with more profiles (Ferguson et al., 2020; Lo et al., 2001). Next, entropy

estimates how separable profiles are from one another and is another source of inferring the putative best-fitting model. Entropy ranges from zero to one where values closer to zero suggests that posterior classification is no better than chance whereas values closer to one suggests closer to perfect posterior classification. However, even with values closer to one, there may be a high degree of misclassification. Notably, classification uncertainty can increase due to chance with more profiles. Therefore, entropy should not be solely used in model selection. All four models significantly improved the preceding n-1 class model according to their LMR-LRT p values. Finally, covariates (i.e., age, sex, race/ethnicity, and family SES) were then added simultaneously to the model following estimation of optimal number of latent profiles to examine their effects as a function of latent profiles and to describe how the covariates differentiate the profiles (Ferguson et al., 2020). The BCH method robustly examines relationships between profiles and covariates and was used to include covariates and auxiliary variables. Multinomial logistic regressions were used at the final step to determine the relationship between profile membership and covariates. For more information on three-step model specification, see Asparouhov & Muthén (2014). After inclusion of the covariates, latent profiles were evaluated again to determine changes in class distribution, which may indicate differential item functioning for profile indicators to the covariates (Masyn, 2017; Nylund-Gibson & Masyn, 2016).

Results

Latent profiles were iteratively fit to the data, starting with one profile and continuing to up to six profiles, at which point relative fit appeared to worsen slightly. Model fit indices for the structure deemed most appropriate are summarized in Table 2.2. Selection of the best fit model was informed by theory and estimated from multiple fit indices. Given that model fit improved only marginally across the increasing number of profiles, a 3-profile solution was selected. The first profile (n = 1576) was conceptualized as High Risk given elevations in family conflict, lower family expression, low family cohesion, low effortful control, high negative affect, and lower

surgency. The second profile (n = 6081) was conceptualized as Normative, reflected in mean levels across all indices. The final profile (n = 3234), considered to be Low Risk, included low family conflict, higher family expression, higher family cohesion, higher effortful control, lower negative affect, and higher surgency. See Figure 2.2 for graphical representation.

Model Validation: Covariates and Levels of Irritability

The BCH approach yielded a significant overall Wald Chi-Square test ($\chi = 1111.27, p < .001$) suggesting significant profile differences, including varying levels of irritability. The High Risk profile was characterized by elevated irritability whereas Normative profile consisted of modest levels of irritability and the Low Risk profile displayed the lowest levels of irritability. Next, we simultaneously introduced age, sex, race-ethnicity, and family income/SES as covariates as a function of the three profiles and evaluated multinomial logistic regressions (summarized on Table 2.4). With the Low Risk profile as the reference group, older and non-white (i.e., “Other” or Hispanic) children were more likely to be in the High Risk profile. Regarding family income, medium, medium/high-, and high-income families were more likely to be in the High Risk profile relative to the Low Risk profile. Also, medium/high and high income families were more likely to be in the Normative profile compared to the Low Risk profile.

Discussion

To identify empirically-based, multivariate profiles of risk for childhood irritability across separate family and temperament dimensions, we utilized data from the ABCD Study, a national longitudinal sample of 11,875 youth (aged nine at baseline). Although there is persuasive evidence that irritability is clinically significant and independently predicts important clinical-health outcomes, surprisingly little is known about its potential risk factors. To reduce its considerable burden, we employed innovative latent profile analyses leveraging naturally-occurring differences in family- (i.e., family conflict, cohesion, expression) and temperament- (i.e., negative affect, effortful control, surgency) based constructs – and their association with irritability. Controlling for age, sex, race-ethnicity, and family income, several key findings were

observed: (a) multiple indicators converged to suggest that a three profile model was a strong fit to the data; (b) the High Risk profile was characterized by elevated family conflict and negative affect, low family expression, family cohesion, effortful control, and modest surgency; (c) the Normative profile was characterized by mean levels of all indicator variables; and (d) the final profile (Low Risk) was characterized by elevated family expression, family cohesion, effortful control, and surgency with low family conflict and negative affect. Overall, the High Risk profile exhibited the highest whereas the Low Risk profile had lowest levels of irritability.

Despite the clinical and etiological heterogeneity of irritability (Stringaris et al., 2017), putative risk factors are myopically considered in isolation; this limitation extends to consideration of the family environment and temperament. In the present study, risk profiles spanning the family environment and temperament were differentially related to levels of childhood irritability. As hypothesized, the High Risk profile (i.e., elevated family conflict and negative affect with low family expression, family cohesion, effortful control, and surgency) exhibited the highest levels irritability relative to individuals in the Moderate and Low Risk profiles. One key distinction of the High Risk profile is the lowest levels of surgency, defined as emotional reactivity in which a person tends towards high levels of positive affect (Rothbart & Bates, 2006). Specifically, surgency has been found to be negatively associated with depression (Kasch et al., 2002) and positively associated with angry/irritable and argumentative/defiant symptom domains (Zastrow et al., 2018). However, surgency was recently unrelated prospectively with childhood irritability specifically (Silver et al., 2023), suggesting a degree of specificity related to risk factors for childhood irritability itself rather than comorbid conditions. In contrast, in the High Risk profile for the current study, lower surgency in addition to other dimensions of family environment and temperament appeared to jointly influence the overall profile and relationship with irritability. On the other hand, sex moderated predictions of internalizing and externalizing problems from irritability (Humphreys et al., 2019), which

suggests that there are likely additional moderators that further specify associations of irritability with psychopathology.

Contextualizing the putative association of surgency with irritability requires consideration of moderating influences, including threat sensitivity, which is defined as the affective, cognitive, behavioral, and physiological responses toward threatening stimuli, information, or cues. Perhaps unsurprisingly, aberrant threat processing or increased orientation to threat is associated with elevated childhood irritability (Stringaris et al., 2017). On the other hand, high levels of surgency may be protective from the consequences of threat sensitivity (Robinson et al., 2005). These distinctions underscore the complex relationship between surgency and irritability beyond simple bivariate associations, signaling a need for innovative approaches to disentangle these associations. For example, experimental manipulation of threat sensitivity – through tasks such as the Fearful Face Detection task (Pessoa et al., 2006) and visual search tasks (Hout et al., 2015) consisting of threatening and nonthreatening conditions – constitutes rigorous tests that further elucidate these associations.

There are notable clinical implications given that adverse family factors and variations in temperament were associated with elevated irritability. Specifically, although intervention science requires strong clinical assessment measures/approaches (Youngstrom & Van Meter, 2016), problematically, there are few standards guiding clinical decision making on what domains to assess. Based on the current study, population-based screening instruments may benefit from focused assessment of temperament and family environment. Related to the High Risk profile, evidence of negative family environments (i.e., high conflict families with low expressive and cohesion) in combination with natural variations of temperament (i.e., high negative affect with low effortful control and surgency), may portend elevated irritability. If replicated, clinically, identified youth may be diverted to prevention-focused programs to prevent psychopathology onset, which is cost-effective and less stigmatizing that interventions specifically targeting dimensions of psychopathology (Le et al., 2021). On the other hand,

individuals from the Normative profile might require different treatment triaging (i.e., lower level of care), including ongoing surveillance. Based on Youngstrom and Van Meter (2016)'s model for clinical management, subgroups from this study that varied across profiles of risk are likely to similarly vary with respect to specific treatment needs (e.g., acute care versus preventative care). This also aligns with Precision Mental Health, a data-driven approach that aims to personalize mental health care through assessment, monitoring, and feedback (Bickman et al., 2016). Moving forward, profiles of risk are well-positioned to inform treatment planning by identifying targeted domains to assess and determining appropriate levels of care based on the assessed risk.

Beyond clinical assessment, clinical decision making may consider profile membership given key familial and temperamental differences. For example, intervention science increasingly recognizes the need to appropriately determine modalities, treatment sequencing, and treatment duration (Keeton et al., 2008; Lambert, 2013; Rush et al., 2004). Heuristically, the Managing and Adapting Practices (MAP) framework identified four prominent components to inform clinical decisions: causal mechanism research (i.e., theories of why something happens), general service research (i.e., findings from evidence-based studies), local aggregate evidence (i.e., what works for specific communities or areas), and case-specific historical information (i.e., what works for specific individuals) (Chorpita et al., 2010). By utilizing information from multiple sources, treatment practices are selected to best suit a client's needs (e.g., parenting practices, emotion regulation skills). Future studies must identify what treatment practices are best suited for individuals from different risk profiles. Similarly, empirically-derived profiles may inform how to sequence multi-modal interventions. For example, in an ongoing sequential multiple assignment randomized control trial, efforts are underway to elucidate sequencing treatment modalities (i.e., formal psychotherapy (CBT) and medication – CBT into medication, CBT into CBT, medication into CBT, medication into medication) that optimizes treatment outcomes for anxious youth (Peterson et al., 2021). Similarly, evidence-based practices for the High Risk

profile might consist of parent management strategies to address the negative family environment and emotion regulation skills to target individual level temperamental factors (Cole et al., 2018; Leijten et al., 2019). However, future research must elucidate if treatment outcome varies as a function of how key intervention practices (e.g., parent management followed by emotion regulation or vice versa) are sequenced.

This study also observed evidence of a unique Low Risk profile consisting of comparatively more family expression, family cohesion, effortful control, and surgency with lower family conflict and negative affect. Individually, there is replicated evidence that these dimensions of family environment and temperament promote adaptive outcomes. For example, surgency is differentially associated with internalizing and externalizing outcomes, although the precise direction of association with these dimensions is unclear (Dolcini-Catania et al., 2020; Zastrow et al., 2018); for example, in a non-treatment sample, children with high surgency showed less aggression in the presence of concurrent elevated social support seeking behaviors (Dollar & Stifter, 2012). This suggests that associations of dimensions of temperament with internalizing and externalizing outcomes must attend to contextual factors and identify other intervening (e.g., moderating, mediating) constructs. As such, while elevated surgency alone may pose as a potential risk factor for psychopathology, in the present study, higher surgency *in combination* with dimensions of family environment and temperament were associated with decreased irritability. This may signal that this specific combination may be protective or may lower overall risk for elevations in irritability.

Although a Low Risk profile suggests few risk indicators, developmentally, these indications may change (including worsening over time). During sensitive periods characterized by new milestones and neurobiological transitions (e.g., adolescence), putatively low risk individuals are not devoid from developing mental health challenges (e.g., adolescent onset conduct disorder) as evidenced by increasing prevalence rates in adolescence (Bitsko et al., 2018). These individuals may still benefit from interventions that would bolster existing strengths

and, it is hoped, prevent onset of new problems and related risk processes (Youngstrom & Van Meter, 2016). Critically, this conceptualization transcends a deficits-based approach and is aligned with a strengths-based approach wherein individual (e.g., temperament and stress) and environmental factors (e.g., family environment, peer influences, natural supports) buffer or prevent poor mental health outcomes. This alleviates burden on healthcare resources by utilizing a prevention model (Mihalopoulos et al., 2011). In a key example of a strengths-based approach, the Positive Youth Development (PYD) framework identified factors or the 5 C's (competence, confidence, connection, character, and caring) that promote resilience and positive outcomes for youth (Lerner, 2009), including decreased rule-breaking, anxiety and depressive symptoms over time (Onyeka et al., 2022). Thus, future studies would continue to benefit from strengths-based approaches that identify risk profiles that may be protective to combat youth mental health problems.

Although important strengths secondary to use of a nationally representative sample with multi-informant/method assessment of key constructs, we emphasize several important limitations. First, although the assessment of childhood irritability is evidence-based (Evans et al., 2019), additional distinctions with respect to irritability (e.g., tonic versus phasic) may reveal different patterns of association (Silver et al., 2021). However, ascertainment of irritability in the present study did not offer such distinctions. Second, the identified profiles associated with varying levels of irritability did not account for potential symptom elevations and functional impairment. That is, despite having a profile associated with elevated irritability, it is unclear whether profiles were related to greater functional impairment or other clinical elevations, which has important clinical implications in initiating treatment. For example, one study identified profiles (e.g., high ADHD/high irritability, low ADHD/high irritability) that divergently predicted outcomes: whereas the high ADHD/ high irritability predicted persistent adolescent ADHD symptoms, externalizing, internalizing, and suicidal outcomes, the low ADHD/ high irritability and high ADHD/ low irritability had weaker associations to these outcomes (Galera et al., 2021).

Thus, despite presenting with high irritability and low ADHD, this combination may incur warrant different treatment considerations relative to youth in the high ADHD/ high irritability profile. To maximize clinical utility, future studies should integrate risk factors, including family environment and temperament, that reflect important clinical, health, and functional outcomes to better discern the consequences of childhood irritability.

Study 2 Tables and Figures

Table 2.1: Descriptive Statistics of Demographic and Study Variables

Variable	Mean (SD)
Age (*SD*)	9.5 (9.0)
Sex (% Male)	52%
Race	
White	52%
Black	15%
Hispanic	20%
Asian	2%
“Other”	11%
Family Income	
Low	15%
Medium	28%
Medium High	45%
High	12%
Irritability	0.80 (1.2)
Family Environment Scale (FES)	
Conflict	2.43 (1.9)
Expression	6.02 (1.5)
Cohesion	7.29 (1.6)
Early Adolescent Temperament Questionnaire (EATQ-R)	
Effortful Control	3.41 (0.62)
Negative Affect	2.41 (0.58)
Surgency	3.39 (0.53)

Table 2.2: Model Fit Comparison Across Class Variance/Covariance Structure

	# Profiles	AIC	BIC	Sample size-adjusted BIC	Entropy	Loglikelihood	LMR Test p value
Invariant Diagonal	3	175139.382	175329.07	175246.446	0.733	-89118.131	<.0001
Invariant Unrestricted	3	171791.146	172090.269	171959.977	**0.832**	-86237.69	<.0001
Variant Diagonal	3	172916.774	173194.01	173073.251	0.652	-87539.26	<.0001
Variant Unrestricted	3	**170205.017**	**170591.689**	**170423.262**	0.627	**-85622.368**	<.0001

Table 2.3: Model Fit Comparison Variant Unrestricted Variance/Covariance Structure

# Profiles	AIC (lower values - better fit)	BIC (lower values - better fit)	Sample size-adjusted BIC (lower - better fit)	% Change	Entropy (higher - better; 0.80 as criteria)	Class Size
2	171324.737	171616.565	171489.48		0.614	c1 = 3220 c2 = 7671
3	170205.017	170591.689	170423.262	**-0.621739596**	**0.627**	c1 = 6081 c2 = 3234 c3 = 1576
4	169439.351	169920.867	169711.127	-0.417862557	0.594	c1 = 1423 c2 = 4176 c3 = 2539 c4 = 2753
5	168804.533	169380.892	169129.841	-0.342514961	0.57	c1 = 1128 c2 = 1894 c3 = 2694 c4 = 2347 c5 = 2828
6	**168590.304**	**169261.507**	**168969.143**	-0.095014575	0.561	c1 = 2082 c2 = 1852 c3 = 1591 c4 = 1876 c5 = 1175 c6 = 2315

Table 2.4: Profile Membership and Descriptive Statistics Per Latent Profile

	n (%)	Conflict M (*SE*)	Expression M (*SE*)	Cohesion M (*SE*)	Effortful Control M (*SE*)	Negative Affect M (*SE*)	Surgency M (*SE*)
Profile 1: High Risk	1576 (14%)	1.08 (*0.058*)	-0.58 (*0.048*)	-1.37 (*0.093*)	-.64 (*0.033*)	0.75 (*0.044*)	-0.28 (*0.033*)
Profile 2: Normative	6081 (56%)	0.043 (*0.044*)	0.057 (*0.036*)	0.052 (*0.030*)	-0.036 (*0.033*)	0.031 (*0.039*)	-0.023 (0.022)
Profile 3: Low Risk	3234 (30%)	-0.76 (*0.051*)	0.25 (*0.044*)	0.75 (*0.046*)	0.47 (*0.033*)	-0.528 (*0.040*)	0.22 (*0.031*)

Table 2.5: Relationship between Profile Membership and Covariates (Odds Ratios)

	Profile 1 vs Profile 3 (ref)		Profile 2 vs Profile 3 (ref)	
	OR	95% CI	OR	95% CI
Age	1.18*	1.01-1.37	0.98	0.86-1.12
Sex (ref = female)	0.89	0.77-1.04	0.92	0.81-1.05
Race – "Other" (ref = White)	1.55*	1.20-2.00	1.25	0.98-1.59
Race – Black (ref = White)	1.11	0.86-1.4	1.02	0.80-1.29
Race – Hispanic (ref = White)	0.67*	0.54-0.85	0.84	0.70-1.00
Race – Asian (ref = White)	1.03	0.61-1.76	0.88	0.55-1.40
Family income – Med (ref – low)	0.48*	0.37-0.61	0.84	0.65-1.08
Family income – Med/High (ref – low)	0.31*	0.24-0.40	0.75*	0.59-0.97
Family income – High (ref – low)	0.26*	0.19-0.36	0.51*	0.38-0.68

Figure 2.1: Proposed Model

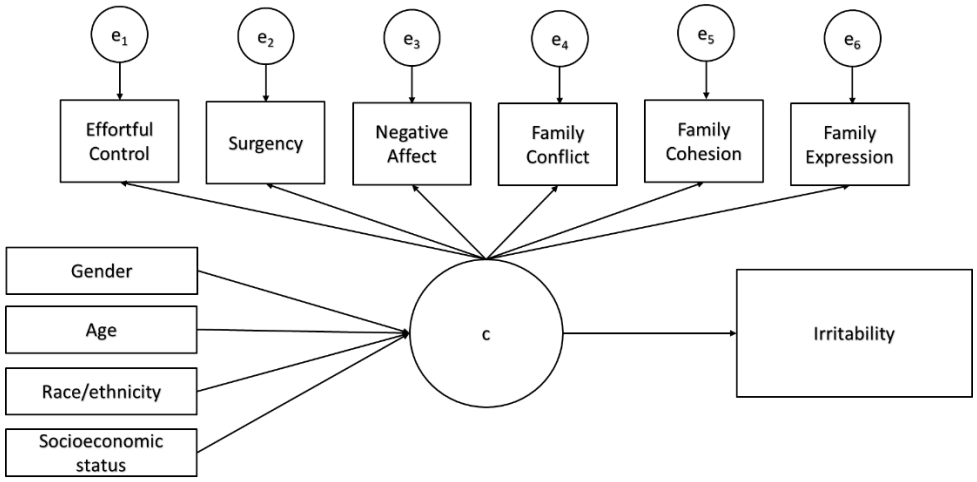
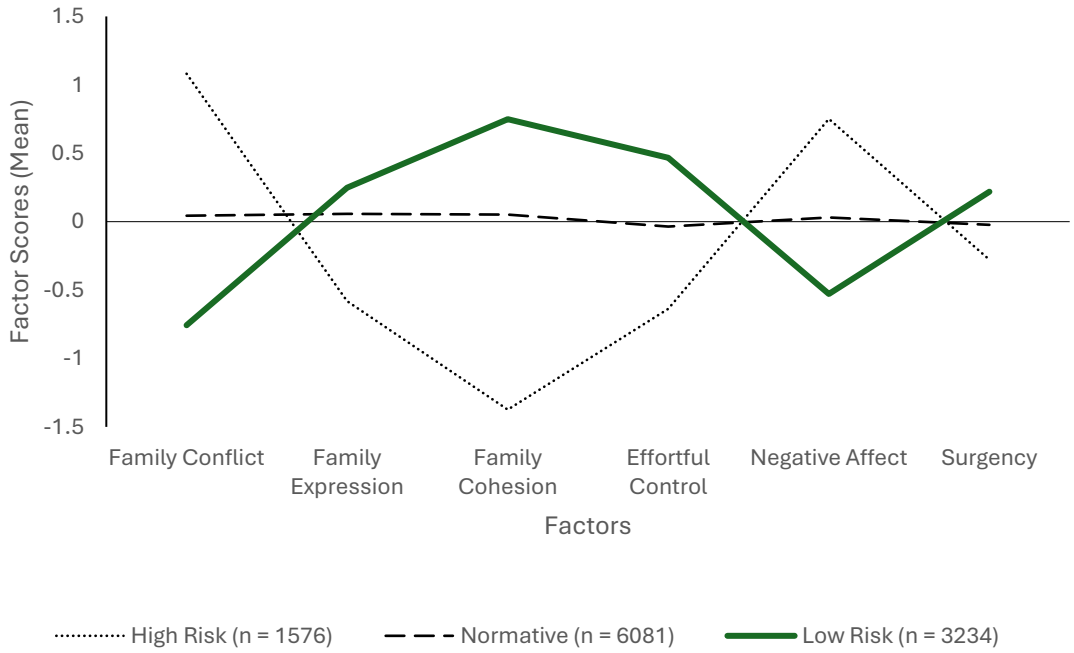


Figure 2.2: Latent Profile Analyses – 3 Factor Model



Conclusion

Recent epidemiological data suggest youth mental health problems are increasing rapidly, including key dimensions of psychopathology peaking acutely during adolescence (Bitsko et al., 2018; Ghandour et al., 2019); these trends were also exacerbated by the COVID-19 pandemic (Hossain et al., 2020). Spanning ADHD and disruptive behavior problems to anxiety and depression, dimensions of psychopathology often co-occur (Ghandour et al., 2019). Typically, co-occurring psychopathology is accompanied by increased impairment, treatment resistance, and greater persistence relative to dimensions presenting singularly (Newman et al., 1998). Overlapping characteristics, such as cognitive impairments and behavioral inhibition, complicate discernment of unique treatment targets, selection of evidence-based interventions, and following treatment protocols to fidelity (Bearman & Weisz, 2015). Identifying modifiable risk factors that precede youth psychopathology is necessary to guide innovations in assessment and intervention for youth mental health problems.

Childhood irritability has garnered increasing clinical and scientific interest (Brotman et al., 2017; Vidal-Ribas et al., 2016). Although developmentally normative, treatment-seeking families endorsed childhood irritability as a top problem, affirming its clinical significance (Evans et al., 2022). However, evidence-based assessment of childhood irritability is still in its infancy; for example, although there are multiple methods to ascertain irritability (e.g., targeted rating scales, subscales from broad measures, and items from diagnostic interviews), there is little evidence guiding clinical decision making with respect to irritability. Although recent studies have evaluated clinical cut-offs for specific rating scales (Wiggins et al., 2017), more fundamentally, there is minimal evidence on how to utilize multi-informant data for childhood irritability. Given that multi-informant data are necessary to capture well-known situational/contextual differences in child behavior, clinicians must ideographically determine how to best combine these data. For example, even when multi-informant data are gathered, they are often treated separately (Orri

et al., 2019) and do not guide crucial clinical inferences (e.g., diagnosis, initiating care; de Los Reyes & Kazdin, 2005).

The present study aimed to address this limitation directly by testing tested specific informant combinations of childhood irritability (i.e., parent only, teacher only, parent “and” teacher, and parent “or” teacher) separately; informants and informant combinations for childhood irritability were subjected to stringent tests of predictive validity based on predictions of independent measures of externalizing and internalizing problems, as well as functional impairment. Parent-rated irritability uniquely predicted clinically significant elevations in important indicators of early adolescent outcomes (Leibenluft et al., 2006; Stringaris et al., 2009). There is replicated evidence that teacher-rated irritability incrementally predicts internalizing and externalizing dimensions (Dwyer et al., 2006; Tandon et al., 2009; Zee & Rudasill, 2011); however, in the current study, teacher ratings specifically predicted externalizing problems (but not internalizing problems and impairment) only. When parent and teacher ratings of irritability were combined using the “or” rule, they prospectively predicted later clinical outcomes. A key challenge going forward will be how and when youth self-report can and should be integrated into these approaches (De Los Reyes & Kazdin, 2005). One promising way of integrating youth report is understanding specific informant *discrepancies*. For example, discrepant maternal and youth report in three domains of parental monitoring behaviors (i.e., child disclosure, parental knowledge, and parent solicitation) were associated with depressive symptoms, even with control of maternal stress, child age, gender, and ethnicity (De Los Reyes et al., 2008). Thus, in addition to understanding the potential clinical utility across different informant combinations, explicit consideration of informant discrepancies, may facilitate case conceptualization and guide treatment recommendations (i.e., ongoing monitoring versus initiating treatment).

To further understand the clinical utility of childhood irritability as an early indicator for later internalizing and externalizing problems as well as functional impairment, we evaluated

informant combinations for childhood irritability with respect to their diagnostic accuracy for clinically elevated outcomes. As a construct, diagnostic accuracy simultaneously reflects performance in detecting true positives but also true negatives (Youngstrom, 2014). First, parent-report only was significantly more accurate than parent “or” teacher ratings for childhood irritability in predicting adolescent internalizing problems. In contrast, diagnostic accuracy was comparable across all four combinations for externalizing outcomes. This aligns with previous findings that parents are often essential to understanding internalizing problems whereas either parent or teachers can be utilized to understand externalizing problems (Sourander & Helstela, 2005). Clinically, this suggests that certain informants may be prioritized depending on the goals of the assessment and/or the outcome of interest. For example, parent report may be prioritized for screening given its clinical utility in detecting internalizing and externalizing outcomes; however, teachers are privy to children’s daily functioning and their data may be prioritized to understand specific domains (e.g., peer relationships, Cullerton-Sen & Crick, 2005).

In the second dissertation study, I aimed to improve evidence on the clinical management of childhood irritability. Despite being considered a top problem in treatment-seeking families, childhood irritability is often monitored in the context of co-occurring disorders (e.g., anxiety, depression, and ODD) and is rarely considered a unique treatment target (Stringaris et al., 2017; Tripepi et al., 2010). Taxonomically, these patterns reflect emerging evidence that dimensions of psychopathology are sensitive to unique and shared risk factors; for example, hierarchical models (HiTOP, Kotov et al., 2017) and bifactor models (general factor of psychopathology – Caspi et al., 2014). However, there is also surprisingly little evidence on specific factors for irritability per se given that irritability is typically evaluated with co-occurring conditions, especially given the reliance on clinic-referred samples. Moreover, beyond individual risk factors, *risk profiles*, reflecting multiple domains, are hypothesized to better reflect the multifactorial nature of childhood irritability. For instance, youth with elevated irritability *and*

depressive/anxious symptoms were twice as vulnerable to exhibit suicidal behaviors relative to youth with low depressive/anxious mood or low irritability (Orri et al., 2018). Thus, the present study utilized a nationally representative sample to understand risk profiles of irritability. A three profile model consisting of both family environmental factors (i.e., family conflict, cohesion, and expression) and individual temperamental factors (i.e., surgency, effortful control, and negative affect) was suggested from the data. This three profile model conferred with varying levels of irritability over and above key demographic variables, including age, sex, race, and family SES, where the High Risk profile demonstrated the highest levels of irritability and the Low Risk profile displayed the lowest levels of irritability. Regarding clinical assessment, population-based screening of temperament and family environment may represent constructs that are observable early into a child's development, thus creating early screening targets that can be utilized across settings. For example, primary care may efficiently assess these factors through routine well child visits to ultimately improve (early) detection and facilitate timely access to mental health services (Nordin et al., 2010; Radez et al., 2020). Therefore, in addition to integrating mental health care within a primary care setting to increase access (Asarnow et al., 2015), creating screening protocols that are specific, easily accessible, and interpretable is a priority moving forward.

Detection of specific risk profiles provides additional guidance towards the clinical management of childhood irritability. Clinical management requires strong clinical assessment to determine possible diagnoses and relevant risk factors. This information can be used to calculate an individual's relative risk, or the likelihood of them having a particular diagnosis, to decide on the need for treatment (i.e., acute treatment, secondary interventions, ongoing monitoring, or preventative care) (Youngstrom & Van Meter, 2016). For instance, the three profile model may represent three probabilities of risk that may require different levels of care given the varying levels of irritability; the High Risk profile had the highest levels whereas the Low Risk profile had the lowest levels. The third profile, Normative, had modest levels of

irritability. The High Risk profile may require acute or secondary treatment, the Normative profile may warrant ongoing monitoring, and the Low Risk profile may capitalize on protective factors that positively affect the youth. Replicated risk profiles may also inform treatment selection and sequencing, including whether treatment outcomes are affected by sequencing of these practices (see Peterson et al., 2021 for a heuristic example of this approach in evaluating sequencing effects of CBT and medication management).

Childhood irritability itself is a clinically significant and impairing problem for children and families. Despite its transdiagnostic significance, including unique predictions of outcome, innovations in intervention and prevention await rigorous evidence on evidence-based clinical decision making (especially with respect to assessment). The present studies innovatively leveraged complementary research strategies – spanning a prospective longitudinal study of well-ascertained children into adolescence and a nationally representative sample of American youth – to build an evidence base for the assessment of childhood irritability and its multivariate risk profiles.

References

Achenbach, T. M., Dumenci, L., Rescorla, L. A., & College, B. M. (2001). *Ratings of Relations Between DSM-IV Diagnostic Categories and Items of the CBCL/6-18, TRF, and YSR*.

Aebi, M., Kuhn, C., Banaschewski, T., Grimmer, Y., Poustka, L., Steinhausen, H. C., & Goodman, R. (2017). The contribution of parent and youth information to identify mental health disorders or problems in adolescents. *Child and adolescent psychiatry and mental health*, *11*, 1-12. https://doi.org/10.1186/s13034-017-0160-9

Agnafors, S., Barmark, M., & Sydsjö, G. (2021). Mental health and academic performance: a study on selection and causation effects from childhood to early adulthood*. Social psychiatry and psychiatric epidemiology*, *56*, 857-866.https://doi.org/10.1007/s00127-020-01934-5

Alexander, L. A., McKnight, P. E., Disabato, D. J., & Kashdan, T. B. (2017). When and How to use Multiple Informants to Improve Clinical Assessments. *Journal of Psychopathology and Behavioral Assessment*, *39*(4), 669–679. https://doi.org/10.1007/S10862-017-9607-9/FIGURES/1

American Academy of Child and Adolescent Psychiatry Committee on Health Care Access and Economics Task Force on Mental Health. (2009). Improving mental health services in primary care: reducing administrative and financial barriers to access and collaboration. *Pediatrics*, *123*(4), 1248-1251.

Asarnow, J. R., Rozenman, M., Wiblin, J., & Zeltzer, L. (2015). Integrated medical-behavioral care compared with usual primary care for child and adolescent behavioral health: A meta-analysis. *JAMA pediatrics*, *169*(10), 929-937. doi:10.1001/jamapediatrics.2015.1141

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Structural equation modeling: A multidisciplinary Journal*, *21*(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Baker-Ericzén, M. J., Jenkins, M. M., & Haine-Schlagel, R. (2013). Therapist, parent, and youth perspectives of treatment barriers to family-focused community outpatient mental health services. *Journal of Child and Family Studies*, *22*, 854-868. https://doi.org/10.1007/s10826-012-9644-7

Barkley, R. A. (2015). Etiologies of ADHD. In R. A. Barkley (Ed.), Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment (4th ed., pp. 356–390). The Guilford Press.

Baxter, A. J., Patton, G., Scott, K. M., Degenhardt, L., & Whiteford, H. A. (2013). Global epidemiology of mental disorders: what are we missing?. *PloS one*, 8(6), e65514. https://doi.org/10.1371/journal.pone.0065514

Bearman, S. K., & Weisz, J. R. (2015). Comprehensive treatments for youth comorbidity–evidence-guided approaches to a complicated problem. *Child and adolescent mental health*, *20*(3), 131-141. https://doi.org/10.1111/camh.12092

Beauchaine, T. P., & Tackett, J. L. (2020). Irritability as a Transdiagnostic Vulnerability Trait: Current Issues and Future Directions. *Behavior Therapy*, *51*(2), 350–364. https://doi.org/10.1016/J.BETH.2019.10.009

Beck, C., Kosnik, C., & Rosales, E. (2017). Longitudinal study of teachers. In Oxford Research Encyclopedia of Education. https://doi.org/10.1093/acrefore/9780190264093.013.117

Bickman, L., Lyon, A. R., & Wolpert, M. (2016). Achieving precision mental health through effective assessment, monitoring, and feedback processes: Introduction to the special issue. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*, 271-276. https://doi.org/10.1007/s10488-016-0718-5

Biederman, J., Monuteaux, M. C., Mick, E., Spencer, T., Wilens, T. E., Silva, J. M., ... & Faraone, S. V. (2006). Young adult outcome of attention deficit hyperactivity disorder: a controlled 10-year follow-up study. *Psychological medicine*, *36*(2), 167-179. https://doi.org/10.1017/S0033291705006410

Biglan, A., Gau, J. M., Jones, L. B., Hinds, E., Rusby, J. C., Cody, C., & Sprague, J. (2015). The role of experiential avoidance in the relationship between family conflict and depression among early adolescents. *Journal of Contextual Behavioral Science*, *4*(1), 30–36. https://doi.org/10.1016/J.JCBS.2014.12.001

Bitsko, R. H., Holbrook, J. R., Ghandour, R. M., Blumberg, S. J., Visser, S. N., Perou, R., & Walkup, J. T. (2018). Epidemiology and Impact of Health Care Provider–Diagnosed Anxiety and Depression Among US Children. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *39*(5), 395. https://doi.org/10.1097/DBP.0000000000000571

Brotman, M. A., Kircanski, K., & Leibenluft, E. (2017). Irritability in children and adolescents. *Annual Review of Clinical Psychology*, *13*, 317-341. https://doi.org/10.1146/ANNUREV-CLINPSY-032816-044941

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/10.1146/annurev-clinpsy-032816-044941

Capaldi, D. M., & Rothbart, M. K. (2016). Development and Validation of an Early Adolescent Temperament Measure: *Http://Dx.Doi.Org/10.1177/0272431692012002002*, *12*(2), 153–173. https://doi.org/10.1177/0272431692012002002

Carver, C. S. (2009). Threat sensitivity, incentive sensitivity, and the experience of relief. *Journal of personality*, *77*(1), 125-138. https://doi.org/10.1111/j.1467-6494.2008.00540.x

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*(2), 119–137. https://doi.org/10.1177/2167702613497473

Chen, W. J., Faraone, S. v., Biederman, J., & Tsuang, M. T. (1994). Diagnostic Accuracy of the Child Behavior Checklist Scales for Attention-Deficit Hyperactivity Disorder: A Receiver-

Operating Characteristic Analysis. *Journal of Consulting and Clinical Psychology*, *62*(5), 1017–1025. https://doi.org/10.1037/0022-006X.62.5.1017

Chorpita, B. F., Daleiden, E. L., & Collins, K. S. (2014). Managing and adapting practice: A system for applying evidence in clinical care with youth and families. *Clinical Social Work Journal, 42*, 134-142. https://doi.org/10.1007/s10615-013-0460-3

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., ... & Starace, N. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice, 18*(2), 154-172.

Cleves, M. A. (2002). From the help desk: Comparing areas under receiver operating characteristic curves from two or more probit or logit models*. The Stata Journal, 2*(3), 301-313. https://doi.org/10.1177/1536867X0200200307

Cohen, J. R., So, F. K., Hankin, B. L., & Young, J. F. (2018). Translating Cognitive Vulnerability Theory into Improved Adolescent Depression Screening: A Receiver Operating Characteristic Approach. *Journal of Clinical Child & Adolescent Psychology*, *48*(4), 582–595. https://doi.org/10.1080/15374416.2017.1416617

Cole, P. M., Lougheed, J. P., Chow, S. M., & Ram, N. (2020). Development of emotion regulation dynamics across early childhood: A multiple time-scale approach. *Affective science*, *1*, 28-41. https://doi.org/10.1007/s42761-020-00004-y

Copeland, W. E., Brotman, M. A., & Costello, E. J. (2015). Normative Irritability in Youth: Developmental Findings from the Great Smoky Mountains Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(8), 635–642. https://doi.org/10.1016/J.JAAC.2015.05.008

Corcoran, J., & Franklin, C. (2002). Multi-systemic risk factors predicting depression, self-esteem, and stress in low SES and culturally diverse adolescents. *Journal of Human Behavior in the Social Environment*, *5*(2), 61-76. https://doi.org/10.1300/J137V05N02_05

Costello, E. J., He, J. P., Sampson, N. A., Kessler, R. C., & Merikangas, K. R. (2014). Services for adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey–Adolescent. *Psychiatric services*, *65*(3), 359-366. https://doi.org/10.1176/appi.ps.201100518

Cornacchio, D., Crum, K. I., Coxe, S., Pincus, D. B., & Comer, J. S. (2016). Irritability and severity of anxious symptomatology among youth with anxiety disorders. Journal of the *American Academy of Child & Adolescent Psychiatry, 55*(1), 54-61. https://doi.org/10.1016/j.jaac.2015.10.007

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological bulletin, 115*(1), 74. https://doi.org/10.1037/0033-2909.115.1.74

Cullerton-Sen, C., & Crick, N. R. (2005). Understanding the effects of physical and relational victimization: The utility of multiple perspectives in predicting social-emotional adjustment. *School psychology review, 34*(2), 147-160. https://doi.org/10.1080/02796015.2005.12086280

de Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The Validity of the Multi-Informant Approach to Assessing Child and Adolescent Mental Health. *Psychological Bulletin*, *141*(4), 858. https://doi.org/10.1037/A0038498

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health*. Psychological bulletin, 141*(4), 858. https://doi.org/10.1037/a0038498

De Los Reyes, A., Goodman, K. L., Kliewer, W., & Reid-Quiñones, K. (2008). Whose depression relates to discrepancies? Testing relations between informant characteristics

and informant discrepancies from both informants' perspectives. *Psychological Assessment, 20*(2), 139–149. https://doi.org/10.1037/1040-3590.20.2.139

de Los Reyes, A., & Kazdin, A. E. (2005). Informant Discrepancies in the Assessment of Childhood Psychopathology: A Critical Review, Theoretical Framework, and Recommendations for Further Study. *Psychological Bulletin*, *131*(4), 483. https://doi.org/10.1037/0033-2909.131.4.483

Dollar, J. M., & Stifter, C. A. (2012). Temperamental surgency and emotion regulation as predictors of childhood social competence*. Journal of Experimental Child Psychology, 112*(2), 178-194. https://doi.org/10.1016/j.jecp.2012.02.004

Dolcini-Catania, L. G., Byrne, M. L., Whittle, S., Schwartz, O., Simmons, J. G., & Allen, N. B. (2020). Temperament and symptom pathways to the development of adolescent depression. *Journal of abnormal child psychology*, *48*, 839-849. https://doi.org/10.1007/s10802-020-00638-3

Dougherty, L. R., Galano, M. M., Chad-Friedman, E., Olino, T. M., Bufferd, S. J., & Klein, D. N. (2021). Using Item Response Theory to Compare Irritability Measures in Early Adolescent and Childhood Samples. *Assessment*, *28*(3), 918–927. https://doi.org/10.1177/1073191120936363

Dougherty, L. R., Smith, V. C., Bufferd, S. J., Kessel, E., Carlson, G. A., & Klein, D. N. (2015). Preschool irritability predicts child psychopathology, functional impairment, and service use at age nine. *Journal of Child Psychology and Psychiatry*, *56*(9), 999–1007. https://doi.org/10.1111/JCPP.12403

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

Dowdy, E., & Kim, E. (2012). *Choosing Informants When Conducting a Universal Screening for Behavioral and Emotional Risk*. School Psychology Forum.

Drabick, D. A., Gadow, K. D., & Loney, J. A. N. (2007). Source-specific oppositional defiant disorder: Comorbidity and risk factors in referred elementary schoolboys. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(1), 92-101. https://doi.org/10.1097/01.chi.0000242245.00174.90

Drabick, D. A. G., & Kendall, P. C. (2010a). Developmental Psychopathology and the Diagnosis of Mental Health Problems among Youth. *Clinical Psychology: A Publication of the Division of Clinical Psychology of the American Psychological Association*, *17*(4), 272. https://doi.org/10.1111/J.1468-2850.2010.01219.X

Drabick, D. A. G., & Kendall, P. C. (2010b). Developmental Psychopathology and the Diagnosis of Mental Health Problems Among Youth. *Clinical Psychology: Science and Practice*, *17*(4), 272–280. https://doi.org/10.1111/J.1468-2850.2010.01219.X

Dwyer, S. B., Nicholson, J. M., & Battistutta, D. (2006). Parent and teacher identification of children at risk of developing internalizing or externalizing mental health problems: A comparison of screening methods. *Prevention science*, *7*, 343-357. https://doi.org/10.1007/s11121-006-0026-5

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

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual review of clinical psychology, 6*, 495-525. https://doi.org/10.1146/annurev.clinpsy.121208.131208

Elvin, O. M., Modecki, K. L., Finch, J., Donnolley, K., Farrell, L. J., & Waters, A. M. (2021). Joining the pieces in childhood irritability: Distinct typologies predict conduct, depressive,

and anxiety symptoms. *Behaviour Research and Therapy, 136*, 103779. https://doi.org/10.1016/j.brat.2020.103779

Evans, S. C., Corteselli, K. A., & Weisz, J. (2022). Is Irritability a Top Problem in Youth Mental Health Care? A Multi-informant, Multi-method Investigation. *Child Psychiatry & Human Development*, 1–15. https://doi.org/10.1007/s10578-021-01301-8

Evans, S. C., Roberts, M. C., Keeley, J. W., Rebello, T. J., de la Pe, F., Lochman, J. E., Burke, J. D., Fite, P. J., Ezpeleta, L., Matthys, W., Youngstrom, E. A., Matsumoto, C., Andrews, H. F., ıa Elena Medina-Mora, M., Ayuso-Mateos, J. L., Khoury, B., Kulygina, M., Robles, R., Sharan, P., … Reed, G. M. (2020). Diagnostic classification of irritability and oppositionality in youth: a global field study comparing ICD-11 with ICD-10 and DSM-5. *Journal of Child Psychology and Psychiatry*, *62*(3), 303–312. https://doi.org/10.1111/jcpp.13244

Evans, S. C., Wei, M. A., Harmon, S. L., & Weisz, J. R. (2021). Modular Psychotherapy Outcomes for Youth with Different Latent Profiles of Irritability and Emotion Dysregulation. *Frontiers in Psychiatry*, *12*, 446. https://doi.org/10.3389/FPSYT.2021.618455/BIBTEX

Ferguson, S. L., Whitney, E., Moore, G., & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*, *44*(5), 458–468. https://doi.org/10.1177/0165025419881721

Forbes, M. K., Rapee, R. M., & Krueger, R. F. (2019). Opportunities for the prevention of mental disorders by reducing general psychopathology in early childhood. *Behaviour Research and Therapy*, *119*, 103411. https://doi.org/10.1016/J.BRAT.2019.103411

Gagne, J. R., van Hulle, C. A., Aksan, N., Essex, M. J., & Goldsmith, H. H. (2011). Deriving Childhood Temperament Measures from Emotion-eliciting Behavioral Episodes: Scale

Construction and Initial Validation. *Psychological Assessment*, *23*(2), 337. https://doi.org/10.1037/A0021746

Galera, C., Orri, M., Vergunst, F., Melchior, M., Van der Waerden, J., Bouvard, M. P., ... & Côté, S. M. (2021). Developmental profiles of childhood attention-deficit/hyperactivity disorder and irritability: Association with adolescent mental health, functional impairment, and suicidal outcomes. *Journal of child psychology and psychiatry, 62*(2), 232-243. https://doi.org/10.1111/jcpp.13270

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., Jernigan, T., Potter, A., Thompson, W., & Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience*, *32*, 16–22. https://doi.org/10.1016/J.DCN.2018.04.004

Ghandour, R. M., Sherman, L. J., Vladutiu, C. J., Ali, M. M., Lynch, S. E., Bitsko, R. H., & Blumberg, S. J. (2019). Prevalence and Treatment of Depression, Anxiety, and Conduct Problems in US Children. *The Journal of Pediatrics, 206*, 256-267.e3. https://doi.org/10.1016/J.JPEDS.2018.09.021

Gloster, A. T., Lamnisos, D., Lubenko, J., Presti, G., Squatrito, V., Constantinou, M., ... & Karekla, M. (2020). Impact of COVID-19 pandemic on mental health: An international study. *PloS one*, *15*(12), e0244809.

Goldsmith, H. H., & Gagne, J. R. (2012). Behavioral Assessment of Temperament. In M. Zentner & R. L. Shiner (Eds.), *Handbook of Temperament* (pp. 209–228).

Goodman, S. H., Lahey, B. B., Fielding, B., Dulcan, M., Narrow, W., & Regier, D. (1997). Representativeness of clinical samples of youths with mental disorders: A preliminary population-based study. *Journal of Abnormal Psychology*, *106*(1), 3–14. https://doi.org/10.1037/0021-843X.106.1.3

Haller, S. P., Stoddard, J., Botz-Zapp, C., Clayton, M., MacGillivray, C., Perhamus, G., ... & Brotman, M. A. (2022). A randomized controlled trial of computerized interpretation bias

training for disruptive mood dysregulation disorder: a fast-fail study. *Journal of the American Academy of Child & Adolescent Psychiatry, 61*(1), 37-45. https://doi.org/10.1016/j.jaac.2021.05.022

Higa-McMillan, C. K., Francis, S. E., Rith-Najarian, L., & Chorpita, B. F. (2016). Evidence base update: 50 years of research on treatment for child and adolescent anxiety. *Journal of Clinical Child & Adolescent Psychology, 45*(2), 91-113. https://doi.org/10.1080/15374416.2015.1046177

Hirsch, E., Davis, K., Cao, Z., & Roy, A. K. (2022). Understanding phasic irritability: Anger and distress in children's temper outbursts. *Child Psychiatry & Human Development, 53*(2), 317-329. https://doi.org/10.1007/s10578-021-01126-5

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive therapy and research*, *36*, 427-440.

Hout, M. C., Walenchok, S. C., Goldinger, S. D., & Wolfe, J. M. (2015). Failures of perception in the low-prevalence effect: Evidence from active and passive visual search. *Journal of Experimental Psychology: Human Perception and Performance, 41*(4), 977. https://doi.org/10.1037/xhp0000053

Howard, M. C., & Hoffman, M. E. (2017). Variable-Centered, Person-Centered, and Person-Specific Approaches: Where Theory Meets the Method. *21*(4), 846–876. https://doi.org/10.1177/1094428117744021

Humphreys, K. L., Schouboe, S. N. F., Kircanski, K., Leibenluft, E., Stringaris, A., & Gotlib, I. H. (2019). Irritability, Externalizing, and Internalizing Psychopathology in Adolescence: Cross-Sectional and Longitudinal Associations and Moderation by Sex. *Journal of Clinical Child & Adolescent Psychology*, *48*(5), 781–789. https://doi.org/10.1080/15374416.2018.1460847

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annu. Rev. Clin. Psychol.*, *3*, 29-51. https://doi.org/10.1146/ANNUREV.CLINPSY.3.022806.091419

Ismatullina, V., & Voronin, I. (2017). Individual Differences in the Relationship between Temperament and Planning Ability in Adolescents. *Procedia - Social and Behavioral Sciences*, *237*, 1455–1461. https://doi.org/10.1016/J.SBSPRO.2017.02.230

Isaac, L., Sutton, T., Kahlon, J., Pathania, P. S., Wolf, B. L., Pearce, R., ... & Zils, M. J. (2024). Child-adolescent emergency psychiatry: addressing false positive admissions. *Frontiers in Psychiatry, 15,* 1321702. https://doi.org/10.3389/fpsyt.2024.1321702

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of abnormal psychology, 111*(4), 589. https://doi.org/10.1037/0021-843X.111.4.589

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., & Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial Reliability and Validity Data. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(7), 980–988. https://doi.org/10.1097/00004583-199707000-00021

Keiley, M. K., Lofthouse, N., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2003). Differential risks of covarying and pure components in mother and teacher reports of externalizing and internalizing behavior across ages 5 to 14. *Journal of abnormal child psychology, 31*, 267-283. https://doi.org/10.1023/A:1023277413027

Kessel, E. M., Dougherty, L. R., Hubacheck, S., Chad-Friedman, E., Olino, T., Carlson, G. A., & Klein, D. N. (2021). Early Predictors of Adolescent Irritability. *Child and Adolescent Psychiatric Clinics*, *30*(3), 475–490. https://doi.org/10.1016/J.CHC.2021.04.002

Keeton, C. P., & Ginsburg, G. S. (2008). Combining and sequencing medication and cognitive-behaviour therapy for childhood anxiety disorders. *International Review of Psychiatry, 20*(2), 159-164. https://doi.org/10.1080/09540260801889047

Kiff, C. J., Lengua, L. J., & Zalewski, M. (2011). Nature and Nurturing: Parenting in the Context of Child Temperament. *Clinical Child and Family Psychology Review 2011 14:3*, *14*(3), 251–301. https://doi.org/10.1007/S10567-011-0093-4

Kishida, K., Tsuda, M., Takahashi, F., & Ishikawa, S. ichi. (2022). Irritability and mental health profiles among children and adolescents: A result of latent profile analysis. *Journal of Affective Disorders*, *300*, 76–83. https://doi.org/10.1016/J.JAD.2021.12.045

Kopala-Sibley, D. C., Olino, T., Durbin, E., Dyson, M. W., & Klein, D. N. (2018). The Stability of Temperament from Early Childhood to Early Adolescence: A Multi-method, Multi-informant Examination. *European Journal of Personality*, *32*(2), 128–145. https://doi.org/10.1002/PER.2151

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M.,…Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*(4), 454–477. http://dx.doi.org/10.1037/abn0000258

Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1997). Coming to Terms with the Terms of Risk. *Archives of General Psychiatry*, *54*(4), 337–343. https://doi.org/10.1001/ARCHPSYC.1997.01830160065009

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry*, *160*(9), 1566–1577. https://doi.org/10.1176/APPI.AJP.160.9.1566/ASSET/IMAGES/LARGE/L03F3.JPEG

Kroneman, L. M., Loeber, R., Hipwell, A. E., & Koot, H. M. (2009). Girls' disruptive behavior and its relationship to family functioning: A review. *Journal of child and family studies*, *18*, 259-273. https://doi.org/10.1007/s10826-008-9226-x

Lahey, B. B. (2009). Public Health Significance of Neuroticism. *The American Psychologist*, *64*(4), 241. https://doi.org/10.1037/A0015309

Lambert, M. J. (Ed.). (2013). Bergin and Garfield's handbook of psychotherapy and behavior change. John Wiley & Sons.

Langley, A. K., Lewin, A. B., Bergman, R. L., Lee, J. C., & Piacentini, J. (2010). Correlates of comorbid anxiety and externalizing disorders in childhood obsessive compulsive disorder. *European Child and Adolescent Psychiatry*, *19*(8), 637–645. https://doi.org/10.1007/S00787-010-0101-0/TABLES/2

Lapalme, M., Bégin, V., le Corff, Y., & Déry, M. (2020). Comparison of Discriminant Validity Indices of Parent, Teacher, and Multi-Informant Reports of Behavioral Problems in Elementary Schoolers. *Journal of Psychopathology and Behavioral Assessment*, *42*(1), 58–68. https://doi.org/10.1007/S10862-019-09782-7/TABLES/2

Lavigne, J. v., Dahl, K. P., Gouze, K. R., LeBailly, S. A., & Hopkins, J. (2015). Multi-Domain Predictors of Oppositional Defiant Disorder Symptoms in Preschool Children: Cross-Informant Differences. *Child Psychiatry and Human Development*, *46*(2), 308–319. https://doi.org/10.1007/S10578-014-0472-4/TABLES/2

Lanza, S. T., & Cooper, B. R. (2016). Latent Class Analysis for Developmental Research. *Child Development Perspectives*, *10*(1), 59–64. https://doi.org/10.1111/CDEP.12163

Le, L. K. D., Esturas, A. C., Mihalopoulos, C., Chiotelis, O., Bucholc, J., Chatterton, M. L., & Engel, L. (2021). Cost-effectiveness evidence of mental health prevention and promotion interventions: A systematic review of economic evaluations. *PLoS medicine, 18*(5), e1003606. https://doi.org/10.1371/journal.pmed.1003606

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, *25*, 1473–1487. https://doi.org/10.1017/S0954579413000722

Leijten, P., Gardner, F., Melendez-Torres, G. J., Van Aar, J., Hutchings, J., Schulz, S., ... & Overbeek, G. (2019). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(2), 180-190. https://doi.org/10.1016/j.jaac.2018.07.900

Lerner, R. M., von Eye, A., Lerner, J. V., & Lewin-Bizan, S. (2009). Exploring the foundations and functions of adolescent thriving within the 4-H study of positive youth development: A view of the issues. *Journal of Applied Developmental Psychology, 30*(5), 567-570. https://doi.org/10.1016/j.appdev.2009.07.002

Lewis, C. C., Simons, A. D., & Kim, H. K. (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology*, *80*(4), 525. https://doi.org/10.1037/a0029131

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

Madigan, S., Korczak, D. J., Vaillancourt, T., Racine, N., Hopkins, W. G., Pador, P., ... & Neville, R. D. (2023). Comparison of paediatric emergency department visits for attempted suicide, self-harm, and suicidal ideation before and during the COVID-19 pandemic: a systematic review and meta-analysis. *The Lancet Psychiatry*, *10*(5), 342-351. DOI:https://doi.org/10.1016/S2215-0366(23)00036-6

Mallidi, A., Meza-Cervera, T., Kircanski, K., Stringaris, A., Brotman, M. A., Pine, D. S., ... & 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

Majdandžić, M., & van den Boom, D. C. (2007). Multimethod Longitudinal Assessment of Temperament in Early Childhood. *Journal of Personality*, *75*(1), 121–168. https://doi.org/10.1111/J.1467-6494.2006.00435.X

Martinelli, A., Ackermann, K., Bernhard, A., Freitag, C. M., & Schwenck, C. (2018). Hostile attribution bias and aggression in children and adolescents: A systematic literature review on the influence of aggression subtype and gender. *Aggression and violent behavior, 39,* 25-32. https://doi.org/10.1016/j.avb.2018.01.005

Masyn, K. E. (2017). Measurement invariance and differential item functioning in latent class analysis with stepwise multiple indicator multiple cause modeling. Structural Equation Modeling: A Multidisciplinary Journal, 24(2), 180-197.

Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. https://doi.org/10.1093/oxfordhb/9780199934898.013.0025

Mihalopoulos, C., Vos, T., Pirkis, J., & Carter, R. (2011). The economic analysis of prevention in mental health programs. *Annual review of clinical psychology, 7,* 169-201. https://doi.org/10.1146/annurev-clinpsy-032210-104601

Millman, Z. B., & Schiffman, J. (2017). False-positives and Clinical Heterogeneity among Youth at Clinical High-risk for Psychosis: Clinical and Ethical Implications for Assessment and Treatment. *Journal of Ethics in Mental Health*.

Moos, R. H. (1994). Family environment scale manual: Development, applications, research. *Consulting Psychologists Press*.

Monducci, E., Battaglia, C., Forte, A., Masillo, A., Telesforo, L., Carlotto, A., ... & Fiori Nastro, P. (2018). Secondary school teachers and mental health competence: Italy–United Kingdom comparison. *Early intervention in psychiatry, 12*(3), 456-463. https://doi.org/10.1111/eip.12345

Naragon-Gainey, K., Rutter, L. A., & Brown, T. A. (2014). The interaction of extraversion and anxiety sensitivity on social anxiety: evidence of specificity relative to depression. *Behavior therapy, 45*(3), 418-429. https://doi.org/10.1016/j.beth.2014.01.004

Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology, 107*(2), 305–311. https://doi.org/10.1037/0021-843X.107.2.305

Nordin, J. D., Solberg, L. I., & Parker, E. D. (2010). Adolescent primary care visit patterns. *Annals of family medicine, 8*(6), 511–516. https://doi.org/10.1370/afm.1188

Nylund-Gibson, K., & Masyn, K. E. (2016). Covariates and mixture modeling: Results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling: A Multidisciplinary Journal, 23(*6), 782-797. https://doi.org/10.1080/10705511.2016.1221313

Onyeka, O., Richards, M., Tyson McCrea, K., Miller, K., Matthews, C., Donnelly, W., ... & Swint, K. (2022). The role of positive youth development on mental health for youth of color living in high-stress communities: A strengths-based approach. *Psychological services, 19*(S1), 72. https://doi.org/10.1037/ser0000593

Orri, M., Galera, C., Turecki, G., Boivin, M., Tremblay, R. E., Geoffroy, M. C., & Côté, S. M. (2019a). Pathways of Association Between Childhood Irritability and Adolescent Suicidality. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(1), 99-107.e3. https://doi.org/10.1016/J.JAAC.2018.06.034

Orri, M., Galera, C., Turecki, G., Boivin, M., Tremblay, R. E., Geoffroy, M. C., & Côté, S. M. (2019b). Pathways of Association Between Childhood Irritability and Adolescent Suicidality. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(1), 99-107.e3. https://doi.org/10.1016/J.JAAC.2018.06.034

Orri, M., Galera, C., Turecki, G., Forte, A., Renaud, J., Boivin, M., ... & Geoffroy, M. C. (2018). Association of childhood irritability and depressive/anxious mood profiles with adolescent suicidal ideation and attempts. *JAMA psychiatry, 75*(5), 465-473. doi:10.1001/jamapsychiatry.2018.0174

Parra, G. R., DuBois, D. L., & Sher, K. J. (2010). Investigation of Profiles of Risk Factors for Adolescent Psychopathology: A Person-Centered Approach. *35*(3), 386–402. https://doi.org/10.1207/S15374424JCCP3503_4

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: a global public-health challenge. *The lancet, 369*(9569), 1302-1313. DOI:https://doi.org/10.1016/S0140-6736(07)60368-7

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafò, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science*, *24*(5), 688–697. https://doi.org/10.1177/0956797612459657

Pepe, M. S., Janes, H., Longton, G., Leisenring, W., & Newcomb, P. (2004). Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American journal of epidemiology*, *159*(9), 882-890. https://doi.org/10.1093/aje/kwh101

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral cortex, 16*(3), 366-375. https://doi.org/10.1093/cercor/bhi115

Peterson, B. S., West, A. E., Weisz, J. R., Mack, W. J., Kipke, M. D., Findling, R. L., ... & Weersing, R. (2021). A Sequential Multiple Assignment Randomized Trial (SMART) study of medication and CBT sequencing in the treatment of pediatric anxiety disorders. *BMC psychiatry, 21*(1), 323. https://doi.org/10.1186/s12888-021-03314-y

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining Discrepant Diagnostic Information from Multiple Sources. *Abnormal Child Psychology*, *20*, 51.

Radez, J., Reardon, T., Creswell, C., Lawrence, P. J., Evdoka-Burton, G., & Waite, P. (2021). Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies.

European child & adolescent psychiatry, 30(2), 183-211. https://doi.org/10.1007/s00787-019-01469-4

Ramsden, S. R., & Hubbard, J. A. (2002). Family expressiveness and parental emotion coaching: Their role in children's emotion regulation and aggression. *Journal of abnormal child psychology, 30*, 657-667.. https://doi.org/10.1023/A:1020819915881

Ravi, S., Havewala, M., Kircanski, K., Brotman, M. A., Schneider, L., Degnan, K., ... & Filippi, C. (2023). Parenting and childhood irritability: Negative emotion socialization and parental control moderate the development of irritability*. Development and Psychopathology, 35*(3), 1444-1453. doi:10.1017/S0954579421001346

Resnick, M. D. (2000). Protective factors, resiliency, and healthy youth development. Adolescent medicine: *State of the art reviews, 11*(1), 157-164.

Rettew, D. C., & McKee, L. (2005). Temperament and Its Role in Developmental Psychopathology. *Harvard Review of Psychiatry*, *13*(1), 14–27. https://doi.org/10.1080/10673220590923146

Rice, M. E., & Harris, G. T. (2005). Comparing Effect Sizes in Follow-Up Studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5). https://doi.org/10.1007/s10979-005-6832-7

Rothbart, M. K., & Bates, J. E. (1988). Temperament. In *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (Vol. 3, pp. 105–176).

Rothbart, M. K., Derryberry, D., Hershey, L., Molfese, V. K., & Molfese, D. L. (2000). *Temperament and personality development across the lifespan*. Lawrence Erbeum Associates.

Rucinski, C. L., Brown, J. L., & Downer, J. T. (2018). Teacher–child relationships, classroom climate, and children's social-emotional and academic development. *Journal of Educational Psychology*, *110*(7), 992. https://doi.org/10.1037/edu0000240

Rush, A. J., Fava, M., Wisniewski, S. R., Lavori, P. W., Trivedi, M. H., Sackeim, H. A., ... & Star* D Investigators Group. (2004). Sequenced treatment alternatives to relieve depression (STAR* D): rationale and design. *Controlled clinical trials, 25*(1), 119-142. https://doi.org/10.1016/S0197-2456(03)00112-0

Saah, A. J., & Hoover, D. R. (1997). “Sensitivity” and “specificity” reconsidered: The meaning of these terms in analytical and diagnostic settings. *Annals of Internal Medicine*, *126*(1), 91–94. https://doi.org/10.7326/0003-4819-126-1-199701010-00026

Shaffer, D., Gould, M. S., Brasic, J., Fisher, P., Aluwahlia, S., & Bird, H. (1983). A Children’s Global Assessment Scale (CGAS). *Archives of General Psychiatry*, *40*(11), 1228–1231. https://doi.org/10.1001/ARCHPSYC.1983.01790100074010

Shemmassian, S. K., & Lee, S. S. (2017). Comparative Validity of DSM-IV and Alternative Empirically Derived Approaches for the Assessment of ADHD. *Journal of Attention Disorders*, *21*(5), 405–415. https://doi.org/10.1177/1087054714522511

Silver, J., Carlson, G. A., Olino, T. M., Perlman, G., Mackin, D., Kotov, R., & Klein, D. N. (2021). Differential outcomes of tonic and phasic irritability in adolescent girls. *Journal of Child Psychology and Psychiatry, 62*(10), 1220-1227. https://doi.org/10.1111/jcpp.13402

Silver, J., Hawes, M., Dougherty, L., Bufferd, S., Kessel, E., Olino, T., ... & Klein, D. (2024). Irritability and temperament: concurrent and prospective relationships in childhood and adolescence. *Journal of Clinical Child & Adolescent Psychology*, *53*(2), 156-168. https://doi.org/10.1080/15374416.2023.2286586

Sorcher, L. K., Goldstein, B. L., Finsaas, M. C., Carlson, G. A., Klein, D. N., & Dougherty, L. R. (2021). Preschool Irritability Predicts Adolescent Psychopathology and Functional Impairment: A 12-Year Prospective Study. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/J.JAAC.2021.08.016

Sourander, A., & Helstelä, L. (2005). Childhood predictors of externalizing and internalizing problems in adolescence: A prospective follow-up study from age 8 to16. *European child & adolescent psychiatry, 14,* 415-423. https://doi.org/10.1007/s00787-005-0475-6

Stocker, C. M., Richmond, M. K., Rhoades, G. K., & Kiang, L. (2007). Family emotional processes and adolescents' adjustment. *Social Development, 16*(2), 310-325. https://doi.org/10.1111/j.1467-9507.2007.00386.x

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*, 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/ASSET/IMAGES/LARGE/U816T5.JPEG

Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., & Brotman, M. A. (2012). The Affective Reactivity Index: a concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry*, *53*(11), 1109–1117. https://doi.org/10.1111/j.1469-7610.2012.02561.x

Stringaris, A., Maughan, B., & Goodman, R. (2010a). What's in a Disruptive Disorder? Temperamental Antecedents of Oppositional Defiant Disorder: Findings from the Avon Longitudinal Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(5), 474–483. https://doi.org/10.1016/j.jaac.2010.01.021

Stringaris, A., Maughan, B., & Goodman, R. (2010b). What's in a Disruptive Disorder? Temperamental Antecedents of Oppositional Defiant Disorder: Findings from the Avon Longitudinal Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(5), 474–483. https://doi.org/10.1016/j.jaac.2010.01.021

Swets, J. A. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*, *240*(4857), 1285–1293. https://doi.org/10.1126/SCIENCE.3287615

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological Science Can Improve Diagnostic Decisions. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, *1*(1), 1–26. https://doi.org/10.1111/1529-1006.001

Stringaris, A., Vidal-Ribas, P., Brotman, M. A., & Leibenluft, E. (2017). *Practitioner Review: Definition, recognition, and treatment challenges of irritability in young people*. https://doi.org/10.1111/jcpp.12823

Tandon, M., Cardeli, E., & Luby, J. (2009). Internalizing disorders in early childhood: A review of depressive and anxiety disorders. *Child and Adolescent Psychiatric Clinics, 18*(3), 593-610. DOI:https://doi.org/10.1016/j.chc.2009.03.004

Tripepi, G., Jager, K. J., Dekker, F. W., & Zoccali, C. (2010b). Selection Bias and Information Bias in Clinical Research. *Nephron Clinical Practice*, *115*(2), c94–c99. https://doi.org/10.1159/000312871

van As, N. M., & Janssens, J. M. (2002). Relationships between child behavior problems and family functioning: A literature review. *International Journal of Child & Family Welfare, 5*(1-2), 40-51.

Van Dis, E. A., Van Veen, S. C., Hagenaars, M. A., Batelaan, N. M., Bockting, C. L., Van Den Heuvel, R. M., ... & Engelhard, I. M. (2020). Long-term outcomes of cognitive behavioral therapy for anxiety-related disorders: a systematic review and meta-analysis. *JAMA psychiatry*, *77*(3), 265-273. doi:10.1001/jamapsychiatry.2019.3986

Verona, E., & Bresin, K. (2015). Aggression proneness: Transdiagnostic processes involving negative valence and cognitive systems. *International Journal of Psychophysiology, 98*(2), 321-329. https://doi.org/10.1016/j.ijpsycho.2015.03.008

Vidal-Ribas, P., & Stringaris, A. (2021). How and Why Are Irritability and Depression Linked? *Child and Adolescent Psychiatric Clinics of North America*, *30*(2), 401–414. https://doi.org/10.1016/j.chc.2020.10.009

Wang, P. S., Berglund, P. A., Olfson, M., & Kessler, R. C. (2004). Delays in Initial Treatment Contact after First Onset of a Mental Disorder. *Health Services Research*, *39*(2), 393–416. https://doi.org/10.1111/J.1475-6773.2004.00234.X

Weist, M. D., Evans, S. W., & Lever, N. A. (Eds.). (2003). Handbook of school mental health: Advancing practice and research (pp. 1-7). New York: Kluwer Academic/Plenum Publishers.

Whelan, Y. M., Leibenluft, E., Stringaris, A., & Barker, E. D. (2015a). Pathways from maternal depressive symptoms to adolescent depressive symptoms: the unique contribution of irritability symptoms. *Journal of Child Psychology and Psychiatry*, *56*(10), 1092–1100. https://doi.org/10.1111/jcpp.12395

Whelan, Y. M., Leibenluft, E., Stringaris, A., & Barker, E. D. (2015b). Pathways from maternal depressive symptoms to adolescent depressive symptoms: the unique contribution of irritability symptoms. *Journal of Child Psychology and Psychiatry*, *56*(10), 1092–1100. https://doi.org/10.1111/jcpp.12395

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.e2. https://doi.org/10.1016/J.JAAC.2017.12.008

Wiggins, J. L., Mitchell, C., Stringaris, A., & Leibenluft, E. (2014). Developmental Trajectories of Irritability and Bidirectional Associations with Maternal Depression. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(11), 1191-1205.e4. https://doi.org/10.1016/J.JAAC.2014.08.005

Youngstrom, E. A. (2013). Future Directions in Psychological Assessment: Combining Evidence-Based Medicine Innovations with Psychology's Historical Strengths to Enhance Utility. *Journal of Clinical Child & Adolescent Psychology*, *42*(1), 139–159. https://doi.org/10.1080/15374416.2012.736358

Youngstrom, E. A. (2014). A Primer on Receiver Operating Characteristic Analysis and Diagnostic Efficiency Statistics for Pediatric Psychology: We Are Ready to ROC. *Journal of Pediatric Psychology*, *39*(2), 204–221. https://doi.org/10.1093/JPEPSY/JST062

Youngstrom, E. A., & Van Meter, A. (2016). Empirically supported assessment of children and adolescents. *Clinical Psychology: Science and Practice, 23*(4), 327. doi:10.1111/cpsp.12172

Zastrow, B. L., Martel, M. M., & Widiger, T. A. (2018). Preschool oppositional defiant disorder: A disorder of negative affect, surgency, and disagreeableness. *Journal of Clinical Child & Adolescent Psychology, 47*(6), 967-977. https://doi.org/10.1080/15374416.2016.1225504

Zee, M., & Rudasill, K. M. (2021). Catching sight of children with internalizing symptoms in upper elementary classrooms. *Journal of school psychology, 87*, 1-17. https://doi.org/10.1016/j.jsp.2021.05.002

Zendarski, N., Galligan, R., Coghill, D., Payne, J. M., De Luca, C. R., & Mulraney, M. (2023). The associations between child irritability, parental distress, parental irritability and family functioning in children accessing mental health services. *Journal of Child and Family Studies, 32*(1), 288-300. https://doi.org/10.1007/s10826-022-02390-2

Zik, J., Deveney, C. M., Ellingson, J. M., Haller, S. P., Kircanski, K., Cardinale, E. M., Brotman, M. A., & Stoddard, J. (2021). Understanding Irritability in Relation to Anger, Aggression, and Informant in a Pediatric Clinical Population. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/J.JAAC.2021.08.012

Zucker, R. A., Gonzalez, R., Feldstein Ewing, S. W., Paulus, M. P., Arroyo, J., Fuligni, A., Morris, A. S., Sanchez, M., & Wills, T. (2018). Assessment of culture and environment in the Adolescent Brain and Cognitive Development Study: Rationale, description of measures, and early data. *Developmental Cognitive Neuroscience*, *32*, 107–120. https://doi.org/10.1016/J.DCN.2018.03.004