Dimensions of irritability in adolescents: longitudinal associations with psychopathology in adulthood

Mariah T. Hawes[1], Gabrielle A. Carlson[2], Megan C. Finsaas[1], Thomas M. Olino[3], John R. Seely[4] and Daniel N. Klein[1]

[1]Stony Brook University, Stony Brook, NY, USA; [2]Stony Brook University Medical Center, Stony Brook, NY, USA; [3]Temple University, Philadelphia, PA, USA and [4]University of Oregon, Eugene, OR, USA

Abstract

Background. There is an emerging consensus in developmental psychopathology that irritable youth are at risk for developing internalizing problems later in life. The current study explored if irritability in youth is multifactorial and the impact of irritability dimensions on psychopathology outcomes in adulthood.

Methods. We conducted exploratory factor analysis on irritability symptom items from a semi-structured diagnostic interview administered to a community sample of adolescents (ages 14–19; 42.7% male; 89.1% white). The analysis identified two factors corresponding to items from the mood disorders *v.* the oppositional defiant disorder (ODD) (Leibenluft and Stoddard) sections of the interview. These factors were then entered together into regression models predicting psychopathology assessed at age 24 ($N = 941$) and again at age 30 ($N = 816$). All models controlled for concurrent psychopathology in youth.

Results. The two irritability dimensions demonstrated different patterns of prospective relationships, with items from the ODD section primarily predicting externalizing psychopathology, items from the mood disorder sections predicting depression at age 24 but not 30, and both dimensions predicting borderline personality disorder symptoms.

Conclusions. These results suggest that the current standard of extracting and compositing irritability symptom items from diagnostic interviews masks distinct dimensions of irritability with different psychopathological outcomes. Additionally, these findings add nuance to the prevailing notion that irritability in youth is specifically linked to later internalizing problems. Further investigation using more sensitive and multifaceted measures of irritability are needed to parse the meaning and clinical implications of these dimensions.

Introduction

Irritability is a transdiagnostic symptom that has garnered a great deal of attention in the developmental psychopathology literature of late. Partially in response to an alarming increase in rates of bipolar disorder diagnoses in young children (Blader and Carlson, 2007), one line of research has explored alternative classification for children who exhibit severe, chronic irritability. Led by Leibenluft and colleagues, this research culminated in the addition of Disruptive Mood Dysregulation Disorder (DMDD) to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013). DMDD is a childhood-onset diagnosis in the depressive disorder section, defined by irritable or angry mood and severe temper outbursts. A related line of research has explored the dimensionality of oppositional defiant disorder (ODD) (Leibenluft and Stoddard) symptoms, identifying a relatively stable irritable symptom dimension characterized by anger, touchiness, and temper outbursts. Compared to other ODD symptom dimensions (e.g. hurtful, headstrong), the irritable symptom dimension demonstrates differential concurrent and longitudinal correlates (e.g. Stringaris and Goodman, 2009, Burke, 2012, Whelan *et al.*, 2013).

Across disorders, irritability has emerged as a clinically significant phenomenon, marking greater severity of psychopathology and functional impairment (Copeland *et al.*, 2015, Wakschlag *et al.*, 2015). There has been particular interest in exploring links between irritability in childhood and adolescence with psychopathology later in life. The consensus emerging from this work is that chronically irritable children and adolescents go on to experience internalizing psychopathology later in life, particularly depression. This work has been summarized in a number of qualitative reviews (Burke and Loeber, 2010; Leibenluft and Stoddard, 2013; Evans *et al.*, 2017) and a recent meta-analysis (Vidal-Ribas *et al.*, 2016). While this evidence base is substantial, the literature has several significant gaps and limitations.

First, there has been little attention to variability in the way that irritability is assessed. Most commonly, irritability symptom items are extracted from diagnostic interviews geared to

assessing DSM criteria for specific disorders. Most studies use an aggregate of the touchy or easily annoyed, angry or resentful and temper tantrum items from the ODD section of interviews (Stringaris and Goodman, 2009; Rowe *et al.*, 2010; Burke, 2012; Stringaris *et al.*, 2012*b*; Whelan *et al.*, 2013; Althoff *et al.*, 2014; Lavigne *et al.*, 2014; Leadbeater and Homel, 2015; Déry *et al.*, 2017; Rice *et al.*, 2017). In addition, some studies have combined these items with the irritability symptom items from the depression, dysthymia, and/or mania sections of the interview (Dougherty *et al.*, 2013; Copeland *et al.*, 2014; Dougherty *et al.*, 2015; Pagliaccio *et al.*, 2018). Studies exploring DMDD most often make the diagnosis *ex post facto* based on endorsement of irritability items from both the ODD and major depression sections (e.g. Dougherty *et al.*, 2016). Finally, one study considered irritability items assessed from a depression-focused measure only (Kouros *et al.*, 2016).

The context in which a symptom is assessed (e.g. in order to make a particular diagnosis) and differences in the way items are worded could impact how the items are rated and, ultimately, what phenomena are captured. This is exemplified by one of the earlier longitudinal studies by Leibenluft's group, which distinguished between irritability assessed with items from the depression and mania section of the interview, which they considered 'episodic irritability,' and irritability assessed with items from the ODD section, labeled 'chronic irritability' (Leibenluft *et al.*, 2006). They found that stability correlations for each type of irritability over time were greater than the correlations between types over time. Further, episodic irritability was a significant predictor of anxiety and mania in late adolescence, and mania only in early adulthood, while chronic irritability predicted ODD and attention-deficit/hyperactive disorder (ADHD) in late adolescence, and major depression in early adulthood. Despite this finding, surprisingly few studies since have considered whether irritability is a multifactorial construct and how that may translate to different trajectories of psychopathology across development.

Another limitation of the literature concerns the assessment of externalizing outcomes. While some reviewers have argued that irritability is primarily related to later internalizing disorders (Vidal-Ribas *et al.*, 2016; Stringaris *et al.*, 2018), these conclusions may be premature given limitations in the assessment of externalizing psychopathology. That is, a substantial portion of studies exploring the longitudinal relationship between irritability and psychopathology have not explored links to externalizing disorders at follow-up (Burke, 2012; Lavigne *et al.*, 2014; Kouros *et al.*, 2016; Déry *et al.*, 2017; Rice *et al.*, 2017). Among studies that do, the focus is largely on ODD, ADHD, and conduct disorder (CD) in adolescents, with relatively less attention to externalizing outcomes in adults, such as substance use disorders (SUD) and antisocial personality disorder (ASPD). Only three prior studies have looked at SUD as an outcome (Pickles *et al.*, 2010), two of which also assessed ASPD (Stringaris *et al.*, 2009; Copeland *et al.*, 2014).

In general, most studies have focused on outcomes in late childhood through adolescence, with few studies following participants into adulthood. Among the studies that have included adults in their follow-up sample, some used irritability items from the ODD-section only (Stringaris and Goodman, 2009; Althoff *et al.*, 2014; Leadbeater and Homel, 2015), while others used measures that combined items from both the depression and ODD sections (Copeland *et al.*, 2014). Some studies did not account for continuity of psychopathology by adjusting for the presence of the outcome disorders at baseline (Pickles *et al.*,

2010; Althoff *et al.*, 2014; Copeland *et al.*, 2014). Finally, several studies employed mixed-age cohorts so that at follow-up, samples included both adolescents and adults (Stringaris and Goodman, 2009; Althoff *et al.*, 2014; Leadbeater and Homel, 2015). These inconsistencies suggest the need for further investigation of the adult outcomes of youth irritability.

A final noteworthy feature of this literature is that a vast majority of studies used parent report of irritability, either exclusively or combined with child or teacher report. Only four studies exploring prospective psychopathology outcomes have used irritability data from youth without considering parent report (Stringaris *et al.*, 2012*b*; Leadbeater and Homel, 2015; Evans *et al.*, in press), one of which also presented results for parent-reported irritability (Stringaris *et al.*, 2009). The predominant focus of prior studies on parent-report is appropriate given that many of these studies assessed irritability in childhood or early adolescence. However, for youths in mid-late adolescence who are more independent and perhaps less open with their parents about their internal struggles, there is particular benefit in understanding how self-report compares in this context.

The current study aims to augment this literature by examining the relationship between self-reported irritability in adolescence and subsequent psychopathology in adulthood, taking into account these issues. Using data from the Oregon Adolescent Depression Project (OADP; Lewinsohn *et al.*, 1993), we first explored dimensionality in the irritability symptom items from the major depression, dysthymia, mania, and ODD sections of a semi-structured diagnostic interview. Next, we examined the relationship between these dimensions of irritability in adolescence and psychopathology assessed at age 24 and 30.

Methods

Participants

The OADP is a large, prospective study of psychopathology in a community sample of three cohorts of adolescents recruited between 1987–1989 (Lewinsohn *et al.*, 1993, Rohde *et al.*, 2013). Participants were randomly selected from nine high schools representative of urban and rural districts in western Oregon. At the initial assessment (T1), the original sample of 1709 participants ranged in age from 14–19 ($M = 16.6$; S.D. = 1.2) and 9% of participants were non-White. All participants were invited for a second assessment (T2) approximately one year later ($M = 13.8$ months; S.D. = 2.3). Of the 1507 participants who returned for the T2 assessment, participants with a history of a mental disorder by T2 ($n = 644$) and a random sample of participants with no history of a mental disorder by T2 ($n = 457$) were invited to participate in a third assessment (T3) when they turned 24 years old. Of these, 85% ($n = 941$) participated in the T3 assessment. A fourth wave of assessments (T4) was conducted when participants turned 30 years old, and 87% ($n = 816$) of participants from T3 returned for this final assessment. T1 assessments were conducted in-person (either in the lab or at the school) while T3 and T4 assessments were conducted over the phone because many participants had moved out of the area. All study procedures were approved by the appropriate institutional review board and written informed consent was obtained from all participants (and their guardians, if applicable) at all assessments. More detailed description of study procedures can be found elsewhere (e.g. Lewinsohn *et al.*, 1993).

For this study, we focused on predicting outcomes at T3 and T4 from T1 to ensure that our outcome samples were exclusively

adult. Attrition from T2 to T3 and from T3 to T4 did not differ as a function of diagnostic status by T2 and T3, respectively. Among the 941 participants in the T3 wave, 57.3% were female and 10.9% were non-White. The distribution of race did not differ between T3 and T4 samples, however, a significantly larger proportion of women returned for the T4 assessment compared to those who did not (58.8 *v.* 47.2%), $\chi^2(1) = 5.99$, $p = 0.01$.

Measures

Irritability was assessed with a version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) that combined features of the epidemiologic version (K-SADS-E; Orvaschel *et al.*, 1982) with the present episode version (K-SADS-P). The interview was conducted with the participant only (i.e. parents were not interviewed about the teen). Three items from the depression, dysthymia, and mania sections that inquire about periods of feeling angry, grouchy, cranky or irritable, and three items from the ODD-section that probe whether the participant is often angry or resentful, often mad or easily annoyed, and often loses temper, were extracted to assess irritability. All six items were rated based on the worst period in their lifetime, so that these items essentially captured whether a participant had experienced a significant period of irritability at any point in their lives up to the time of assessment. We chose to use the symptom items rated based on the worst period instead of the current episode ratings, because current episode ratings for the irritability items from the ODD section were not available. The irritability symptom items were coded as present or absent. All six items were assessed in all participants (i.e. they were not affected by skip-outs).

Axis I diagnoses were assessed solely with the K-SADS at T1 and jointly with the Longitudinal Interval Follow-up Evaluation (LIFE; Keller *et al.*, 1987) at T3. At T4, diagnoses were jointly assessed with the LIFE and the Structured Clinical Interview for DSM-IV (SCID; First *et al.*, 1997). T1 diagnoses were based on the DSM-III-R (American Psychiatric Association, 1987) criteria, while T3 and T4 diagnoses were based on DSM-IV criteria (American Psychiatric Association, 2000). We included diagnoses if they were rated as threshold or not otherwise specified (NOS) and if they occurred at any point in the lifetime up to the time of assessment for T1. Outcome variables counted psychopathology reported since the last assessment (i.e. since T2 for T3 outcomes and since T3 for T4 outcomes).

We combined disorders into diagnostic groups coded as any disorder present/absent because of the relatively low rates of many individual disorders. The diagnostic groups of interest for the T3 and T4 outcomes included depressive disorders, anxiety disorders, and SUD. For depression, we counted major depression, dysthymia, and depressive disorder NOS. Anxiety disorders included panic disorder with and without agoraphobia, specific and social phobia, separation anxiety, generalized anxiety disorder, and anxiety disorder NOS. SUD included abuse or dependence of alcohol or any illicit (at the time) drug. Finally, we created a disruptive behavior disorder (DBD) category, which included ODD, ADHD, and CD. DBDs were not explored as outcomes at T3 and T4 but were included as T1 covariates in some analyses. Diagnostic group variables were coded as any disorder present or absent.

Personality Disorder Symptoms were assessed with the International Personality Disorder Evaluation (IPDE) at T3 and T4. The IPDE is a semi-structured interview of personality disorder symptoms based on DSM-IV criteria. ASPD and borderline

personality disorder (BPD) were the only personality disorders assessed because they are the most widely researched. Symptoms were summed to create dimensional measures of ASPD and BPD because of the relatively low base rates of these disorders in our sample. We excluded the symptoms referring to behavior occurring before age 15 from the ASPD symptom total because these could have occurred prior to the T1 assessment and thus would not reflect adult outcomes. We also excluded the item probing 'lack of remorse' from both T3 and T4 because it was not assessed at T4 (Table 1).

Data analysis

All analyses were weighted to account for overselection of participants reporting a mental disorder by T2 in the follow-up samples. Analyses were conducted in Mplus version 8.1 (Muthén and Muthén, 1998–2010).

First, an exploratory factor analysis (EFA) using robust weighted least squares (WLSMV) estimation was conducted to evaluate dimensionality of the irritability symptom items assessed at T1. WLSMV estimation is recommended for categorical indicators (Rhemtulla *et al.*, 2012). Factor solutions with eigenvalues greater than one were extracted and oblique geomin rotation was applied to allow factors to correlate. Although parallel analysis is generally recommended to determine the number of factors to retain (Hayton *et al.*, 2004), Muthén and Muthén (2013–2017) have found it to be inaccurate with categorical indicators. Model selection was based on χ^2 difference testing and evaluation of fit indices in accordance with Muthén and Muthén's recommendation for WLSMV. Irritability symptom dimensions to be used in later analyses were created based on the sum of irritability items loading on each factor. Results of the EFA conducted on the sample of participants who came for the T3 follow-up were reported, as the T4 follow-up sample was completely nested within the T3 sample and results did not differ substantively between the two samples.

Next, the irritability symptom dimensions were entered into binary logistic regression models predicting any SUD, depressive, or anxiety disorder at later assessments. Separate models predicted T3 and T4 outcomes in order to explore whether outcomes differed as a function of age at follow-up. The personality disorder symptom variables followed a count-like distribution (i.e. positive integer values concentrated at or near zero) with over-dispersion (i.e. conditional variance > conditional mean), so we used negative binomial regression to model BPD and ASPD symptoms at T3 and T4 (Coxe *et al.*, 2009). We initially ran each model controlling for gender, race, age at T1, and all T1 psychopathology diagnostic groups (i.e. depressive disorders, anxiety disorders, SUD, and DBD). Including many uninformative covariates can be problematic, (e.g. leading to numerically unstable estimates and large standard errors; Bursac *et al.*, 2008), so covariates with $p > 0.10$ were dropped from the final models. Results did not differ substantively between the initial and final models. The initial models containing all covariates can be found in the online Supplementary appendix.

Missing data

In our larger T3 sample, 84 (8.9%) participants were missing irritability symptom items, 39 (4.1%) participants were missing BPD symptom scores, and 37 (3.9%) of participants were missing ASPD symptom scores. We addressed missing data by using multiple imputation, creating 100 'complete' datasets. All predictor

Table 1. Sample characteristics

	T1 (*N* = 941)	T3 (*N* = 941)	T4 (*N* = 816)
Age, mean (s.d.) (age range)	16.63 (1.2) (14–19)	24.22 (0.6) (23–27)	29.70 (0.7) (28–33)
Gender: male (%)	–	402 (42.7)	336 (41.2)
Race			
White (%)	–	838 (89.1)	728 (89.2)
Black (%)	–	10 (1.1)	9 (1.1)
Asian (%)	–	24 (2.6)	22 (2.7)
Indian (%)	–	24 (2.6)	21 (2.6)
Hispanic (%)	–	28 (3.0)	22 (2.7)
Other (%)	–	17 (1.8)	14 (1.7)
Psychopathology			
Any depressive disorder (%)	477 (50.7)	375 (39.9)	265 (32.5)
Any anxiety disorder (%)	104 (11.1)	63 (6.7)	76 (9.3)
Any SUD (%)	104 (11.1)	318 (33.8)	198 (24.3)
Any DBD (%)	78 (8.3)		
ASPD symptoms, mean (s.d.)		1.37 (2.27)	1.03 (1.77)
BPD symptoms, mean (s.d.)		1.09 (2.46)	0.96 (1.95)
IRRM-OOD, mean (s.d.)	0.63 (0.80)		
IRR-ODD, mean (s.d.)	0.35 (0.75)		

T1 demographics are reported for the subset of the original sample who returned for at least the T3 follow-up, therefor the distribution of gender and race is identical across T1 and T3. s.d., standard deviation; SUD, substance use disorder; DBD, disruptive behavioral disorder; ASPD, antisocial personality disorder; BPD, borderline personality disorder; IRR-MOOD, the mood disorder irritability dimension; IRR-ODD, the oppositional defiant disorder irritability dimension.

and outcome variables were included in the imputation models to improve the accuracy of the missing data estimates. This process was completed separately for T3 and T4 outcomes. Regression analyses were conducted in each imputed dataset and results were pooled across analyses. As Mplus does not allow for multiple imputation with EFA, this analysis was conducted on pairwise complete data.

Results

Exploratory factor analysis of the irritability items

Factor loadings and model fit indices are reported in Table 2. One-factor and 2-factor solutions produced eigenvalues greater than one and a χ^2 difference test suggested that the 2-factor solution provided a significantly better fit, $\chi^2(5) = 168.822$ ($p < 0.001$). Additionally, approximate fit indices (Comparative Fit Index (CFI) = 0.997; Tucker Lewis Index (TLI) = 0.988; Root Mean Square Error of Approximation (RMSEA) = 0.039, 95% confidence interval = 0.006-0.071) surpassed commonly cited thresholds (Hooper *et al.*, 2008), suggesting that the 2-factor model fit the data well.

The irritability items from the major depression, dysthymia, and mania sections all loaded on the first factor, suggesting that this factor embodies a mood disorder dimension. Although the

Table 2. Exploratory factor analysis of irritability items

	1 factor model	2 factor model	
χ^2(df)	206.50 (9)*	9.76 (4)*	
CFI	0.693	0.997	
TLI	0.822	0.988	
RMSEA	0.153 (0.135, 0.171)	0.039 (0.006, 0.071)	
Factor loadings	1	1	2
Irritability/anger (MDD)	0.86*	0.95*	−0.01
Irritability/anger (dysthymia)	0.92*	0.94*	0.01
Irritability/anger (mania)	0.27*	0.27*	0.06
Loses temper (ODD)	0.70*	−0.05	0.86*
Angry/resentful (ODD)	0.75*	0.18*	0.74*
Mad/easily annoyed (ODD)	0.74*	0.00	0.88*

MDD, major depressive disorder; ODD, oppositional defiant disorder.
*Significant at $p < 0.05$; the 2-factor model fit significantly better than the 1 factor model $\chi^2(5) = 168.822$, $p < 0.001$.

mania item loading was relatively small (loading = 0.27), it did not cross-load on the second factor and thus we retained it in the mood disorder dimension. These three items were summed to create a mood disorder irritability dimension, which we will refer to as *IRR-MOOD* for efficiency of communication. All three irritability items from the ODD section loaded strongly on the second factor (loadings = 0.74-0.88). The angry or resentful item loaded significantly on the first factor as well, however, this loading was much smaller (0.18 *v*. 0.74), suggesting that it was a better indicator of the second factor. The three irritability items from the ODD section were summed to create an ODD irritability dimension, which we will refer to as *IRR-ODD*. The IRR-MOOD and IRR-ODD factors were moderately correlated ($r = 0.41$), further suggesting that these items are probing distinct, but related dimensions of irritability.

Regression analyses predicting psychopathology in early-mid adulthood

Results of the binary logistic regression analyses predicting depression, anxiety, and SUD at T3 and T4 are displayed in Table 3. The IRR-MOOD dimension predicted depression at T3 (but not T4) and anxiety at T4 (but not T3), but it did not predict SUD at either timepoint. In contrast, the IRR-ODD dimension predicted SUD at both timepoints, and anxiety at T3 (but not T4), but did not predict depression at either timepoint. Males had a significantly lower risk for depression and anxiety and a higher risk for SUD, compared to females. Younger age at T1 was associated with increased risk of SUD at T3 (but not T4). Additionally, participants with a SUD at T1 were at increased risk of depression at T3 and T4 and anxiety at T3 (but not T4).

Results of the negative binomial regression analyses predicting ASPD and BPD symptoms at T3 and T4 are reported in Table 4. Neither irritability dimension predicted ASPD symptoms at T3, while IRR-ODD (but not IRR-MOOD) predicted ASPD

Table 3. Logistic regression models predicting depression, anxiety, and substance use disorders

T3			T4		
	OR (95% CI)	p		OR (95% CI)	p
Any depressive disorder					
Sex	0.69 (0.51–0.93)	<0.005	Sex	0.40 (0.28–0.57)	<0.001
T1 depression	1.87 (1.36–2.57)	<0.001	T1 depression	1.76 (1.21–2.55)	<0.005
T1 SUD	1.52(0.99–2.33)	0.05	T1 anxiety	1.93 (1.33–2.80)	0.01
IRR-MOOD	1.27 (1.03–1.56)	0.03	IRR-MOOD	1.12 (0.88–1.43)	0.36
IRR-ODD	1.03 (0.85–1.25)	0.77	IRR-ODD	1.07 (0.85–1.34)	0.55
Any anxiety disorder					
Sex	0.42 (0.21–0.82)	0.01	Sex	0.47 (0.26–0.87)	0.02
T1 anxiety	3.31 (1.86–5.90)	<0.001	T1 anxiety	2.68 (1.44–5.00)	<0.005
T1 SUD	2.65 (1.36–5.18)	<0.005			
IRR-MOOD	1.15 (0.81–1.62)	0.44	IRR-MOOD	1.52 (1.09–2.11)	0.01
IRR-ODD	1.35 (1.05–1.73)	0.02	IRR-ODD	0.88 (0.60–1.29)	0.52
Any substance use disorder					
Sex	2.03 (1.47–2.79)	<0.001	Sex	1.95 (1.36–2.79)	<0.001
T1 age	0.84 (0.74–0.95)	<0.005			
T1 depression	1.37 (0.97–1.94)	0.07			
T1 SUD	3.07 (1.94–4.87)	<0.001	T1 SUD	4.38 (2.59–7.38)	<0.001
IRR-MOOD	0.97 (0.78–1.21)	0.78	IRR-MOOD	1.09 (0.84–1.40)	0.52
IRR-ODD	1.40 (1.14–1.72)	0.001	IRR-ODD	1.37 (1.06–1.76)	0.02

OR, odds ratio; CI, confidence interval; IRR-MOOD, the mood disorder irritability dimension; IRR-ODD, the oppositional defiant disorder irritability dimension; SUD, substance use disorder.

symptoms at T4. Male sex and T1 DBD and SUD were relatively strong predictors of ASPD symptoms at both T3 and T4. Additionally, younger age at T1 assessment predicted greater severity of ASPD symptoms at T4. At T3, IRR-MOOD predicted BPD symptoms and IRR-ODD approached significance ($p = 0.06$). At T4, both IRR-ODD and IRR-MOOD predicted BPD symptoms. T1 SUD and DBD predicted BPD symptoms at T3 and T4, while T1 depression predicted BPD symptoms at T3 only. Examination of the variance inflation factor revealed that multicollinearity was not an issue for any of the models.

Discussion

To our knowledge, this is the first paper to explore dimensionality in irritability symptoms assessed across multiple sections of a diagnostic interview using a data-driven approach. The results of our EFA are in line with Leibenluft et al. (2006) theory-driven distinction between irritability symptoms from the mood disorder section and irritability symptoms from the ODD section, which they labeled 'episodic' and 'chronic' irritability, respectively. We referred to these dimensions as the IRR-MOOD and the IRR-ODD so as to remain neutral about the meaning of the latent factors.

As reviewed above, most investigators have concluded that irritable youth go on to develop internalizing psychopathology, particularly depression. The results of our prospective analyses exploring the association between irritability assessed in adolescence and psychopathology assessed in early-mid adulthood yielded more nuanced findings.

First, the IRR-ODD dimension did not predict depression at either follow-up. This is particularly surprising given that a number of studies using an ODD-only assessment of irritability have found this relationship to be significant (e.g. Stringaris et al., 2009). Several factors likely contributed to this discrepancy. First, our study accounted for covariance between the two irritability dimensions. We observed that the IRR-MOOD dimension predicted depression in early adulthood over and above IRR-ODD, suggesting that significant associations in previous studies using the ODD items alone or when aggregated with the mood disorder items could have been driven, in part, by shared variance between irritability symptom items from both sections of the interview.

Additionally, few previous studies have assessed psychopathology in adulthood, and most of those that have focused on the period of late adolescence through early adulthood. The fact that neither irritability dimension predicted depression at T4 (age 30) could indicate that irritability does not predict increased risk for depression after early adulthood, over and above prior depression. Alternatively, it could be that, specifically, *self-reported* irritability does not predict depression in adulthood, given that Stringaris et al. (2009) found that parent but not child-reported irritability in adolescence predicted depression assessed around age 33. Further study of adult outcomes of irritability is needed to clarify which factors are driving these discrepancies.

Regardless, there is evidence to suggest that this relationship is sensitive to developmental period. For example, a twin study parsing additive genetic effects of irritability and symptoms of depression/anxiety across late-childhood through early-adulthood found

Table 4. Negative binomial regression models predicting personality disorder symptoms

	T3				T4		
	B (s.e.)	β	p		B (s.e.)	β	p
Antisocial personality disorder symptoms							
Sex	0.80 (0.11)	0.73	<0.001	Sex	0.59 (0.13)	0.60	<0.001
				T1 age	−0.10 (0.05)	−0.25	0.07
T1 SUD	0.72 (0.14)	0.36	<0.001	T1 SUD	0.86 (0.17)	0.45	<0.001
T1 DBD	0.85 (0.20)	0.38	<0.001	T1 DBD	0.72 (0.21)	0.34	<0.001
IRR-MOOD	0.02 (0.07)	0.02	0.83	IRR-MOOD	0.01 (0.09)	0.01	0.96
IRR-ODD	0.11 (0.07)	0.14	0.13	IRR-ODD	0.23 (0.10)	0.31	0.02
Borderline personality disorder symptoms							
Race	0.13 (0.05)	0.21	0.01	Race	0.15	0.26	0.001
T1 depression	0.39 (0.15)	0.37	0.01				
T1 SUD	0.61 (0.18)	0.26	0.001	T1 SUD	0.76 (0.20)	0.38	0.01
T1 DBD	0.57 (0.23)	0.31	0.01	T1 DBD	0.86 (0.23)	0.38	<0.004
IRR-MOOD	0.23 (0.10)	0.33	0.02	IRR-MOOD	0.24 (0.09)	0.36	0.01
IRR-ODD	0.22 (0.11)	0.28	0.06	IRR-ODD	0.29 (0.10)	0.38	<0.005

s.e., standard error; IRR-MOOD, the mood disorder irritability dimension; IRR-ODD, the oppositional defiant disorder irritability dimension; SUD, substance use disorder; DBD, disruptive behavioral disorder; race is coded as a nominal variable with 5 levels.

that, although there was significant genetic overlap across development, this overlap was strongest at ages 13–14 and the relationship between irritability and depression/anxiety was bidirectional at some, but not all, ages (Savage *et al.*, 2015). Additionally, Leadbeater and Homel (2015) found that, although irritability was moderately stable across adolescence-early adulthood, irritability only predicted internalizing symptoms in the transition from ages 18–19 to ages 20–21. This suggests that the nature of the relationship between irritability and internalizing symptoms changes across development. Thus, irritability may play a central role in the development of internalizing pathology in childhood and adolescence, but this relationship may attenuate after early adulthood.

The results for SUD and ASPD also extend previous research. No prior study has reported that irritability significantly predicted SUD or ASPD in adulthood. In contrast, IRR-ODD significantly predicted ASPD at T4 and SUD at both follow-up waves in our study. These findings challenge the assumption that irritability has a specific association with later internalizing psychopathology. Rather, it seems that this relationship depends on developmental stage and how irritability is assessed. Notably, none of the prior studies exploring this relationship in purely adult samples appeared to use the ODD irritability symptom items alone, which may explain this discrepancy. However, there are very few studies exploring externalizing outcomes in adulthood at all. This is likely due to the fact that, by convention, fewer externalizing disorders are assessed in adult diagnostic interviews. ADHD and CD were only recently added to the adult SCID for DSM-5 as optional modules not included in the main interview.

Our finding that IRR-ODD did not predict T3 ASPD is consistent with previous work exploring conduct problems in late adolescence-early adulthood using ODD-only measures of irritability (Stringaris and Goodman, 2009, Althoff *et al.*, 2014, Leadbeater and Homel, 2015). That IRR-ODD predicts ASPD

symptoms in mid, but not early, adulthood may be explained by Moffitt's dual taxonomy of antisocial behavior (Moffitt, 1993). Observing that prevalence of antisocial behavior peaks during adolescence, Moffitt has argued for two subgroups of individuals engaging in antisocial behavior; a life-course persistent group who engages in antisocial behavior throughout the lifespan and is characterized by more severe externalizing problems and a less severe adolescence-limited group whose antisocial behavior emerges during adolescence and desists by adulthood. It could be the case that our T3 assessment of ASPD was capturing a mix of antisocial subgroups, whereas our T4 assessment was more exclusively represented by life-course persistent individuals more likely to have experienced IRR-ODD during adolescence.

The results of the analyses predicting anxiety disorders at T3 and T4 are more challenging to explain. Only the IRR-ODD dimension predicted anxiety disorders at T3, while only the IRR-MOOD dimension predicted anxiety disorders at T4. Post-hoc analyses revealed little overlap in participants with anxiety disorders at T3 *v.* T4, however, exploratory analyses did not reveal notable differences in the composition of specific anxiety disorders or the patterns of comorbidity. Future research should elucidate whether these are spurious findings or capture meaningful relationships.

Surprisingly, the literature has largely ignored BPD as a prospective outcome of irritability in youth. Affective instability, and particularly, proneness to hostility and anger are central features of BPD. Further, there is evidence that ODD symptoms in childhood predict BPD symptoms in adolescence (Stepp *et al.*, 2012), however, this study did not distinguish between irritable and defiant/hurtful symptoms of ODD. One study compared the relationship between ODD symptom dimensions and BPD (Burke and Stepp, 2012), and found that the oppositional behavior dimension, but not the negative affective dimension, predicted BPD at age 24. Notably, this study defined negative affectivity

using the spiteful symptom and two of the three irritability symptoms we have included, and they included the 'loses temper' symptom in the oppositional behavior dimension. In our study, both IRR-MOOD and IRR-ODD predicted severity of BPD symptoms at T3 and T4, although the relationship between IRR-ODD and BPD at T3 fell just short of the threshold for significance. Taken together, this suggests that irritability in youth, and perhaps both dimensions of irritability, may be precursors of BPD in adulthood. This is consistent with structural analyses indicating that BPD straddles both internalizing and externalizing dimensions of psychopathology (Eaton et al., 2011). Further investigation of the role of irritability in the development of BPD is warranted.

It is notable that, although the 2-factor model provided superior fit compared to the 1-factor model, all irritability items loaded significantly onto the single factor and, in the 2-factor model, the factors were moderately correlated. This suggests that there is moderate overlap in the two dimensions of irritability and therefore it is unsurprising that many researchers have combined them. It is possible that the stronger relationship among irritability items within than between diagnostic sections reflects shared method variance rather than differences in the underlying trait. For instance, interviewers may have a lower threshold for positively rating the irritability item from the depression sections if other depression symptoms are present. Additionally, the items from the same section are administered more closely together in time than items from different sections. Nevertheless, the two dimensions demonstrated different patterns of longitudinal correlates, both in our study and in Leibenluft et al. (2006), even when holding concurrent psychopathology constant. Further, Leibenluft and colleagues found that chronic and episodic irritability followed different trajectories across development, with episodic irritability increasing linearly from late-childhood through early adulthood and chronic irritability peaking in mid-adulthood. This suggests that shared method variance cannot entirely account for this distinction and that the two dimensions of irritability symptoms are likely capturing meaningfully distinct phenomena.

The question remains of how to interpret the two dimensions of irritability symptoms. Apart from context/method variance effects, one possibility is that the irritability items from the mood disorder section reflect subjective affective experiences such as grouchy, grumpy or angry mood, while the items from the ODD section capture behavioral expressions of irritability, such as temper tantrums. Carlson and Klein (2018) note that the affective and behavioral aspects of irritability can have differing impact on functioning and different clinical implications. Another, not incompatible, possibility is that the probes in the mood disorder sections inquire about specific periods of time (i.e. 'periods of feeling angry, grouchy, cranky, or irritable'), suggesting episodicity, whereas the questions in the ODD section stipulate that the symptoms occur 'often,' implying some degree of chronicity. Hence, the two sets of items may reflect differences in course, as suggested by Leibenluft et al. (2006). Most likely, the two dimensions identified in our data reflect a combination of differences in affective v. behavioral expression of irritability, course, and method factors. We echo the call from Stringaris et al. (2012a) and others for more specific measures of irritability that will be able to provide a clearer characterization of the components (e.g. behavioral v. affective, episodic v. chronic) of this important clinical phenomenon, independent of a diagnostic context.

Limitations

The many strengths of this study should be considered in light of its limitations. First, because we used a community sample, rates of many individual disorders were relatively low and could not be assessed as independent outcomes. Second, although T1 assessments were conducted in adolescence, the diagnostic interviews covered the lifespan up to the time of assessment. As such, we were unable to distinguish between participants who experienced irritability at earlier (e.g. early childhood) v. later (e.g. mid-adolescence) developmental periods. Additionally, we were unable to compare the timing of mood symptoms relative to ODD symptoms, however, symptoms within a diagnostic section were rated based on the same period. Third, as is the case of all studies that employ retrospective reports, there is risk of recall bias. Fourth, we were limited to using existing diagnostic interviews that were designed to assess irritability as a symptom of a larger syndrome. This is a common approach, however, future research would benefit from using more specific measures that assess irritability in a transdiagnostic manner. Fifth, although irritability is a feature of generalized anxiety disorder, the irritability item from this section could not be included in the T1 irritability composites because it is not part of the screener and thus not assessed unless the cardinal symptoms are endorsed. Finally, we were limited to interviews with adolescents. As parent-report also has limitations, it would be ideal to obtain both parent and adolescent report for comparison.

Conclusions

Despite a large literature exploring longitudinal outcomes of irritable youth, few of these studies have considered the multidimensionality of irritability. The results of the present study suggest that the irritability symptom items in a widely used diagnostic interview represent distinct dimensions that are associated with different psychopathological outcomes in adulthood. These findings challenge the assertion that irritable youth grow specifically into internalizing adults; rather, different facets of irritability appear to predict internalizing and externalizing outcomes. Exploration of the irritability dimensions in other samples is needed to confirm these, as irritability is clearly a marker of persisting impairment that warrants more nuanced assessment that takes its multidimensionality into account.

References

Althoff RR, Kuny-Slock AV, Verhulst FC, Hudziak JJ and van der Ende J (2014) Classes of oppositional-defiant behavior: concurrent and predictive validity. *Journal of Child Psychology and Psychiatry* **55**, 1162–1171.

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders*, 3rd Edn. rev. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Press.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-5*®). Washington, DC: American Psychiatric Publishing.

Blader JC and Carlson GA (2007) Increased rates of bipolar disorder diagnoses among US child, adolescent, and adult inpatients, 1996–2004. *Biological Psychiatry* **62**, 107–114.

Burke J (2012) An affective dimension within oppositional defiant disorder symptoms among boys: personality and psychopathology outcomes into early adulthood. *Journal of Child Psychology and Psychiatry* **53**, 1176–1183.

Burke J and Loeber R (2010) Oppositional defiant disorder and the explanation of the comorbidity between behavioral disorders and depression. *Clinical Psychology: Science and Practice* **17**, 319–326.

Burke JD and Stepp SD (2012) Adolescent disruptive behavior and borderline personality disorder symptoms in young adult men. *Journal of Abnormal Child Psychology* **40**, 35–44.

Bursac Z, Gauss CH, Williams DK and Hosmer DW (2008) Purposeful selection of variables in logistic regression. *Source code for Biology and Medicine* **3**, 17.

Carlson GA, Klein DN (2018) Commentary: frying pan to fire? Commentary on Stringaris et al. (2018). *Journal of Child Psychology and Psychiatry* **59**, 740–743.

Copeland WE, Shanahan L, Egger H, Angold A and Costello EJ (2014) Adult diagnostic and functional outcomes of DSM-5 disruptive mood dysregulation disorder. *American Journal of Psychiatry* **171**, 668–674.

Copeland WE, Brotman MA and Costello EJ (2015) Normative irritability in youth: developmental findings from the Great Smoky Mountains Study. *Journal of the American Academy of Child & Adolescent Psychiatry* **54**, 635–642.

Coxe S, West SG and Aiken LS (2009) The analysis of count data: a gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment* **91**, 121–136.

Déry M, Lapalme M, Jagiellowicz J, Poirier M, Temcheff C and Toupin J (2017) Predicting depression and anxiety from oppositional defiant disorder symptoms in elementary school-age girls and boys with conduct problems. *Child Psychiatry & Human Development* **48**, 53–62.

Dougherty LR, Smith VC, Bufferd SJ, Stringaris A, Leibenluft E, Carlson GA and Klein DN (2013) Preschool irritability: longitudinal associations with psychiatric disorders at age 6 and parental psychopathology. *Journal of American Academy of Child & Adolescent Psychiatry* **52**, 1304–1313.

Dougherty LR, Smith VC, Bufferd SJ, Kessel E, Carlson GA and Klein DN (2015) Preschool irritability predicts child psychopathology, functional impairment, and service use at age nine. *Journal of Child Psychology and Psychiatry* **56**, 999–1007.

Dougherty LR, Smith VC, Bufferd SJ, Kessel EM, Carlson GA and Klein DN (2016) Disruptive mood dysregulation disorder at the age of 6 years and clinical and functional outcomes 3 years later. *Psychological Medicine* **46**, 1103–1114.

Eaton NR, Krueger RF, Keyes KM, Skodol AE, Markon KE, Grant BF and Hasin DS (2011) Borderline personality disorder co-morbidity: relationship to the internalizing–externalizing structure of common mental disorders. *Psychological Medicine* **41**, 1041–1050.

Evans SC, Burke JD, Roberts MC, Fite PJ, Lochman JE, Francisco R and Reed GM (2017) Irritability in child and adolescent psychopathology: an integrative review for ICD-11. *Clinical Psychology Review* **53**, 29–45.

Evans SC, Blossom JB and Fite PJ (in press) Exploring longitudinal mechanisms of irritability in children: implications for cognitive-behavioral intervention. *Behavior Therapy*.

First MB, Spitzer RL, Gibbon M and Williams JB (1997) *User's guide for the structured clinical interview for DSM-IV axis I disorders SCID-I: clinician version*. American Psychiatric Pub.

Hayton JC, Allen DG and Scarpello V (2004) Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organizational Research Methods* **7**, 191–205.

Hooper D, Coughlan J and Mullen M (2008) Structural equation modeling: guidelines for determining model fit. *Electronic Journal of Business Research Methods* **6**, 53–60.

Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P and Andreasen NC (1987) The longitudinal interval follow-up evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

Kouros CD, Morris MC and Garber J (2016) Within-person changes in individual symptoms of depression predict subsequent depressive episodes in adolescents: a prospective study. *Journal of Abnormal Child Psychology* **44**, 483–494.

Lavigne JV, Gouze KR, Bryant FB and Hopkins J (2014) Dimensions of oppositional defiant disorder in young children: heterotypic continuity with anxiety and depression. *Journal of Abnormal Child Psychology* **42**, 937–951.

Leadbeater BJ and Homel J (2015) Irritable and defiant sub-dimensions of ODD: their stability and prediction of internalizing symptoms and conduct problems from adolescence to young adulthood. *Journal of Abnormal Child Psychology* **43**, 407–421.

Leibenluft E and Stoddard J (2013) The developmental psychopathology of irritability. *Development and Psychopathology* **25**, 1473–1487.

Leibenluft E, Cohen P, Gorrindo T, Brook JS and Pine DS (2006) Chronic versus episodic irritability in youth: a community-based, longitudinal study of clinical and diagnostic associations. *Journal of Child and Adolescent Psychopharmacology* **16**, 456–466.

Lewinsohn PM, Hops H, Roberts RE, Seeley JR and Andrews JA (1993) Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III–R disorders in high school students. *Journal of Abnormal Psychology* **102**, 133–144.

Moffitt TE (1993) Adolescent limited and life course persistent antisocial behavior: A developmental taxonomy. *Psychological Review* **100**, 674–701.

Muthén LK and Muthén BO (1998–2010) *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.

Muthén BO and Muthén LK (2013–2017) Parallel analysis for categorical data. Available at http://www.statmodel.com/discussion/messages/8/11966.html?1504133952 (Accessed 30 March 2019).

Orvaschel H, Puig-Antich J, Chambers WJ, Tabrizi MA and Johnson R (1982) Retrospective assessment of prepubertal major depression with the Kiddie-SADS-E. *Journal of the American Academy of Child Psychiatry* **21**, 392–397.

Pagliaccio D, Pine DS, Barch DM, Luby JL and Leibenluft E (2018) Irritability trajectories, cortical thickness, and clinical outcomes in a sample enriched for preschool depression. *Journal of the American Academy of Child Adolescent Psychiatry* **57**, 336–342, e6.

Pickles A, Aglan A, Collishaw S, Messer J, Rutter M and Maughan B (2010) Predictors of suicidality across the life span: the Isle of Wight study. *Psychological Medicine* **40**, 1453–1466.

Rhemtulla M, Brosseau-Liard PÉ and Savalei V (2012) When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods* **17**, 354–373.

Rice F, Sellers R, Hammerton G, Eyre O, Bevan-Jones R, Thapar AK, Collishaw S, Harold GT and Thapar A (2017) Antecedents of new-onset major depressive disorder in children and adolescents at high familial risk. *JAMA Psychiatry* **74**, 153–160.

Rohde P, Lewinsohn PM, Klein DN, Seeley JR and Gau JM (2013) Key characteristics of major depressive disorder occurring in childhood, adolescence, emerging adulthood, and adulthood. *Clinical Psychological Science* **1**, 41–53.

Rowe R, Costello EJ, Angold A, Copeland WE and Maughan B (2010) Developmental pathways in oppositional defiant disorder and conduct disorder. *Journal of Abnormal Psychology* **119**, 726.

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

Stepp SD, Burke JD, Hipwell AE and Loeber R (2012) Trajectories of attention deficit hyperactivity disorder and oppositional defiant disorder symptoms as precursors of borderline personality disorder symptoms in adolescent girls. *Journal of Abnormal Child Psychology* **40**, 7–20.

Stringaris A and Goodman R (2009) Longitudinal outcome of youth oppositionality: irritable, headstrong, and hurtful behaviors have distinctive predictions. *Journal of the American Academy of Child Adolescent Psychiatry* **48**, 404–412.

Stringaris A, Cohen P, Pine DS and Leibenluft E (2009) Adult outcomes of youth irritability: a 20-year prospective community-based study. *American Journal of Psychiatry* **166**, 1048–1054.

Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E and Brotman MA (2012*a*) The affective reactivity index: a concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry* **53**, 1109–1117.

Stringaris A, Zavos H, Leibenluft E, Maughan B and Eley TC (2012*b*) Adolescent irritability: phenotypic associations and genetic links with depressed mood. *American Journal of Psychiatry* **169**, 47–54.

Stringaris A, Vidal-Ribas P, Brotman MA and Leibenluft E (2018) Practitioner review: definition, recognition, and treatment challenges of irritability in young people. *Journal of Child Psychology and Psychiatry* **59**, 721–739.

Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E and Stringaris A (2016) The status of irritability in psychiatry: a conceptual and quantitative review. *Journal of the American Academy of Child Adolescent Psychiatry* **55**, 556–570.

Wakschlag LS, Estabrook R, Petitclerc A, Henry D, Burns JL, Perlman SB, Voss JL, Pine DS, Leibenluft E and Briggs-Gowan ML (2015) Clinical implications of a dimensional approach: the normal: abnormal spectrum of early irritability. *Journal of the American Academy of Child Adolescent Psychiatry* **54**, 626–634.

Whelan YM, Stringaris A, Maughan B and Barker ED (2013) Developmental continuity of oppositional defiant disorder subdimensions at ages 8, 10, and 13 years and their distinct psychiatric outcomes at age 16 years. *Journal of the American Academy of Child Adolescent Psychiatry* **52**, 961–969.