Journal of Psychopathology and Behavioral Assessment (2022) 44:679–697
https://doi.org/10.1007/s10862-022-09974-8

Irritability and Emotional Impulsivity as Core Feature of ADHD and ODD in Children

Michaela Junghänel[1] · Ann-Kathrin Thöne[1] · Claudia Ginsberg[1] · Anja Görtz-Dorten[1,2] · Franziska Frenk[1] · Kristina Mücke[1] · Anne-Katrin Treier[1] · Sara Zaplana Labarga[1] · Tobias Banaschewski[3] · Sabina Millenet[3] · Jörg M. Fegert[4] · Dorothee Bernheim[4] · Charlotte Hanisch[5] · Michael Kölch[4,6] · Anne Schüller[7] · Ulrike Ravens-Sieberer[8] · Anne Kaman[8] · Veit Roessner[9] · Julian Hinz[9] · Manfred Döpfner[1,2]

Accepted: 13 April 2022 / Published online: 11 May 2022

Abstract

The categorical approach of diagnosing mental disorders entails the problem of frequently occurring comorbidities, suggesting a more parsimonious structure of psychopathology. In this study, we therefore aim to assess how affective dysregulation (AD) is associated with attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD) in children. To assess AD in children aged 8–12 years ($n = 391$), we employed the parent version of a newly constructed parent rating scale. Following item reduction, we conducted exploratory and confirmatory factor analyses to establish a factorial structure of AD. One core dimension was identified, comprising irritability and emotional impulsivity, and two smaller dimensions, comprising positive emotionality and exuberance. Subsequently, we examined five different latent factor models – a unidimensional model, a first-order correlated factor model, a second-order correlated factor model, a traditional bifactor model, and a bifactor S-1 model, in which the first-order factor AD-Irritability/Emotional Impulsivity (II) was modeled as the general reference factor. A bifactor S-1 model with the *a priori* defined general reference domain AD-II provided the best fit to our data and was straightforward to interpret. This model showed excellent model fit and no anomalous factor loadings. This still held true, when comparing it to bifactor S-1 models with ADHD/ODD-related reference factors. Differential correlations with emotion regulation skills and the established Parent Proxy Anger Scale validate the interpretation of the different dimensions. Our results suggest that irritability/emotional impulsivity might be a common core feature of ADHD and ODD.

Keywords ADHD · ODD · Affective dysregulation · Irritability · Emotional impulsivity · Bifactor models

✉ Michaela Junghänel
Michaela.junghaenel@uk-koeln.de

[1] School of Child and Adolescent Cognitive Behavior Therapy (AKiP), Faculty of Medicine, University Hospital Cologne, University of Cologne, Pohligstraße 9, 50969 Cologne, Germany

[2] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Medical Faculty, University of Cologne, Cologne, Germany

[3] Department of Child and Adolescent Psychiatry and Psychotherapy, Medical Faculty Mannheim, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany

[4] Department of Child and Adolescent Psychiatry/ Psychotherapy, University of Ulm, Ulm, Germany

[5] Department of Special Education, Faculty of Human Sciences, University of Cologne, Cologne, Germany

[6] Department of Child and Adolescent Psychiatry, Neurology, Psychosomatics, and Psychotherapy, University Medical Center Rostock, Rostock, Germany

[7] Department of Child and Adolescent Psychiatry and Psychotherapy, Brandenburg Medical School Theodor Fontane, Neuruppin, Germany

[8] Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Research Uni Child Public Health, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[9] Department of Child and Adolescent Psychiatry and Psychotherapy, TU Dresden, Dresden, Germany

The classification of psychiatric symptoms into categorical mental disorders, as is currently the case in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association [APA], 2013) and the 11th editions of the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2020), is useful for many reasons. It is useful for research, as the nosology makes it possible to integrate empirical findings, allows for communication and comparison of research findings, and guides further empirical studies. Ultimately, moreover, this research guides treatment. The categorical approach of classifying mental disorders brings the advantages of identifying risk and protective factors, of enabling prognoses, and of deciding which form of treatment has the best chance of success for a particular disorder (Caspi et al., 2014; Malhi & Bell, 2019).

However, one of the major challenges inherent in the categorical approach is the observation that comorbidities are the rule rather than the exception. Newman et al., (1998) found that comorbidities roughly conform to the rule of 50%, describing that half of individuals meeting diagnostic criteria for one disorder also meet criteria for a second disorder at the same time; and of these, 50% meet the criteria for a third mental disorder, and so on. Approaches attempting to explain this non-negligible number of comorbidities have focused either on the underlying etiological mechanisms in a bottom-up fashion (e.g. Research Domain Criteria [RDoC] by the National Institute of Mental Health; Cuthbert & Insel 2013; Insel et al., 2010) or on the observed pattern of covariation among psychopathological symptoms and traits in a top-down fashion (e.g. Hierarchical Taxonomy of Psychopathology [HiTOP]; Kotov et al., 2017). Though the RDoC and HiTOP differ in their approach to the reorganization of psychopathological symptom complexes, they pursue the same goal of moving away from diagnostic categories (Michelini et al., 2021). The neuroscientific RDoC approach aims to elucidate biobehavioral systems underlying a range of mental disorders (Cuthbert & Insel, 2013; Insel et al., 2010; Michelini et al., 2021). The approach is based on the assumption that mental disorders are, in fact, disorders of brain circuits, whose (dys-) functions can be assessed with neuroscientific tools, which will ultimately yield biosignatures that improve the understanding of the associations between symptom complexes (Insel et al., 2010). The extensively researched dimensional framework HiTOP (Kotov et al., 2017) aims to incorporate broader dimensions, potentially explaining comorbidity, as well as specific dimensions, accommodating heterogeneity *within* a disorder as well as symptom overlap *between* disorders (Kotov et al., 2017; Michelini et al., 2021). As both approaches come with some disadvantages (RDoC: large number of symptoms requiring clinical attention are missing (Michelini et al., 2021); HiTOP: underlying biological mechanisms are not considered), Michelini et al. (2020) proposed a promising RDoC-HiTOP interface, in which psychometrically robust clinical targets are suggested by HiTOP and can then be examined in terms of potential biological underpinnings.

While a more parsimonious, dimensional approach of classifying mental disorders comes with many advantages, it is not without criticism (Carragher et al., 2015; Ruggero et al., 2019; Zimmerman, 2021). As the ultimate goal of psychopathological classification systems is the selection and application of the ideal treatment, a restructuring of the classification system would have to improve precisely this process. The necessary relearning and retraining would require a significant amount of time and money (Zimmermann 2021, Carragher 2014) and it remains to be seen whether this process is worth the increased resources, as patients may show a considerable response to non-specific treatment aspects, independent of their diagnosis (Zimmermann, 2021). Related to this, due to time constraints, it might be difficult for acute settings to employ a fully dimensional approach (Ruggero et al., 2019). While an accurate classification system is undoubtedly of interest, and the relearning process would only be a temporary issue, the impact it would have on clinical outcomes will have to be evaluated.

In this article, we focus on affective dysregulation (AD) and how it might potentially explain the frequently observed comorbidities between attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). AD is a criterion for many diagnoses in children in the DSM-5 (APA, 2013) and the ICD-11 (WHO, 2020), and is therefore often characterized as a transdiagnostic dimension (Evans et al., 2017). It is generally understood as entailing an affective component (anger) and a behavioral component (aggression; Leibenluft & Stoddard 2013). Precise definitions differ, and range from rather narrow to very broad conceptualizations, leading to reported prevalence rates between 0.8% and 6.6% (Brotman et al., 2006; Copeland et al., 2013; Holtmann et al., 2008). Though the terms irritability, AD, and emotion dysregulation are sometimes used interchangeably when referring to the same or a highly similar construct (Evans et al., 2017; Leibenluft, 2011; Shaw et al., 2014), the mere definition of AD based on irritability is rather restricted, as emotion dysregulation and AD generally also include at least an impulsive component. In their influential review on irritability, Evans et al., (2017) summarized that irritability constitutes a diagnostic or at least an associated feature of a large number of psychological conditions, in particular but not limited to internalizing disorders such as depression and general anxiety disorder, and may be able to explain a good proportion of the frequently found comorbidities in children and adolescents. The authors also concluded that

particularly children with a combination of irritability with anger and temper outbursts, which corresponds to the definition of AD as entailing an affective as well as a behavioral component, show a pattern of correlates and outcomes differentiating them from other children with the same diagnosis. These findings led them to suggest that instead of an independent diagnosis for the combination of irritability and temper outbursts, a more parsimonious solution is needed (Evans et al., 2017). Compared to irritability, AD might therefore serve as an even more transdiagnostic concept, as irritability shows especially strong associations with disorders within the internalizing spectrum (Evans et al., 2017), whereas temper outbursts and impulsivity are, by definition, associated features of externalizing disorders. The addition of impulsivity might also be of particular interest given that in the majority of published associations, the p-factor is interpreted as a general tendency to react impulsively (Carver et al., 2017). AD-associated impulsivity differs from ADHD-associated impulsivity insofar as it contains an emotional component (e.g. *"often loses temper"*) that is not inherent in ADHD-associated impulsivity (e.g. *"talks a lot"*), which is so closely associated with hyperarousal that they are counted as *one* dimension in the DSM-5 (APA, 2013). The concept of emotional impulsivity has previously been suggested as an additional feature to the two established ADHD dimensions inattention and hyperactivity/impulsivity (Barkley & Fischer, 2010) and has been defined as "impatience, low frustration tolerance, hot-temperedness, quickness to anger, irritability and easily emotional excitability" (Barkley & Fischer, 2010, p.503), which closely corresponds to our definition of AD-associated impulsivity In this article, AD-associated emotional impulsivity and ADHD-associated impulsivity will therefore be treated as two separate constructs, in line with the general notion that impulsivity might be more of an "umbrella concept" (Berg et al., 2015, p.1129) referring to a large yet not definitively known number of dimensions (Berg et al., 2015).

AD symptomatology still poses a challenge for diagnostic classification systems (Evans et al., 2017). It is related to the DSM-5 (APA, 2013) diagnosis of disruptive mood dysregulation disorder (DMDD; Waltereit et al., 2019) as both concepts include chronic irritability as well as emotional impulsivity. As opposed to the broad, dimensional concept of AD symptomatology, DMDD is defined as a distinct categorical disorder. This has been strongly criticized (Evans et al., 2017; Lochman et al., 2015), as studies have failed to show a clear distinction from numerous other disorders, and mainly from ODD and ADHD (Evans et al., 2017). Based on recommendations by Lochman et al., (2015) and Evans et al., (2017), the ICD-11 (WHO, 2020) refrained from including AD symptomatology as a distinct categorical disorder and instead added a specifier of chronic irritability to the ODD diagnosis. The categorical conceptualization in the ICD-11 (WHO, 2020) and the DSM-5 (APA, 2013) is therefore countered by a dimensional view insofar as AD symptomatology is seen as a feature that is present in most individuals to varying degrees as well as a transdiagnostic feature in several psychopathological symptom complexes (Brotman et al., 2017). The same most likely holds true for ODD and ADHD symptomatology, as has been shown by a number of studies applying latent factor analyses in clinical and community samples to examine the underlying structure of these two disorders Burns et al., 2020; Junghänel et al., 2020; Thöne et al., 2021).

Previous studies assessing the comorbidities between ADHD and ODD have frequently applied bifactor or higher-order factor models to specify a common overarching factor Burns et al., 2020; Junghänel et al., 2020; Thöne et al., 2021; Waldman et al., 2021). Of the two ODD dimensions, particularly the dimension of defiant behavior has often being linked to ADHD (Evans et al., 2017; Stringaris & Goodman, 2009; Waldman et al., 2021). Emotion regulation, a concept that is inherent to all of the aforementioned symptom complexes, might explain the strong associations that are frequently found on this level (Ambrosini et al., 2013; Shaw et al., 2014). Similarly, in their trait-impulsivity theory, Beauchaine et al., (2010) suggested, that impulsivity combined with emotion dysregulation might be predecessors of the comorbid expression of ADHD and ODD.

Due to its particularly close associations with ADHD and ODD, AD may help to explain the frequently found comorbidities between these two disorders (Willcutt, 2012). To date, DMDD is the best researched conceptualization of chronic irritability and resembles the broad AD conceptualization to some extent, as it includes irritability as well as emotional impulsivity. For DMDD, strong associations with ODD and ADHD have been reported (Copeland et al., 2013; Evans et al., 2017; Leibenluft, 2011; Mayes et al., 2016; Mulraney et al., 2016). ADHD has been associated with the two main components of AD symptomatology separately as well. Barkley & Fischer (2010) identified emotional impulsivity as an important add-on concept to the ADHD diagnosis. In addition, (Karalunas et al., 2019) found ADHD symptomatology to be associated with irritability and identified a subgroup of children with ADHD and irritability that could not be reduced to the combination of ADHD and ODD. This finding led the authors to suggested a specifier of irritability to the ADHD diagnosis. Similarly, Eyre et al., (2017) showed that almost all children with an ADHD diagnosis displayed at least one symptom of irritability. Regarding AD and ODD symptomatology, the chronicity of irritability and the severity of temper tantrums supposedly differentiate AD from ODD. The close proximity between DMDD and ODD can also be seen in the DSM-5, where a

diagnostic hierarchy does not allow for an additional ODD diagnosis once the criteria for DMDD are fulfilled. The reason for the frequent diagnostic and clinical overlap between ODD and DMDD might lie in the current conceptualization of ODD. A number of independent research groups have demonstrated that ODD consists of at least two different dimensions – irritability and defiant/argumentative behavior, which lead to differential outcomes (Evans et al., 2017). The current conceptualization of ODD appears to confound these two frequently co-occurring but distinct dimensions (Runions et al., 2016). This combination of dimensions into a categorical diagnosis is a hallmark of the current classification systems and is contrasted by approaches attempting to explain comorbidities between diagnoses by reorganizing symptoms constituting DSM-5/ICD-11 diagnosis into dimensions and to model their associations with one another (e.g. HiTOP; Kotov et al., 2017). Further research is necessary to examine the exact composition of AD symptomatology and its association with dimensions of ODD and ADHD. As we were interested in the broad spectrum of AD and not in a specific diagnosis or conceptualization, for the purpose of the present study it seemed most appropriate to employ a broad definition of AD symptomatology, assessing all potentially associated features.

One method to address the research question of how (dimensions of) mental disorders are related to one another is latent factor analysis. Latent factor analysis can highlight core dimensions accounting for observed symptomatology and comorbidity between disorders (Eaton, 2015) and has been applied with increasing frequency within psychological research in recent years. Applying latent factor analysis also enables us to move away from diagnostic-level to symptom-level analyses, contributing to the development of potentially more valid and parsimonious nosologies (Eaton, 2015). In particular, higher-order factor models and bifactor models, which decompose true score variance and assign it to a general (g-) or a specific (s-) factor have provided useful insights into the latent factor structure of psychopathology (Eid et al., 2017). The variety of latent factor models come with advantages and disadvantages, provide differential information, and can therefore answer different kinds of research questions. In the following, we will briefly outline the kind of models, also employed in this study, that are frequently used to answer research questions related to the overall theme of examining associations between dimensions of mental disorders. First-order correlated factor models (CFO) are a good basis for higher-order models and yield initial insights into how the dimensions or disorders are related to one another (Eid, 2020). In our case, assessing AD, ODD, and ADHD dimensions, such models would allow us to examine to what degree these dimensions are correlated. If high correlations between dimensions are found, the CFO ends up being an incomplete model, as correlations between dimensions are left unmodeled. These correlations can then either be accounted for in a unidimensional model (i.e., assuming that an externalizing spectrum, which has been found frequently, though with different subfacets (Kotov et al., 2017; Krueger et al., 2005) captures the corresponding symptomatology better than the distinct diagnoses ADHD, ODD and AD) or a model that assumes some sort of overarching factor. The models with an overarching factor are either higher-order factor models, such as the HiTOP (Kotov et al., 2017), or bifactor models, which model an additional first-order factor instead of a higher-order factor, which is associated with the *items* from all dimensions. The idea behind bifactor models is that *"only with the general factor variance removed can we have a clear window into the remaining covariance patterns among the symptoms in our measure. Only with specific measures unconfounded by the general factor can we have a clear window into the etiological or prognostic associations"* (Hartman, 2021, p.72). However, despite the popularity of these models, it often remains unclear what this so-called g-factor really stands for (Heinrich et al., 2020). Additional problems include the proportionality constraint in the higher-order factors model, which describes the problem that, by definition, all items from one dimension show the same association ratio with the lower- and higher-order factors (Brunner et al., 2012; Gignac et al., 2016), whereas in bifactor models, associations between the s-factors as well as between the g- and the s-factors cannot be assessed as they are constrained to zero. In addition, bifactor models often yield weakly defined s-factors and interpretation difficulties of the factors and their relation to one another arise due to anomalous factor loadings, such as negative or non-significant factor loadings or variances (Burns et al., 2020; Eid et al., 2017). A relatively new version of a bifactor model – the bifactor S-1 model – offers a solution to the aforementioned problems. By modeling one of the s-factors as a general reference factor (i.e. there is no s-factor modeled for items of the reference factor), there are now "pure indicators of the general factor" (Markon, 2021, p.67). The choice of this reference factor should be theoretically derived or correspond to a domain of greater interest (Eid, 2020). The remaining s-factors are orthogonal to this general reference factor and are allowed to correlate with each other (Eid et al., 2017). The bifactor S-1 model therefore allows for a straightforward interpretation of the g- and s-factors as well as their relations to one another. The initial goal of bifactor models to identify a general overarching psychopathology factor has to be dismissed when applying this version of the bifactor model. Nevertheless, it appears that traditional bifactor models cannot reach this goal either, and frequently lead to

a misinterpretation of the g-factor as a general psychopathology factor (Heinrich et al., 2020).

Aim of Study

The current study aimed to establish a factorial structure of AD and to examine how AD relates to the defiant dimension of ODD and ADHD. As ODD has been shown to consist of a defiant and an irritable dimension and irritability is a main component of AD, we only assessed associations of AD with the defiant dimension of ODD (ODD-D). For ADHD, associations with inattention (ADHD-IN) and hyperactivity/impulsivity (ADHD-HI) were examined. In a first step we assessed item descriptive statistics and item redundancy in order to shorten the AD questionnaire, which in its original form was a combination of items from several existing questionnaires assessing AD. This was followed by exploratory (EFA) and confirmatory factor analyses (CFA) to establish the factorial structure of AD. In a next step, we examined five different latent factor models (Fig. 1): (a) a unidimensional model, corresponding to the idea of a general externalizing spectrum, in which all items from all dimensions loaded onto one common overarching factor (Uni), (b) a correlated factor model with the aforementioned correlated first-order factors (CFO) to assess the associations between dimensions, (c) a second-order factor model, in which these first-order factors loaded onto one second-order factor (SOF), (d) a bifactor model, in which the items from all dimensions loaded onto one general factor as well as onto *one* specific factor (BI), and (e) a bifactor S-1 model, in which we suggest the core AD-factor, which we presumed to be describing irritability/emotional impulsivity, as a candidate for the general reference factor (BI S-1). Importantly, at this step, any dimension of interest could have been chosen as the general reference factor. We chose the core AD dimension, as previous studies have frequently found measures of emotional impulsivity, a concept closely related to AD symptomatology, to be strongly associated with the p-factor and to predict the onset and progression of symptoms (Carver et al., 2017). In line with this, Beauchaine et al., (2010) suggested emotion dysregulation/impulsivity as a predecessor of the combined presentation of ADHD and ODD, contributing to the hypothesis of AD core symptomatology serving as an ideal candidate for the general reference factor of AD, ODD and ADHD. Subsequently, we examined the associations of the different dimension with emotion regulation skills, assessed using the German FRUST questionnaire as well as the Parent Proxy Anger Scale, which is part of the Patient-Reported Outcomes Measurement Information System (PROMIS; Irwin et al., 2012). In a final step, in order to validate our assumption of AD core symptomatology serving as an ideal reference factor, we tested competing hypotheses, i.e. models with ODD-D, ADHD-IN, or ADHD-HI as alternative general reference factors. Since AD symptomatology is currently conceptualized as a specifier to the ODD diagnosis in the ICD-11 (WHO, 2020), it seemed important to assess ODD-D as a general factor as well. As for ADHD, the finding of Karalunas et al., (2019) regarding a specifier of irritability to the ADHD diagnosis and the results of Barkley & Fischer (2010) of emotional impulsivity being an important add-on concept to the ADHD diagnoses, led us to the decision to also test the hypotheses of both ADHD dimensions serving as the general reference factor.

Hypotheses

1) Item reduction: As items were taken from several existing questionnaires assessing broad AD symptomatology in children, we expected the results of item descriptive statistics and item redundancy to allow for a shortening of the newly developed 38-item AD questionnaire before further validating it.
2) Factorial structure of AD: In line with the conceptualization of AD symptomatology as a specifier of irritability/anger suggested by Evans et al., (2017) and as DMDD (DSM-5, APA, 2013), we expected to identify an AD-core factor, defined by items describing irritability and emotional impulsivity. As items were taken

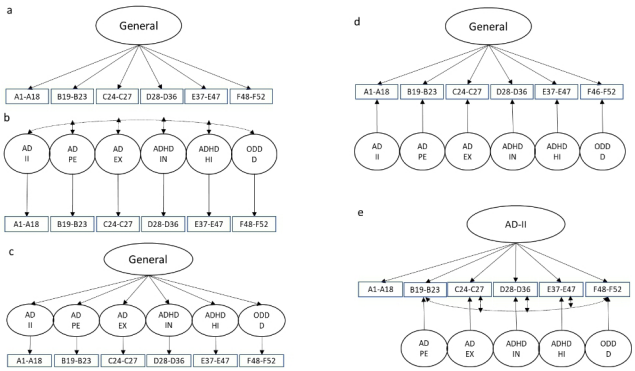

Fig. 1 Latent Factor Models of AD, ADHD and ODD Dimensions. (**Fig. 1a** Unidimensional model (Uni), **b** Factor model with six correlated first-order factors (CFO), **c** Factor model with six correlated first-order factors and one second-order factor (SOF), **d** Bifactor model (BI) **e** Bifactor S-1 model (BI S-1)
Note. In 1b and 1e, all first-order factors are allowed to correlate (indicated by dotted arrows). Item numbers are displayed in the boxes and residuals are not shown for clarity of presentation. AD = affective dysregulation, II = irritability/emotional impulsivity, PE = positive emotionality, EX = exuberance, ADHD = attention-deficit/hyperactivity disorder, IN = inattention, HI = hyperactivity/impulsivity, ODD-D = oppositional defiant disorder – only defiant dimension)

Table 1 Sample and Descriptive Statistics

Sample Statistics		
Total sample		*n* = 391
Age: mean (*SD*)		10.64 (1.33)
Male: *n* (%)		220 (56)
Group [*n* (%)]		
AD		244 (62)
NoAD		147 (38)
Diagnoses [*n* (%)]		
DMDD		41 (11)
ODD		93 (24)
ADHD		62 (16)
ADHD, combined type		19 (5)
ADHD, predominantly inattentive type		30 (8)
ADHD, predominantly hyperactive-impulsive type		13 (3)
CD		6 (2)
MD		5 (1)
Comorbid diagnoses [*n* (%)]		
AD + ODD		93 (24)
AD + ADHD		61 (16)
AD + ADHD, combined type		19 (5)
AD + ADHD, predominantly inattentive type		29 (7)
AD + ADHD, predominantly hyperactive-impulsive type		13 (3)
AD + CD		6 (2)
AD + MD		5 (1)
Dimensional Statistics		
Dimension (*n* = 386–390)	M (*SD*)	α
AD-irritability/emotional impulsivity	1.13 (0.55)	0.96
AD-positive emotionality	2.23 (0.49)	0.72
AD-exuberance	0.89 (046)	0.81
ADHD-inattention	0.93 (0.79)	0.94
ADHD-hyperactivity/impulsivity	0.56 (0.63)	0.92
ODD-defiant dimension	0.68 (0.64)	0.84
FRUST-adaptive emotion regulation strategies	2.00 (0.72)	0.89
FRUST-maladaptive emotion regulation strategies	1.61 (1.00)	0.78
PROMIS	1.40 (0.93)	0.91

Note. AD = affective dysregulation, ADHD = attention-deficit/hyperactivity disorder, DMDD = disruptive mood dysregulation disorder, ODD = oppositional defiant disorder, CD = conduct disorder, MD = major depressive episode, M = mean, *SD* = standard deviation, α = Cronbach's Alpha, FRUST = "Questionnaire on the regulation of unpleasant moods in children" (Fragebogen zur Regulation unangenehmer Stimmungen von Kindern), PROMIS = Patient-Reported Outcomes Measurement Information System - Parent Proxy Anger Scale

from several questionnaires, assessing a broad range of AD symptomatology in children, we expected to find smaller AD-factors in addition to that core factor.

3) Latent factor analysis of AD, ADHD and ODD symptomatology: We expect the extension of our theoretically derived bifactor S-1 model to AD symptomatology with AD-core symptoms as 'pure indicators' of the reference factor to fit the data best in terms of global model fit indices, the significance of factor loadings and general interpretability. This is consistent with previous studies showing the clear interpretability of the bifactor S-1 model when assessing the structure of ADHD and ODD (Burns et al., 2020; Junghänel et al., 2020; Thöne et al., 2021). On top of that, this approach adds to previous findings reporting high comorbidities between AD, ADHD and ODD symptomatology (Evans et al., 2017; Eyre et al., 2017; Mulraney et al., 2016) and a strong association of irritability with ADHD and ODD (Ambrosini et al., 2013; Shaw et al., 2014).
4) Divergent and convergent validity: We expected to find differential correlations of all AD, ADHD and ODD dimensions with emotion regulation strategies as well as the PROMIS Parent Proxy Anger Scale, thus supporting the external validation of our model.
5) Comparison of reference factors: Based on studies suggesting measures of emotional impulsivity, which are strongly related to AD core symptomatology, as an ideal candidate for our general reference factor (Beauchaine et al., 2010; Carver et al., 2017), we expected the bifactor S-1 model with the reference factor AD-II to provide a better fit to the data in terms of explained common variance of the reference factor, global model fit, and the pattern of factor loadings, compared to models with the reference factors ODD-D, ADHD-IN or ADHD-HI.

Methods

Participants and Procedure

Data collection took place within the ongoing ADOPT (Affective Dysregulation in Childhood—Optimizing Prevention and Treatment; Döpfner et al., 2019) study. The *ADOPT* project is a multicenter research project encompassing Seven study centers located in Germany and the six subprojects coordination, epidemiology, neurobiology, online, treatment and institution. It entails several measurement time points (Döpfner et al., 2019). The aim of this multicenter study was to optimize clinical diagnostics, prevention and treatment of AD and investigates an evidence-based, individualized treatment program based on behavioral interventions for children with AD. For the

present study, we analyzed baseline data (T1) of a sample screened in the community, collected between August 2018 and September 2019, which included 391 children aged 8–12 years ($M = 10.6$, $SD = 1.3$; 56.3% males). An initial screening, obtained through a parent screening questionnaire, categorized participants into an AD and a NoAD group. Children with AD symptoms in the top 10% of the sample were allocated to the AD group, children with AD symptoms in the bottom 10% of the sample were allocated to the NoAD group (for further details regarding the screening procedure see Otto et al., 2022). All families in the AD group were then offered further participation in the ADOPT study. A random sample was drawn from the NoAD group. Clinical child and parent interviews (Görtz-Dorten, Döpfner & Thöne, 2022) were conducted with participating families. AD/NoAD group assignment was confirmed through the clinical parent interview. Main inclusion criteria include the age of the child (8;0–12;11 at T1), the residence of the child (child lives with at least one natural or adoptive parent), and clinician-rated AD symptomatology of the child (based on the clinical parent interview) as well as the families' willingness and ability to participate in the study. As can be seen in Tables 1 and 244 (62.4%) of study participants were categorized into the AD group. Regarding additional diagnoses, 41 study participants (10.5%) fulfilled the DSM-5 criteria for DMDD, 93 (23.8%) for ODD, 62 (15.9%) for ADHD, six (1.5%) for CD and five (1.3%) for a major depressive episode (MD). In almost 100% of ODD, ADHD, MD and CD comorbid AD was present (Table 1). All diagnoses were based on clinical interviews (Görtz-Dorten, Döpfner & Thöne, 2022), which are part of the Diagnostic System of Mental Disorders in Children and Adolescents based on the ICD-10 and DSM-5 [*Diagnostik-System für psychische Störungen nach ICD-10 und DSM-5 für Kinder und Jugendliche – III*] (DISYPS–III; Döpfner & Görtz-Dorten (2017).

Measures

FBB-ADHS/FBB-SSV

The mothers or fathers of the 391 participants completed the German Symptom Checklist for Attention-Deficit/Hyperactivity Disorder (*Fremdbeurteilungsbogen für Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen;* FBB-ADHS) as well as the Symptom Checklist for Disruptive Behavior Disorders (*Fremdbeurteilungsbogen für Störungen des Sozialverhaltens*; FBB-SSV) from the DISYPS-III (Döpfner & Görtz-Dorten, 2017). Both rating scales are based on the symptom criteria of the DSM-5 and ICD-10. The FBB-ADHS assesses IN symptoms with nine items and HI symptoms with eleven items. All items are rated on a 4-point Likert scale (0 [not at all) – 3 [very much]). The FBB-SSV assesses ODD, CD and callous-unemotional symptoms. Following the questionnaire's instruction to assess CD and CU symptoms only for children aged eleven or older, we only used the items assessing ODD for the present study. To avoid symptom overlap, and as we were interested in how dimensions of psychopathology (similar to HiTOP; Kotov et al., 2017) might be able to explain comorbidities between current diagnostic categories such as ADHD and ODD, we only considered five of the items, which assess the defiant dimension of ODD (ODD-D), as the other three ODD symptoms of the FBB-SSV assessing the irritability dimension were already included in the assessment of AD. Both questionnaires have shown good psychometric properties in terms of reliability and validity of the scale scores (Döpfner et al., 2008; Erhart et al., 2008; Görtz-Dorten et al., 2014).

DADYS parent rating

The parent version of the DADYS [**D**iagnostikum für **A**ffektive **DYS**regulation (Diagnostic System for Affective Dysregulation)] is used to assess AD and comprises 38 items from several existing questionnaires assessing irritability/anger and affective dysregulation/emotional regulation in general. Items from the DADYS included in the present analysis were taken from the Emotion Regulation Checklist (Shields & Cicchetti, 1997), the FBB-SSV (Döpfner & Görtz-Dorten, 2017) and the Affective Reactivity Index (Stringaris et al., 2012). All items are rated on a 4-point Likert scale (0 [not at all) – 3 [very much]).

PROMIS parent Proxy anger scale

The Parent Proxy Anger Scale is part of the Patient-Reported Outcomes Measurement Information System (PROMIS; Irwin et al., 2012) and entails five items assessing rage and anger in children. Items are rated by the parent on a 5-point Likert scale, ranging from 0 (never) to 4 (almost always). The scale score of the Anger Scale has shown good psychometric properties (Varni et al., 2012).

FRUST

The "Questionnaire on the regulation of unpleasant moods in children" (Fragebogen zur Regulation unangenehmer Stimmungen von Kindern; FRUST) by (Görtz-Dorten et al., 2019, unpublished manuscript) is adapted from the FEEL-KJ questionnaire (Grob & Smolenski, 2005), which is frequently applied to assess emotion regulation skills in children and adolescents. As opposed to the FEEL-KJ, which assesses the regulation skills for fear, anger and

sadness separately with 30 item per emotion ("If I am scared, I…"/"If I am angry, I…"/"If I am sad, I…") the FRUST assesses the regulation of unpleasant emotions in general with 30 items all together ("If I feel bad, I…") We additionally designed and validated a parent version of the FRUST, which was used in this study. This version contained only 14 items, as we excluded items assessing primarily internal processes that are difficult to observe from the outside. The FRUST showed good psychometric qualities (Junghänel et al., in preparation). Items are rated on a 5-point Likert-scale ranging from 0 (almost never) to 4 (almost always). A two-factorial structure of the FRUST has been established, consisting of one adaptive emotion regulation factor (FRUST-A; 10 items) and one maladaptive emotion regulation factor (FRUST-M; 4 items) (Junghänel et al., in preparation). The internal consistency was good, with Cronbach's $\alpha = 0.89$ for FRUST-A and $\alpha = 0.78$ for FRUST-M (Junghänel et al., in preparation).

Statistical analyses

For descriptive analyses and calculations of internal consistency, we used SPSS version 26. All other analyses were conducted using Mplus version 8.4 (Muthén & Muthén, 2017). We used the weighted least square mean and variance adjusted (WLSMV) estimator (delta parameterization), which is suggested for modeling ordinal data and does not assume normally distributed variables (Li, 2016). Missing data were handled with the Mplus default strategy *pairwise deletion.* Due to increasing computational demands in analyses with five or more factors, the full information maximum likelihood (FIML) estimation has been found to be impracticable (Forero & Maydeu-Olivares, 2009). In addition, global model fit cannot be assessed in most cases with Mplus when fitting ordinal factor analysis models with FIML (Shi et al., 2020). In the case of a low number of missing values, high reliability of the scales, as well as the assumption that missing values are missing completely at random, which all held true in our study, pairwise deletion has been found to work well (Shi et al., 2020; Tsikriktsis, 2005). The amount of missing data per item was below 1% for all items. Covariance coverage was above 0.995 for all items.

For the SOF model, we applied the Schmid-Leiman transformation, which is a mathematical transformation of the standardized factor loadings that can be used to estimate the direct impact of the first-order and the higher-order factors on manifest item scores in higher-order factor models (Brunner et al., 2012; Gignac, 2016). For the impact of the higher-order factor, the standardized factor loading of each item was multiplied by the factor loading of the corresponding first-order factor on the second-order factor. To estimate the impact of the specific first-order factor, the factor loading of each item was multiplied by the standard deviation of the corresponding factor (for a detailed explanation see Brunner et al., 2012).

To evaluate model fit, we predominantly relied on the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residuals (SRMR). For model fit to be considered good, the CFI and TLI should be ≥ 0.95 (Hu & Bentler, 1999) and RMSEA and SRMR should be ≤ 0.05 (Browne & Cudeck, 1992; Hooper et al., 2008). For adequate model fit, RMSEA and SRMR should be ≤ 0.08 (Browne & Cudeck, 1992; Hooper et al., 2008). Differences of > 0.010 in CFI would indicate a significant difference between global model fit (Cheung & Rensvold, 2002).

There is a vast array of indices that can be used for the evaluation of dimensionality in bifactor models on the item level, the factor level, and the model level (for a detailed summary and explanation see Rodriguez et al., 2016). Frequently reported are the omega statistics. Omega (ω) describes the amount of reliable variance accounted for by the g-factor and all s-factors taken together, whereas omega hierarchical (ωH) describes the amount of reliable variance accounted for by the g-factor (Brunner et al., 2012; Reise, 2012). Omega hierarchical subscale (ωHS) is the equivalent to ωH for the individual s-factors. It has been recommended that ωH/ωHS should not be lower than 0.50 or ideally 0.75 in order to be interpreted reliably (Reise et al., 2013). Another important index supporting the correct evaluation of dimensionality is the explained common variance (ECV), which indicates the proportion of all common variance explained by that factor. Bifactor-relevant indices were calculated with the help of the Bifactor Indices Calculator by Dueber (2017).

Analytic Plan

Item reduction of the DADYS Questionnaire

For item reduction several criteria were considered: Skewness, kurtosis, usage of all response categories, distribution of responses (how often was an item answered with 0 or 1), item-item-correlations $r < .30$, item-item-correlations $r > .80$, redundancy of item content assessed by three clinical raters, and correspondence with DSM-5 criteria for DMDD. For the clinical assessment of item redundancy each of the three clinical raters created item pools, consisting of items that they regarded to assess the same content. Items were excluded for three reasons: 1) *Exclusion due to saliences in*

descriptive statistics. Items were excluded if they showed at least four of the following saliencies: (a) Skewness/kurtosis larger than one/two standard deviation(s) (counted as one and two saliences, respectively), (b) not all response categories were used, (c) salient distribution of responses (> 90% answered this item with 0 or 1), (d) the item correlated with more than 50% (= 17 items) of the other items with $r < .30$. 2) *Exclusion due to item redundancy*. Items were excluded if they correlated with at least one other item with $r > .80$ and the content was additionally rated as redundant by at least two of the three clinical raters. The selection regarding which of the redundant items was kept in the questionnaire was based on two aspects: If one of the redundant items was a DSM-5 criterion for DMDD, this item was selected to remain in the questionnaire. If none of the items rated as redundant were DSM-5 criteria for DMDD, the item with the highest number of item-item correlations $r > .80$ was kept in the questionnaire. 3) *Exclusion due to low associations with the other items*. Items were excluded if they correlated with more than 90% (= 34 items) of the other items with $r < .30$. If the to-be-excluded item was a DSM-5 criterion for DMDD, it was kept in the questionnaire. As our assessment of AD symptomatology is similar to, yet broader than DMDD, we aimed for our AD construct to fully include DMDD and therefore decided to keep the previously evaluated DSM-5 criteria for DMDD in our questionnaire whenever possible.

Exploratory and confirmatory factor analysis of the DADYS Questionnaire

For EFA and CFA, items were treated as ordinal and the WLSMV estimator was used. In a first step, the sample was divided randomly in two groups, each containing $n = 195$ children. We then applied EFA in one half of the sample with a principal axes factor analysis, as we did not expect for all variance to be explained by the extracted factors. We chose the oblique GEOMIN rotation, which allows for correlations between factors. EFA is advantageous in situations where no clear established structure for a construct exists, as cross-loadings are freely estimated, which can provide novel insight into the data structure. We then applied CFA in the other half of the sample in order to examine if the previously extracted factorial structure of AD can be confirmed. We additionally let Mplus calculate modification indices (MI) to suggest potential changes to our model, which were carefully checked and assessed with regard to content-based meaningfulness in every case.

Confirmatory factor analysis of AD, ADHD and ODD dimensions

In a next step, items from all AD-dimensions, ADHD-IN, ADHD-HI and ODD-D were tested in a unidimensional model (Uni), a first-order correlated factor model (CFO), a second-order correlated factor model (SOF), a traditional bifactor model (BI) and a bifactor S-1 model (BI S-1) to examine how AD symptomatology was associated with ADHD and ODD dimensions (Fig. 1).

Associations with External correlates

To externally validate our factorial structure, correlations of the dimensions of our optimally fitting model with the PROMIS Parents Proxy Anger Scale as well as with the FRUST-A and the FRUST-M were calculated.

Comparison of reference factors

In addition to the previously examined BI S-1 model, which had AD-II as reference factor, we additionally tested three bifactor S-1 models with ODD-D (BI S-1; ODD-D), ADHD-IN (BI S-1; ADHD-IN) and ADHD-HI (BI S-1; ADHD-HI) as reference factors.

Results

Descriptive information

Descriptive information can be found in Table 1. Internal consistencies for all scales, including the reduced ODD-D scale, were good ($\alpha \geq 0.72$).

Item reduction

The exclusion process with all its criteria can be found in Table S1. Seven items (10, 12, 16, 18, 19, 24 and 30) fulfilled the criteria for previously described salient descriptive statistics. We excluded all of these items with the exception of item 30 ("*prolonged physically aggressive tantrums*") as this item was included in the DSM-5 as a criterion for DMDD and therefore of interest to keep in the questionnaire. Four items (8, 26, 36, 37) were excluded due to item redundancy. Of the items that correlated highly with one another we kept item 29 as this was included in the DSM-5 as a criterion for DMDD. One item (23) was excluded as it showed low item-item correlations of $r < .30$ with at least 90% of the other items. Altogether, we excluded eleven items and reduced the DADYS questionnaire from 38 to 27

items selected fur further validation, thus confirming our first hypothesis.

Factorial structure of AD

The three factor-solution of the EFA in one half of the sample was the first that yielded good model-fit (CFI = 0.988, TLI = 0.985, RMSEA = 0.056 (90% CI: 0.046-0.066), SRMR = 0.052) and clear interpretability of the factors. Factor 1 constitutes the AD-core factor and comprises 18 items describing irritability and emotional impulsivity (AD-II). Factor 2 is composed of five items, describing positive emotionality (AD-PE), whereas the four items of factor 3 characterize exuberance (AD-EX), which in distinction to impulsive behavior is free of value and also comprises positive outbursts of emotions. Factor loadings can be found in Table S2. We found significant correlations ($p < .05$) between AD-II and AD-PE ($r = -.50$) and AD-II and AD-EX ($r = .38$).

We then assessed the three-factor structure applying CFA in the other half of the sample. Items 4, 5, 9 and 11 were recoded to load positively on their respective factor. Model fit indices were in an adequate to good range, expect for the RMSEA (CFI = 0.964, TLI = 0.961, RMSEA 0.090 (90% CI: 0.083-0.098), SRMR = 0.077) and all items loaded significantly ($p < .001$) on their corresponding factor (Table S3). A careful inspection of the MI's revealed that the two largest MI's suggested allowing for the additional residual correlations between items 3 ("*responds positively to adults*") and 7 ("*responds positively to peers*") [MI: 145.48], as well as between items 5 (*"Calms down after being angry"*) and 34 ("*Stays angry"*) [MI: 66.74], due to their very close proximity in content. After this adaptation, model fit indices were all in an adequate to good range (CFI = 0.978, TLI = 0.976, RMSEA = 0.071 (90% CI: 0.062-0.079), SRMR = 0.066). Items 3 and 7 as well as items 5 and 34 correlated significantly ($p < 0001$), with $r = .70$, and $r = .67$, respectively. The internal consistency for all dimensions was good to excellent, with Cronbach's α ranging from 0.72 for AD-PE to 0.96 for AD-II. This stable three-factor solution with an AD-core factor (AD-II), as well as two additional smaller AD-factors (AD-PE; AD-EX) confirms our second hypothesis.

Table 2 Comparison of Model Fit Indices

Model	χ^2 *(df)*	CFI	TLI	RMSEA (90% CI)	SRMR
Uni	4670.937* (1272)	0.918	0.914	0.083 (0.080, 0.085)	0.095
CFO	2451.262* (1257)	0.971	0.970	0.049 (0.046, 0.052)	0.060
SOF	2728.702* (1266)	0.965	0.963	0.054 (0.052, 0.057)	0.069
BI	2436.548* (1220)	0.971	0.968	0.051 (0.048, 0.053)	0.061
BI S-1 (Ref. = AD-II)	2294.591* (1228)	0.974	0.972	0.047 (0.044, 0.050)	0.055
BI S-1 (Ref. = ODD-D)	2103.000* (1215)	0.979	0.977	0.043 (0.040, 0.046)	0.049
BI S-1 (Ref. = ADHD-IN)	2274.056* (1219)	0.975	0.972	0.047 (0.044, 0.050)	0.052
BI S-1 (Ref. = ADHD-HI)	2148.761* (1221)	0.978	0.976	0.044 (0.041, 0.047)	0.050

Note. χ^2 = Chi-Square, *df* = degrees of freedom, CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation, CI = confidence interval, SRMR = standardized root mean square residuals, Uni = unidimensional model, CFO = factor model with six correlated first-order factors, SOF = factor model with six correlated first-order factors and one second-order factor, BI = bifactor model, BI S-1 = bifactor S-1 model (see Fig. 1), Ref. = reference factor. AD-II = affective dysregulation – irritability/emotional impulsivity, ODD-D = oppositional defiant disorder – only defiant dimension, ADHD = attention-deficit/hyperactivity disorder, IN = inattention, HI = hyperactivity/impulsivity

* = $p < .001$

Latent factor analysis including ADHD and ODD

In terms of global model fit, the third hypothesis was not supported. As indicated by the model fit indices alone, the global model fit of the CFO, the SOF, the BI and the BI S-1 did not differ significantly from one another. Only the unidimensional model yielded inadequate fit and will therefore not be discussed further (Table 2). For the CFO model (Table S4), all factor loadings on their respective factor were high and significant ($p < .001$). All six factors correlated strongly ($r = -.48$ to $r = .85$) and significantly ($p < .001$) with each other (Table 3). For the SOF model (Table S4), all factor loadings on the first-order factors were high and significant ($p < .001$). As shown in Table S5, all first-order factors loaded significantly ($p < .001$) on the second-order factor (-0.70 for AD-PE to 0.90 for AD-II). All first-order factors showed significant ($p < .001$) residual variance (0.18 for AD-EX to 0.51 for AD-PE). The Schmid-Leiman transformation (Table S6) showed that in total, item loadings were stronger on the second-order factor than on the first-order factors (*Mean[M]* = 0.61 for the second-order vs. $M = .30$ for the s-factors). An examination of the g/s loading ratios revealed that this pattern was not equally pronounced for all dimensions: The g/s loading ratio was highest for the dimension AD-EX (2.16), followed by AD-II (2.01) and ODD-D (1.99). The ratio was closer to 1 for the dimension

Table 3 Correlations Between the First-Order Factors

Variable	1	2	3	4	5	6
1. AD-II	–					
2. AD-PE	−0.71*	–				
3. AD-EX	0.84*	−0.56*	–			
4. ADHD-IN	0.70*	−0.54*	0.69*	–		
5. ADHD-HI	0.73*	−0.48*	0.82*	0.79*	–	
6. ODD-D	0.85*	−0.64*	0.74*	0.67*	0.71*	–

Note. AD = affective dysregulation, II = irritability/emotional impulsivity, PE = positive emotionality, EX = exuberance, ADHD = attention-deficit/hyperactivity disorder, IN = inattention, HI = hyperactivity/impulsivity, ODD-D = oppositional defiant disorder – only defiant dimension

* = $p < .001$

AD-HI (1.66), ADHD-IN (1.40) and AD-PE (-0.98). The BI model yielded some anomalies, such as non-significant factor loadings and negative residual items variances (Table S7). The bifactor-specific indices ωH/ ωHS and ECV indicated a strong g-factor and weakly defined s-factors. ωHS ranged between 0.19 for ODD-D to 0.44 for AD-PE, thus remaining below the recommended cut-off of 0.50 by Reise et al., (2013). The ECV for all s-factors combined lay at 29% and ranged from 2% for ODD-D to 8% for AD-II. The BI S-1 model showed significant factor loadings ($p < .001$) and no anomalies (Table S8). Similar to the BI model, ωH/ ωHS and ECV suggested a strong g-factor as well as weak s-factors (Table 4). Compared to the BI model, ωHS values were higher, albeit still below the recommended cut-off and ranged between 0.26 for ODD-D and 0.47 for ADHD-IN. The ECV for all s-factors combined lay at 30%, with values ranging from 3% for AD-PE, AD-EX and ODD-D to 0.11% for ADHD-HI. As shown in Table S9, some residual correlations between the first-order factors remained significant ($p < .05$), the two highest being the correlation between the two ADHD dimensions ($r = .58$) and the correlation between ADHD-HI and AD-EX ($r = .56$). The good model fit of the BI S-1 model, the significant factor loadings of items from all dimensions on the reference factor AD-II, the absence of anomalous factor loadings, the remaining significant residual correlations as well as the *a priori* defined general reference factor and the resulting straightforward interpretation of g- and s-factors and their relation to one another, support our third hypothesis that the BI S-1 model captures the data best.

Convergent and divergent validity

To assess convergent and divergent validity, we included the PROMIS Parent Proxy Anger Scale, FRUST-A and FRUST-M in our BI S-1 model and computed (residual) correlations. (Table S10). The PROMIS Parent Proxy Anger Scale correlated significantly ($p < .05$) with all other residual factors, except for ODD-D. The highest correlation was found with the reference factor AD-II ($r = .90$), all other correlations were small and ranged from $r = .08$ to $r = .14$. FRUST-A correlated significantly negative ($p < .05$) with the reference factor AD-II ($r = -.71$) as well as the residual factors AD-PE ($r = -.46$) and the PROMIS Scale ($r = -.59$). FRUST-M correlated significantly ($p < .05$) with all residual factors, as well as the PROMIS scale and FRUST-A. The highest correlations were found with AD-II ($r = .82$), PROMIS ($r = .74$) and FRUST-A ($r = -.77$). All other correlations were small and ranged from $r = -.08$ to $r = .18$. The observed differential correlations of the PROMIS scale and emotion regulation strategies with all AD, ADHD and ODD dimensions confirm our fourth hypothesis.

Comparison of reference factors

Out of the four BI S-1 models, the model with ODD-D as reference factor showed the highest ECV of the reference factor (72%), followed by the models with AD-II (70%), ADHD-HI (64%) and ADHD-IN (60%) as reference factors. The models did not show significant differences in model fit. The BI S-1 model with ODD-D as reference factor showed non-significant and negative factor loadings on the s-factor AD-II. Taken together, our fifth hypothesis can be confirmed, as the model explaining the most common variance (BI S-1; ODD-D) showed a number of anomalous factor loadings, with that limiting straight-forward interpretability of all factors.

Discussion

In the present study, we performed item reduction of the DADYS questionnaire, investigated the factorial structure of AD in a sample screened in the community, including children with and without AD, assessed associations between AD, ADHD and ODD dimensions, examined convergent and divergent validity of our dimensions and compared bifactor S-1 models with different reference factors.

To establish a factorial structure of our broadly defined AD concept, we first excluded eleven items from the questionnaire due to salient item descriptive statistics or item redundancy. Item redundancy was to be expected as the

questionnaire was created by combining items from several questionnaires assessing AD. Some other items (e.g. "*takes pleasure in distress of others*") that were excluded due to salient descriptive statistics may have been too negatively connotated for our sample consisting of relatively young children (8–12 years), about half of whom did not meet criteria for AD.

Through EFA and CFA, including the process of cross-validation, we established a stable factorial structure of AD, which comprised one core-factor, describing irritability/emotional impulsivity, as well as two smaller AD factors, representing positive emotionality and exuberance. This core factor of irritability and emotional impulsivity is in line with the chronic irritability/anger specifier suggested by Evans et al., (2017). Similarly, the core criteria of DMDD are severe temper tantrums and persistent irritability or anger, corresponding to our core-factor AD-II. In our sample, we could not find the differentiation between the trait component of chronic irritability and the state component of temper tantrums. In view of the fact that the conceptualization and assessment of AD have not yet been unequivocally determined, our results contribute to the standardization of this concept and suggest a broader conceptualization, with a strong core factor of irritability and emotional impulsivity.

In order to examine how AD relates to the externalizing disorders ADHD and ODD and to assess whether irritability/emotional impulsivity might be the common core feature underlying ADHD and ODD symptomatology, we specified our hypothesized bifactor S-1 model with 'pure indicators' of AD core symptomatology as general reference factor as a potential improvement to the traditional bifactor model, as well as four additional latent factor models to test for competing theories such as the assumption of an externalizing spectrum (Achenbach & Edelbrock, 1981) or a HiTOP-like organization of symptoms and dimensions (Kotov et al., 2017). High correlations between all the factors observed in the CFO suggested a common factor connecting these disorders within the externalizing spectrum, which we subsequently assessed with a variety of models. The good to adequate model fit of all models apart the unidimensional model should never be understood as a "decision-maker" *for* a model, but can only help us in our decision process, mainly by excluding models with an inadequate fit. This especially holds true for bifactor models, which tend to overfit (Bonifay et al., 2017) as a result of being less restrictive compared to other latent factor models. The unidimensional model showing an inadequate fit suggests that a model capturing general and well as specific aspects is better suited. Results from the higher-order and the bifactor models demonstrate that the specification of a common factor is justified. Compared to higher-order factor models, bifactor models come with the advantage that here, s-factors model dimensionality *beyond* the general factor, whereas in higher-order factor models, the lower-order factors model dimensionality *within* the general factor (Hartman, 2021). A problem arose from the varying loading ratios for g- and s-factors in the SOF model, as well as the differential ECV in the traditional bifactor model, which indicated that the associations of the individual domains with g differed, giving rise to the question of what this second-order g-factor really represents. As frequently observed in bifactor models (Eid et al., 2017), anomalous factor loadings in the traditional bifactor model impeded the interpretation of g- and s-factors. ECV values indicated that the g-factor did not represent all dimensions equally well, changing its meaning from a truly general factor to a mainly AD-II/ODD defined factor. Moreover, we found weakly defined s-factors, with ωHS values below the recommended cut-off of .50 (Reise et al., 2013). Though omega statistics are popular bifactor-specific indices, they are not without criticism (for a detailed discussion regarding problems with the ωHS see Rodriguez et al., 2016). According to the ωHS values alone, s-factors are often considered unstable and are dismissed for that reason. Hartman (2021), however, pointed out that unstable s-factors are only a reflection of what has already been well established from factor analytic effort – namely that dimensions free from the dominant general factor frequently display a "chaotic covariance structure of high instability" (p. 72). Their suggestion is to use bifactor models in a top-down fashion, choosing theoretically derived relevant item clusters, and in a next step to the design strong measures of these s-factors, instead of dismissing weak s-factors altogether. Interpreting the usefulness of specific factors for clinical practice or the individual research question based on other indicators such as the ECV, in combination with theoretical considerations, might therefore be a necessary first step, instead of solely relying on predefined global cut-off values. In the BI model, the ECV for all s-factor combined lay at 29% after partialling out the influence of the g-factor, therefore explaining an important part of the variance. These findings suggest that the s-factors are still meaningful, despite not meeting the cut-off criteria for ωHS. As we were interested in how well AD-core symptomatology defined through irritability and emotional impulsivity was able to explain the associations between ADHD and ODD, we applied the bifactor S-1 model with AD-II symptoms as pure indicators of the general reference factor. The bifactor S-1 model combines some of the advantages of a first-order correlated factor model and the traditional bifactor model: It retains the straightforward interpretability of all factors and their relation to each other of the CFO, and allows for the specification of a factor that explains variance common to all other s-factors, albeit to different extents, of the traditional bifactor model. In the bifactor S-1 model, all other s-factors can

be interpreted in relation to the general reference domain. Additionally, residual correlations can be interpreted meaningfully as partial correlations. Significant factor loadings and no anomalous results were observed, which greatly facilitated interpretation and is in accordance with previous studies pointing to a facilitated and less ambiguous interpretation of bifactor S-1 models compared to traditional bifactor models (Burns et al., 2020; Junghänel et al., 2020). Regarding the omega statistics, a similar, though slightly improved pattern compared to the traditional bifactor model was observed. All s-factors combined explained 30% of the variance in this model, after partialling out the influence of the reference factor, with ADHD-IN (10%) and ADHD-HI (11%) explaining the largest amount of the residual variance. The particularly high residual correlations between ADHD-IN and ADHD-HI showed that beyond the irritable/impulsive component, there is a remaining ADHD-specific component. Taken together, this model nicely demonstrates that the broadly defined AD-II factor captures important aspects of both ADHD dimensions, as well as the ODD-D dimension. At the same time, important disorder-specific aspects and dimensions remain and help us obtain a more nuanced picture of the associations between the symptom complexes. The differentiation of g- and s-factors brings along a number of potential advantages for research and clinical practice. Measures based on an array of problems tend to be particularly good prognosticators (Achenbach, 2021) and might be especially valuable for improving the diagnostic process (Lahey et al., 2021). In line with this, Pettersson et al., (2021) hypothesizes that the g-factor of psychopathology might be as useful for the psychiatric domain as the g-factor of intelligence has proven to be for the educational domain. It has been suggested that the g-factor, capturing correlations between different psychopathological dimensions, might result mainly from nonspecific etiological factors (Lahey et al., 2017) – a hypothesis that potentially provides a great target for further RDoC research. The s-factors might differentiate better between patients, especially those with a broad range of problems, who frequently show elevated scores on a range of scales (Pettersson et al., 2021). Longitudinal studies that examine the differential pathways, including specific protective and risk factors for people scoring high on different s-factors, are necessary.

The observation of remaining significant ODD and ADHD aspects that appear to be mostly free of irritability, emotional impulsivity, and emotion dysregulation is interesting. In future research, it would be worthwhile to examine more closely what these residual symptom complexes represent and what they are associated with. With regard to ODD, our results strengthen previous findings (Evans et al., 2017; Runions et al., 2016) reporting a distinction between a defiant and an irritable dimension of ODD, thus showing differential pathways and associations. Compared to the irritable dimension, the defiant dimension, which we identified as an important ODD-remaining aspect, has been shown to be associated more strongly with the odds of violence as an adult (Althoff et al., 2014) as well as disruptive and aggressive behavior (Burke et al., 2021). Regarding the ADHD remaining aspects that are free of irritability and emotional impulsivity, research has suggested particularly strong associations of the ADHD-IN dimension with sluggish cognitive tempo (Hartman et al., 2004), two concepts that are significantly related to academic functioning (Langberg et al., 2014). One could hypothesize that the ADHD-HI dimension now mainly captures hyperactivity and physical impulsivity, which is in line with the general understanding of impulsivity as a multidimensional construct (Berg et al., 2015). In future studies, it would be of great interest to ascertain, whether these residual ADHD- and ODD-specific aspects are now associated even more strongly with their respective external correlates. In a next step, and in line with Hartman (2021), strong measures that specifically assess these aspects could be designed.

To assess the convergent and divergent validity of our bifactor S-1 model, we added the three factors PROMIS Parent Proxy Anger Scale, FRUST-A and FRUST-M to our model and computed correlations with the reference factor as well as residual correlations with all s-factors. The strong correlation of the established PROMIS Parent Proxy Anger Scale with our AD-II reference factor, in combination with the low correlations of this scale with all other factors of our model, suggest that in line with our assumption, all scales except for AD-II indeed measure distinct constructs, independent of anger/rage. The correlations between the dimensions and emotion regulation strategies differ greatly in strength, with AD-II showing strong correlations with both, FRUST-A and FRUST-M. This constitutes an important validation of our model, because emotion dysregulation, has been strongly associated with ADHD and ODD (Ambrosini et al., 2013; Shaw et al., 2014) and corresponds here to our reference factor AD-II, while showing only small correlations with most of the remaining factors. Interestingly, there is one exception, as AD-PE correlate moderately with FRUST-A but not with FRUST-M. This supports the frequent finding that adaptive and maladaptive emotion regulation are distinct and not simply opposing constructs (Aldao & Nolen-Hoeksema, 2010; Grob & Smolenski, 2005). Indeed, the use of adaptive emotion regulation strategies might even be a protective factor for ADHD- and ODD-related symptomatology, beyond the mere absence of maladaptive emotion regulation strategies. Correspondingly, adaptive emotion regulation strategies, such as reappraisal

Table 4 Omega Statistics and Explained Common Variance of all Bifactor S-1 Models

AD-II	AD-II	AD-PE	AD-EX	ADHD-IN	ADHD-HI	ODD-D
ω	**0.98**	0.78	0.91	0.97	0.96	0.91
ωH	**0.90**					
ωHS		0.41	0.33	0.47	0.46	0.26
ECV	**0.70**	0.03	0.03	0.10	0.11	0.03
ODD-D						
ω	0.98	0.76	0.91	0.96	0.96	**0.98**
ωH						**0.90**
ωHS	0.11	0.40	0.26	0.43	0.35	
ECV	0.06	0.03	0.03	0.09	0.08	**0.72**
ADHD-IN						
ω	0.97	0.77	0.89	**0.98**	0.96	0.92
ωH				**0.82**		
ωHS	0.51	0.54	0.43		0.35	0.51
ECV	0.19	0.04	0.04	**0.60**	0.08	0.06
AD-HI						
ω	0.97	0.77	0.87	0.96	**0.98**	0.92
ωH					**0.85**	
ωHS	0.48	0.61	0.18	0.37		0.45
ECV	0.18	0.05	0.02	0.07	**0.64**	0.05

Note. Values for reference domain are marked in bold. AD = Affective dysregulation, II = irritability/emotional impulsivity, PE = positive emotionality, EX = exuberance, ADHD = attention-deficit/hyperactivity disorder, IN = inattention, HI = hyperactivity/impulsivity, ODD-D = oppositional defiant disorder – only defiant dimension, ω = omega (amount of variance accounted for by the g- and s-factors taken together), ωH = omega hierarchical (amount of variance accounted for by the g-factor), ωHS = omega hierarchical subscale (amount of variance accounted for by the s-factors), ECV = explained common variance (proportion of all common variance explained by that factor; for specific factors, the ECV computes the strength of a specific factor relative to all explained variance only of all items, even those not loading on the specific factor).

and acceptance, have been associated with fewer symptoms of psychopathology (Aldao et al., 2010, Braet et al., 2014).

In a final step, we compared bifactor S-1 models with different reference factors in order to examine, if AD-II indeed was the ideal core component of ADHD and ODD as hypothesized. All models yielded a good model fit and explained between 60% (ADHD-IN) to 72% (ODD-D) of the common variance (Table 4). The high ECV of all four BI S-1 models supports the general suggestion of a more parsimonious structure of psychopathology, as intended by the specifier approach in the ICD-11 (WHO, 2020) for ODD and AD symptomatology, instead of frequently diagnosing several comorbid disorders, such as ADHD *and* ODD or ADHD *and* DMDD. Despite the high ECV for the respective general factor, we observed a large gain of information through the assessment of all additional dimensions, indicated by the ECV for the s-factors in all models (Tables S8, S11-S13). Based on model fit alone, no model could be excluded at this point, which is unsurprising given the high correlations between all dimensions. The model with ODD-D as general reference factor demonstrates a slightly higher ECV (72%) than the model with AD-II (70%) as general reference factor. However, in the ODD-D model, the s-factor AD-II collapsed, as indicated by several anomalous factor loadings as well as the outstandingly low ωHS value (ωHS = 0.11), exacerbating straightforward interpretability and suggesting that this model does not capture the data in the most ideal way. The model with AD-II as reference factor explains a similar amount of common variance (70%) and shows no anomalous factor loadings or collapsing s-factors, and is straightforward to interpret. As pointed out above, a model should never be chosen based on statistical indices alone. The model with AD-II is derived based on theoretical considerations and fits well. The model with ODD-D as general reference factor shows good model fit, but does come with a number of statistical and theoretical problems, suggesting that AD-II might be a better reference factor than ODD-D: First, it has been suggested that the defiant dimension of ODD appears only later in life, as a consequence of impulsivity and weak emotion regulation skills (Beauchaine et al., 2010). Therefore, conceptualizing ODD-D as core feature, might not capture the problem behavior of all children, especially not of younger ones. This will have to be assessed in longitudinal studies. Second, when conceptualizing irritability/emotional impulsivity as core factor, this more parsimonious structure could potentially be extended into the internalizing spectrum. This seems worth examining for several reasons. In the DSM-5, DMDD is classified as a unipolar mood disorder, and previous studies have shown a strong association of AD symptomatology with internalizing symptomatology such as depression/dysthymia and anxiety (Copeland et al., 2013; Evans et al., 2017; Leibenluft, 2011; Waldman et al., 2021). The high correlations of AD-II with emotion regulation also suggest a potential association with

the internalizing spectrum, as mood and anxiety disorders have been strongly associated with more maladaptive and less adaptive emotion regulation strategies (Carthy et al., 2010; Silk et al., 2003). We therefore suggest an adaption and extension of the specifier approach. Our results indicate that irritability/emotional impulsivity could be the crucial factor explaining the high correlations frequently found between ADHD and ODD and could therefore be modelled as the core factor for ADHD- and ODD-related symptomatology. Related specifiers, such as "*with inattention*", "*with defiant behavior*" and "with *hyperarousal*" could be added to this core factor, leading to a more parsimonious structure of psychopathology instead of major symptom overlap and a number of comorbid diagnoses.

Limitations

One limitation of our study is that due to the relatively small age range of the children (8–12 years), we are unable to draw any conclusions about emotional impulsivity and irritability in children beyond that age range. In future research, it will be important to assess whether emotional impulsivity and irritability are similarly strongly associated with an array of disorders in other age groups. This is important as the AD-II factor showed a strong correlation with adaptive and maladaptive emotion regulation strategies, which has been suggested to be age dependent, potentially in the sense of a maladaptive shift, describing a reduction of adaptive strategies during adolescence (Cracco et al., 2017). Additionally, it will be important to assess whether similar results can be obtained in samples from other cultural and linguistic backgrounds. Another limitation is that we employed a parent rating scale to examine the associations between the relevant dimensions as our only measure, which will have to be expanded upon in future studies. We did not include any internalizing disorders (e.g. anxiety or depressive disorders) in our analyses. AD, however, has often been referred to as a transdiagnostic dimension, and it would be of great interest to examine how it relates to disorders in the internalizing spectrum. Furthermore, external correlates of the residual ADHD and ODD dimensions, such as sluggish cognitive tempo and violent behavior, should be examined in future research. Due to our cross-sectional study design, we cannot draw any conclusions relating to the onset and progression of AD and other externalizing disorders or the identification of risk or protective factors. It would be interesting to examine if the number of specifiers increases with age in irritable/emotionally impulsive children and if so, which specifiers appear at what time. Moreover, it would be useful to examine whether adaptive emotion regulation strategies might be a protective factor and maladaptive strategies a general liability factor for irritability and emotional impulsivity, as this could result in targeted prevention programs at an early age.

Conclusions

In the present study, we examined the factorial structure of AD, and found a stable structure, entailing one core-component, describing irritability/emotional impulsivity and two smaller factors describing positive emotionality and exuberance. We found a bifactor S-1 model with AD-II as general reference factor to lead to a straightforward interpretation of the associations between all dimensions. AD-II captured a major part of the shared variance of all AD, ADHD and ODD dimensions and at the same time, all dimensions explained important additional variance. Correlations with external correlates validated our model. When comparing models with different reference factors, AD-II as reference factor captured the data better than models with ADHD-/ODD-dimensions as reference factor. Our results support the specifier-approach adopted by the ICD-11 for ODD and AD symptomatology. We suggest an adaptation and extension of this approach in the future by selecting irritability/emotional impulsivity as the core diagnostic dimension, to which specifiers such as "*with inattention*" "*with defiant behavior*" "*with hyperarousal*", could be added. Additional specifiers, potentially extending into the internalizing spectrum should be examined in future studies and could eventually lead to an even more parsimonious structure of psychopathology. Our results suggest that comorbidities arise largely through how diagnoses are currently determined in the DSM-5 (APA, 2013) and the ICD-11 (WHO, 2020) – namely the combination of several domains within one disorder. The specifier approach might offer us a more accurate, richer, less stigmatizing and at the same time more parsimonious description of patients (Ruggero et al., 2019), which could additionally improve communication in research and clinical practice and lead to better treatment approaches.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10862-022-09974-8.

Acknowledgements Other members of the ADOPT Consortium: Aggensteiner Pascal, Bienioschek Stefanie, Brandeis Daniel, Breier Maurice, Dobler Veronika, Giller Franziska, Groh Monja, Heintz Stefan, Hellmich Martin, Hohmann Sarah, Igel Christine, Kaiser Anna, Katmer-Amet Betül, Katzmann Josepha, Koppisch Katrin, Kuhnke Kristin, Nickel Theresa, Otto Christiane, Ritschel Anne, Rodova-Ghasemi Elisaveta, Samaras Angelina, Schreiner Anne, Schroth Jennifer, Steiner Marie, Steiner Marion, Steinhauser Susanne and Winkler Matthias. This trial was supported by the Clinical Trials Centre Cologne (CTCC), Medical Faculty, University of Cologne, which

performed the monitoring. Study data were collected and managed using REDCap electronic data capture tools hosted at the Clinical Trials Centre Cologne (https://projectredcap.org/resources/citations/). Data preparation was conducted by Ingrid Becker from the Institute of Medical Statistics and Computational Biology (IMSB), Faculty of Medicine and University Hospital Cologne.

Authors' contributions Anja Görtz-Dorten, Tobias Banaschewski, Jörg M. Fegert, Charlotte Hanisch, Michael Kölch, Ulrike Ravens-Sieberer, Veit Roessner and Manfred Döpfner contributed to the study conception and design. Michaela Junghänel, Claudia Ginsberg, Franziska Frenk, Kristina Mücke, Sara Zaplana Labarga, Dorothee Bernheim, Anne Schüller, Anne Kaman and Julian Hinz performed data collection. Anne-Katrin Treier and Sabina Millenet coordinated the study and were responsible for data management. Analysis were performed by Michaela Junghänel. Statistical analyses were reviewed by Ann-Kathrin Thöne. The first draft of the manuscript was written by Michaela Junghänel and revised by Ann-Kathrin Thöne and Manfred Döpfner and all authors critically reviewed and approved the final manuscript.

Funding This work was supported by the research consortium on affective dysregulation, ADOPT and was funded by the German Federal Ministry of Education and Research (FKZ 01GL1741D).

Open Access funding enabled and organized by Projekt DEAL.

Availability of data and material: available upon reasonable request.

Code Availability available upon reasonable request.

Declarations

Conflicts of Interest/Competing interests Manfred Döpfner and Anja Görtz-Dorten are authors of DISYPS and receive royalties.

Ethics approval Approval was obtained from the ethics committee of the University of Cologne (18–033), the University Hospital of Dresden (EK 35409218), the University Hospital of Mannheim (2018-554 N-MA) and the University Hospital of Ulm (297/18). The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Trial registration ADOPT Online: German Clinical Trials Register (DRKS) DRKS00014963. Registered 27 June 2018.

Consent to participate The parents of all participating children and adolescents provided informed consent.

Consent for publication The parents of all participating children and adolescents provided consent for publication.

References

Achenbach, T. M. (2021). Hierarchical dimensional models of psychopathology: yes. *but… World Psychiatry*, 20(1), 64–65. https://doi.org/10.1002/wps.20810

Achenbach, T. M., & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs of the Society for Research in Child Development*, 1–82. https://doi.org/10.2307/1165983

Aldao, A., & Nolen-Hoeksema, S. (2010). Specificity of cognitive emotion regulation strategies: a transdiagnostic examination. *Behaviour Research and Therapy*, 48(10), 974–983. https://doi.org/10.1016/j.brat.2010.06.002

Althoff, R. R., Kuny-Slock, A. V., Verhulst, F. C., Hudziak, J. J., & Van Der Ende, J. (2014). Classes of oppositional-defiant behavior: Concurrent and predictive validity. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 55(10), 1162–1171. https://doi.org/10.1111/jcpp.12233

Ambrosini, P. J., Bennett, D. S., & Elia, J. (2013). Attention deficit hyperactivity disorder characteristics: II. Clinical correlates of irritable mood. *Journal of Affective Disorders*, 145(1), 70–76. https://doi.org/10.1016/j.jad.2012.07.014

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Barkley, R. A., & Fischer, M. (2010). The unique contribution of emotional impulsiveness to impairment in major life activities in hyperactive children as adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(5), 503–513. https://doi.org/10.1016/j.jaac.2010.01.019

Beauchaine, T. P., Hinshaw, S. P., & Pang, K. L. (2010). Comorbidity of attention-deficit/hyperactivity disorder and early-onset conduct disorder: Biological, environmental, and developmental mechanisms. *Clinical Psychology: Science and Practice*, 17(4), 327–336. https://doi.org/10.1111/j.1468-2850.2010.01224.x

Berg, J. M., Latzman, R. D., Bliwise, N. G., & Lilienfeld, S. O. (2015). Parsing the heterogeneity of impulsivity: A meta-analytic review of the behavioral implications of the UPPS for psychopathology. *Psychological Assessment*, 27(4), 1129–1146. https://doi.org/10.1037/pas0000111

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5(1), 184–186. https://doi.org/10.1177/2167702616657069

Braet, C., Theuwis, L., Van Durme, K., Vandewalle, J., Vandevivere, E., Wante, L., … & Goossens, L. (2014). Emotion regulation in children with emotional problems. *Cognitive Therapy and Research*, 38(5), 493–504. https://doi.org/10.1007/s10608-014-9616-x

Brotman, M. A., Schmajuk, M., Rich, B. A., Dickstein, D. P., Guyer, A. E., Costello, E. J. … Leibenluft, E. (2006). Prevalence, clinical correlates, and longitudinal course of severe mood dysregulation in children. *Biological Psychiatry*, 60(9), 991–997. https://doi.org/10.1016/j.biopsych.2006.08.042

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

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2), 230–258. https://doi.org/10.1177/0049124192021002005

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, 80(4), 796–846. https://doi.org/10.1111/j.1467-6494.2011.00749.x

Burke, J. D., Johnston, O. G., & Butler, E. J. (2021). The irritable and oppositional dimensions of oppositional defiant disorder: Integral factors in the explanation of affective and behavioral psychopathology. *Child and Adolescent Psychiatric Clinics of North America*, 30(3), 637–647. https://doi.org/10.1016/j.chc.2021.04.012

Burns, G. L., Geiser, C., Servera, M., Becker, S. P., & Beauchaine, T. P. (2020). Application of the bifactor S – 1 model to multisource ratings of ADHD/ODD symptoms: an appropriate bifactor model for symptom ratings. *Journal of Abnormal Child Psychology*, 48(7), 881–894. https://doi.org/10.1007/s10802-019-00608-4

Carragher, N., Krueger, R. F., Eaton, N. R., & Slade, T. (2015). Disorders without borders: current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, 50(3), 339–350. https://doi.org/10.1007/s00127-014-1004-z

Carthy, T., Horesh, N., Apter, A., & Gross, J. J. (2010). Patterns of emotional reactivity and regulation in children with anxiety disorders. *Journal of Psychopathology and Behavioral Assessment*, 32(1), 23–36. https://doi.org/10.1007/s10862-009-9167-8

Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a functional view of the p factor in psychopathology. *Clinical Psychological Science*, 5(5), 880–889. https://doi.org/10.1177/2167702617710037

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S. … 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

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902

Copeland, W. E., Angold, A., Costello, E. J., & Egger, H. (2013). Prevalence, comorbidity, and correlates of DSM-5 proposed disruptive mood dysregulation disorder. *American Journal of Psychiatry*, 170(2), 173–179. https://doi.org/10.1176/appi.ajp.2012.12010132

Cracco, E., Goossens, L., & Braet, C. (2017). Emotion regulation across childhood and adolescence: evidence for a maladaptive shift in adolescence. *European Child & Adolescent Psychiatry*, 26(8), 909–921. https://doi.org/10.1007/s00787-017-0952-8

Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: The seven pillars of RDoC. *BMC Medicine*, 11(1), 1–8. https://doi.org/10.1186/1741-7015-11-126

Döpfner, M., Breuer, D., Wille, N., Erhart, M., & Ravens-Sieberer, U. (2008). How often do children meet ICD-10/DSM-IV criteria of attention deficit-/hyperactivity disorder and hyperkinetic disorder? Parent-based prevalence rates in a national sample - results of the BELLA study. *European Child and Adolescent Psychiatry*, 17(1), 59–70. https://doi.org/10.1007/s00787-008-1007-y

Döpfner, M., & Görtz-Dorten, A. (2017). *Diagnostik-System f?r psychische Störungen nach ICD-10 und DSM-5 für Kinder und Jugendliche ? III [Diagnostic System of Mental Disorders in Children and Adolescents based on the ICD-10 and DSM-5] (DISYPS?III)*. Göttingen: Hogrefe

Döpfner, M., Katzmann, J., Hanisch, C., Fegert, J. M., Kölch, M., Ritschel, A., … & Banaschewski, T. (2019). Affective dysregulation in childhood - Optimizing prevention and treatment: Protocol of three randomized controlled trials in the ADOPT study. *BMC Psychiatry*, 19(1), 1–20. https://doi.org/10.1186/s12888-019-2239-8

Dueber, D.M. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. https://doi.org/10.13023/edp.tool.01 [Available at http://sites.education.uky.edu/apslab/resources/]

Eaton, N. R. (2015). Latent variable and network models of comorbidity: toward an empirically derived nosology. *Social Psychiatry and Psychiatric Epidemiology*, 50(6), 845–849. https://doi.org/10.1007/s00127-015-1012-7

Eid, M. (2020). Multi-faceted constructs in abnormal psychology: Implications of the bifactor S – 1 model for individual clinical assessment. *Journal of Abnormal Child Psychology*, 48(7), 895–900. https://doi.org/10.1007/s10802-020-00624-9

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in g -factor models: Explanations and alternatives. *Psychological Methods*, 22(3), 541–562

Erhart, M., Döpfner, M., & Ravens-Sieberer, U. (2008). Psychometric properties of two ADHD questionnaires: comparing the conners' scale and the FBB-HKS in the general population of german children and adolescents - results of the BELLA study. *European Child and Adolescent Psychiatry*, 17(1), 106–115. https://doi.org/10.1007/s00787-008-1012-1

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

Eyre, O., Langley, K., Stringaris, A., Leibenluft, E., Collishaw, S., & Thapar, A. (2017). Irritability in ADHD: Associations with depression liability. *Journal of Affective Disorders*, 215, 281–287. https://doi.org/10.1016/j.jad.2017.03.050

Forero, C. G., & Maydeu-Olivares, A. (2009). Estimation of IRT graded response models: limited versus full information methods. *Psychological Methods*, 14(3), 275–299. https://doi.org/10.1037/a0015825

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, 55, 57–68. https://doi.org/10.1016/j.intell.2016.01.006

Görtz-Dorten, A., Ise, E., Hautmann, C., Walter, D., & Döpfner, M. (2014). Psychometric properties of a german parent rating scale for oppositional defiant and conduct disorder (FBB-SSV) in clinical and community samples. *Child Psychiatry and Human Development*, 45(4), 388–397. https://doi.org/10.1007/s10578-013-0409-3

Görtz-Dorten, Thöne, & Döpfner (2022). DISYPS-ILF Interview-Leitfäden zum Diagnostik-System für psychische Störungen nach DSM-5 für Kinder und Jugendliche. Bern: Hogrefe

Grob, A., & Smolenski, C. (2005). FEEL-KJ. Fragebogen zur Erhebung der Emotionsregulation bei Kindern und Jugendlichen. Bern: Hans Huber

Hartman, C. A. (2021). The important gain is that we are lumpers and splitters now; it is the splitting that needs our hard work. *World Psychiatry*, 20(1), 72–73. https://doi.org/10.1002/wps.20816

Hartman, C. A., Willcutt, E. G., Rhee, S. H., & Pennington, B. F. (2004). The relation between sluggish cognitive tempo and DSM-IV ADHD. *Journal of Abnormal Child Psychology*, 32(5), 491–503. https://doi.org/10.1023/B:JACP.0000037779.85211.29

Heinrich, M., Geiser, C., Zagorscak, P., Burns, G.L., Bohn, J., Becker, S.P., … & Knaevelsrud, C. (2020). On the meaning of the general factor of psychopathology ("p-factor") in symmetrical bifactor models. *Preprint*. https://doi.org/10.31234/osf.io/syj9k

Holtmann, M., Goth, K., Wöckel, L., Poustka, F., & Bölte, S. (2008). CBCL-pediatric bipolar disorder phenotype: severe ADHD or bipolar disorder? *Journal of Neural Transmission*, 115(2), 155–161. https://doi.org/10.1007/s00702-007-0823-4

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60. https://doi.org/10.21427/D79B73

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D., Quinn, K. … Wang, P. (2010). Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379

Irwin, D. E., Gross, H. E., Stucky, B. D., Thissen, D., Dewitt, E. M., Lai, J. S. … Dewalt, D. A. (2012). Development of six PROMIS pediatrics proxy-report item banks. *Health and Quality of Life Outcomes*, 10(1), 1–13

Junghänel, M., Rodenacker, K., Dose, C., & Döpfner, M. (2020). Applying the bifactor S-1 model to ratings of ADHD/ODD symptoms: A commentary on Burns et al. (2019) and a re-analysis. *Journal of Abnormal Child Psychology*, *48*(7), 905–910. https://doi.org/10.1007/s10802-020-00637-4

Karalunas, S. L., Gustafsson, H. C., Fair, D., Musser, E. D., & Nigg, J. T. (2019). Do we need an irritable subtype of ADHD? Replication and extension of a promising temperament profile approach to ADHD subtyping. *Psychological Assessment*, 31(2), 236–247. https://doi.org/10.1037/pas0000664

Kotov, R., Waszczuk, M. A., Krueger, R. F., Forbes, M. K., Watson, D., Clark, L. A. … Zimmerman, S. E., M (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454–477. https://doi.org/10.1037/abn0000258

Krueger, R. F., Markon, K. E., Patrick, C. J., & Iacono, W. G. (2005). Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for DSM-V. *Journal of Abnormal Psychology*, 114(4), 537–550. https://doi.org/10.1037/0021-843X.114.4.537

Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, 143(2), 142–186. https://doi.org/10.1037/bul0000069.A

Lahey, B. B., Moore, T. M., Kaczkurkin, A. N., & Zald, D. H. (2021). Hierarchical models of psychopathology: empirical support, implications, and remaining issues. *World Psychiatry*, 20(1), 57–63. https://doi.org/10.1002/wps.20824

Langberg, J. M., Becker, S. P., & Dvorsky, M. R. (2014). The association between sluggish cognitive tempo and academic functioning in youth with attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology*, 42(1), 91–103. https://doi.org/10.1007/s10802-013-9722-3

Leibenluft, E. (2011). Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. *American Journal of Psychiatry*, 168(2), 129–142. https://doi.org/10.1176/appi.ajp.2010.10050766

Leibenluft, E., & Stoddard, J. (2013). The developmental psychopathology of irritability. *Development and Psychopathology*, 25(4pt2), 1473–1487. https://doi.org/10.1017/S0954579413000722

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Lochman, J. E., Evans, S. C., Burke, J. D., Roberts, M. C., Fite, P. J., Reed, G. M. … Garralda, E., M (2015). An empirically based alternative to DSM-5's disruptive mood dysregulation disorder for ICD-11. *World Psychiatry*, 14(1), 30–33. https://doi.org/10.1002/wps.20176

Malhi, G. S., & Bell, E. (2019). Fake views: DMDD, indeed!. *Australian and New Zealand Journal of Psychiatry*, 53(7), 706–710. https://doi.org/10.1177/0004867419863162

Markon, K. E. (2021). On hierarchically-informed measures of psychopathology. *World Psychiatry*, 20(1), 66–67. https://doi.org/10.1002/wps.20812

Mayes, S. D., Waxmonsky, J. D., Calhoun, S. L., & Bixler, E. O. (2016). Disruptive mood dysregulation disorder symptoms and association with oppositional defiant and other disorders in a general population child sample. *Journal of Child and Adolescent Psychopharmacology*, 26(2), 101–106. https://doi.org/10.1089/cap.2015.0074

Michelini, G., Palumbo, I. M., DeYoung, C. G., Latzmann, R. D., & Kotov, R. (2021). Linking RDoC and HiTOP: A new interface for advancing psychiatric nosology and neuroscience. *Clinical Psychology Review*, 102025. https://doi.org/10.1016/j.cpr.2021.102025

Mulraney, M., Schilpzand, E. J., Hazell, P., Nicholson, J. M., Anderson, V., Efron, D. … Sciberras, E. (2016). Comorbidity and correlates of disruptive mood dysregulation disorder in 6–8-year-old children with ADHD. *European Child and Adolescent Psychiatry*, 25(3), 321–330. https://doi.org/10.1007/s00787-015-0738-9

Muthén, L. K., & Muthén, B. O. (2017). Mplus User's Guide. Eighth Edition. *Los Angeles, CA: Muthén & Muthén*. https://doi.org/10.1111/j.1600-0447.2011.01711.x

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

Otto, C., Kaman, A., Barkmann, C., Döpfner, M., Görtz-Dorten, A., Ginsberg, C., … & Ravens-Sieberer, U. (2022). The DADYS-Screen: Development and Evaluation of a Screening Tool for Affective Dysregulation in Children. *Assessment*, 10731911221082709

Pettersson, E., Larsson, H., & Lichtenstein, P. (2021). Psychometrics, interpretation and clinical implications of hierarchical models of psychopathology. *World Psychiatry*, 20(1), 68–69. https://doi.org/10.1002/wps.20813

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129–140. https://doi.org/10.1080/00223891.2012.725437

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223–237. https://doi.org/10.1080/00223891.2015.1089249

Ruggero, C. J., Kotov, R., Hopwood, C. J., First, M., Clark, L. A., Skodol, A. E. … Zimmermann, J. (2019). Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. *Journal of Consulting and Clinical Psychology*, 87(12), 1069–1084. https://doi.org/10.1037/ccp0000452

Runions, K. C., Stewart, R. M., Moore, J., Martinez Ladino, Y., Rao, P., & Zepf, F. D. (2016). Disruptive mood dysregulation disorder in ICD-11: a new disorder or ODD with a specifier for chronic irritability? *European Child and Adolescent Psychiatry*, 25(3), 331–332. https://doi.org/10.1007/s00787-015-0789-y

Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 171(3), 276–293. https://doi.org/10.1176/appi.ajp.2013.13070966

Shi, D., Lee, T., Fairchild, A. J., & Maydeu-Olivares, A. (2020). Fitting ordinal factor analysis models with missing data: A comparison between pairwise deletion and multiple imputation. *Educational and Psychological Measurement*, 80(1), 41–66. https://doi.org/10.1177/0013164419845039

Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Developmental Psychology*, 33(6), 906. https://doi.org/10.1037/0012-1649.33.6.906

Silk, J. S., Steinberg, L., & Morris, A. S. (2003). Adolescents' emotion regulation in daily life: links to depressive symptoms and problem behavior. *Child Development*, 74(6), 1869–1880. https://doi.org/10.1046/j.1467-8624.2003.00643.x

Stringaris, A., & Goodman, R. (2009). Three dimensions of oppositionality in youth. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(3), 216–223. https://doi.org/10.1111/j.1469-7610.2008.01989.x

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 and Allied Disciplines*, 53(11), 1109–1117. https://doi.org/10.1111/j.1469-7610.2012.02561.x

Thöne, A. K., Junghänel, M., Görtz-Dorten, A., Dose, C., Hautmann, C., Jendreizik, L. T. … Döpfner, M. (2021). Disentangling symptoms of externalizing disorders in children using multiple measures and informants. *Psychological Assessment*, 33(11), 1065–1079. https://doi.org/10.1037/pas0001053

Tsikriktsis, N. (2005). A review of techniques for treating missing data in OM survey research. *Journal of Operations Management*, 24(1), 53–62. https://doi.org/10.1016/j.jom.2005.03.001

Varni, J. W., Thissen, D., Stucky, B. D., Liu, Y., Gorder, H., Irwin, D. E., … & DeWalt, D. A. (2012). PROMIS® Parent Proxy Report Scales: an item response theory analysis of the parent proxy report item banks. *Quality of Life Research*, 21(7), 1223–1240. https://doi.org/10.1007/s11136-011-0025-2

Waldman, I. D., Rowe, R., Boylan, K., & Burke, J. D. (2021). External validation of a bifactor model of oppositional defiant disorder. *Molecular Psychiatry*, 26(2), 682–693. https://doi.org/10.1038/s41380-018-0294-z

Waltereit, R., Giller, F., Ehrlich, S., & Roessner, V. (2019). Affective dysregulation: a transdiagnostic research concept between ADHD, aggressive behavior conditions and borderline personality traits. *European Child and Adolescent Psychiatry*, 28(12), 1551–1553. https://doi.org/10.1007/s00787-019-01438-x

Willcutt, E. G. (2012). The prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A meta-analytic review. *Neurotherapeutics*, 9(3), 490–499. https://doi.org/10.1007/s13311-012-0135-8

World Health Organization (2020). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

Zimmerman, M. (2021). Why hierarchical dimensional approaches to classification will fail to transform diagnosis in psychiatry. *World Psychiatry*, 20(1), 70–71. https://doi.org/10.1002/wps.20815