Disentangling Heterogeneity of Childhood Disruptive Behavior Problems Into Dimensions and Subgroups

Koen Bolhuis, MD, Gitta H. Lubke, PhD, Jan van der Ende, MSc, Meike Bartels, PhD, Catharina E.M. van Beijsterveldt, PhD, Paul Lichtenstein, PhD, Henrik Larsson, PhD, Vincent W.V. Jaddoe, MD, PhD, Steven A. Kushner, MD, PhD, Frank C. Verhulst, MD, PhD, Dorret I. Boomsma, PhD, Henning Tiemeier, MD, PhD

Objective: Irritable and oppositional behaviors are increasingly considered as distinct dimensions of oppositional defiant disorder. However, few studies have explored this multidimensionality across the broader spectrum of disruptive behavior problems (DBPs). This study examined the presence of dimensions and distinct subgroups of childhood DBPs, and the cross-sectional and longitudinal associations between these dimensions.

Method: Using factor mixture models (FMMs), the presence of dimensions and subgroups of DBPs was assessed in the Generation R Study at ages 6 (n = 6,209) and 10 (n = 4,724) years. Replications were performed in two population-based cohorts (Netherlands Twin Registry, n = 4,402, and Swedish Twin Study of Child and Adolescent Development, n = 1,089) and a clinical sample (n = 1,933). We used cross-lagged modeling in the Generation R Study to assess cross-sectional and longitudinal associations between dimensions. DBPs were assessed using mother-reported responses to the Child Behavior Checklist.

Results: Empirically obtained dimensions of DBPs were oppositional behavior (age 6 years), disobedient behavior, rule-breaking behavior (age 10 years), physical aggression, and irritability (both ages). FMMs suggested that one-class solutions had the best model fit for all dimensions in all three population-based cohorts. Similar results were obtained in the clinical sample. All three dimensions, including irritability, predicted subsequent physical aggression (range, 0.08–0.16).

Conclusion: This study showed that childhood DBPs should be regarded as a multidimensional phenotype rather than comprising distinct subgroups. Incorporating multidimensionality will improve diagnostic accuracy and refine treatment. Future studies need to address the biological validity of the DBP dimensions observed in this study; herein lies an important opportunity for neuroimaging and genetic measures.

Key words: disruptive behavior disorder, classification, *DSM-5*, irritable mood, factor mixture model

J Am Acad Child Adolesc Psychiatry 2017;56(8):678–686.

The nosology of childhood disruptive behavior disorders has given rise to considerable academic debate, even since before the disorders were operationalized by the *DSM* in 1980.[1-5] Many studies using different informants, instruments, and study populations have addressed the heterogeneity and developmental continuities of disruptive behavior disorders.[6-11] More recently, with the development of the *DSM-5*, several changes in the criteria for oppositional defiant disorder (ODD) and conduct disorder (CD) were made. For example, it can now be specified whether CD had its onset before the age of 10 years, which is indicative of a poorer prognosis.[9,12] Another important change is the possibility to differentiate irritable from oppositional ODD subtypes.[13,14] However, to our knowledge, these studies have not assessed whether the ODD dimensions can be discerned on a broader spectrum of disruptive behavior problems (DBPs) beyond a priori−defined *DSM* criteria, which would strengthen our current diagnostic frameworks with an empirical basis.

Recent research established that ODD subtypes have divergent developmental courses; most notably, irritability is associated with later depression and anxiety.[13-15] Findings from behavioral genetics studies have provided further support for distinguishing irritable from oppositional symptoms,[16] and the *DSM-5* now allows for better classification of ODD symptoms along these dimensions. It is still unclear, however, how irritability is related to other DBPs. So far, significant associations of irritability with oppositionality, but less so with CD, have been described.[17] However, earlier work from this cohort demonstrated that both irritable and headstrong dimensions predicted later ODD, CD, and depression to a similar extent.[18] On the basis of these studies that emphasize the distinct developments of ODD and CD,[10,19] the *DSM-5* posits CD as a disorder of

CG Clinical guidance is available at the end of this article.

Supplemental material cited in this article is available online.

physical violence and delinquency, and ODD as a disorder of oppositionality and irritability. However, comorbidity between these disorders is common.[8] It might well be that irritability is a distinct dimension on the broad spectrum of DBPs that influences the development of other DBP dimensions, for example, aggression and noncompliance. Indeed, developmental studies have provided preliminary evidence for this, as reviewed by Wakschlag *et al.*[20] They discussed that problematic defiance/rule-breaking is often associated with negative affect, but most of what is known about this association is derived from small observational studies. Unfortunately, empirical studies investigating irritability across the broad spectrum of DBPs are lacking.

Developmental scientists have stressed the importance of disentangling the heterogeneity of disruptive behavior,[5-11,20] and many studies have addressed this with various approaches. One option is classifying DBPs by age of onset, as proposed by the *DSM-5*. However, this will in practice be less useful for clinicians, as retrospective symptoms recall is often unreliable.[4,21] Furthermore, it is still unclear whether early-onset DBPs will be limited to childhood or whether these children will continue to have problems later in life.[9] Thus, an empirically based refinement of diagnosis based on the pattern of symptoms a child or adolescent exhibits could be more promising. Given the recent interest in the irritability subtype of ODD, it is important to examine how irritability is associated with other DBPs. A seminal meta-analysis of factor analytic studies by Frick *et al.* has demonstrated that DBPs can be classified along two principal axes, namely, overt/covert, and destructive/nondestructive,[6] with different developmental trajectories.[7,21] So far, no study has identified a distinct irritability dimension on the broader spectrum of DBPs, even though irritability has been found to be a distinguishable dimension of ODD specifically. In addition, it remains to be studied more thoroughly how irritability is associated with other DBP dimensions over time. To move toward a more developmentally sensitive nosology of DBPs that would transcend current diagnostic boundaries, it is crucial to examine these symptom patterns across ages.[18,20,21]

This study had two aims. First, we empirically assessed the multidimensionality across the whole spectrum of childhood DBPs, while simultaneously examining whether meaningful subgroups could be discerned. Hereby we wish to extend recent research that has focused on ODD symptoms specifically, with the goal of testing whether a distinct irritability dimension can be distinguished on a broader DBPs spectrum. The majority of studies on the heterogeneity of DBPs used either dimensional (e.g., factor analysis) or categorical (e.g., latent class analysis) statistical methods.[6,9,14,16-18] We performed factor mixture models (FMMs), which allow the presence of both dimensional and categorical latent variables and are therefore appropriate for studying the heterogeneity of psychiatric problems.[22,23] This is important, as recent studies have examined the latent structures of DBPs without clearly characterizing the dimensional or categorical latent structures of DBPs.[6,13,14,16,17] Second, longitudinal associations between the different dimensions of DBPs were studied using a cross-lagged model. Data from three population-based cohorts were used, as replication is important for FMMs. In addition, we explored consistency in a sample of clinically referred children in order to test generalizability. Although different population subgroups might be present in clinical samples due to referral bias,[22] these analyses will aid translation to clinical practice. Hereby our findings could be more easily interpreted by clinicians treating children with DBPs.

METHOD

Study Populations

This study was conducted using data from three population-based cohorts that collaborate under the FP7-ACTION consortium. Primary analyses were conducted in the Generation R Study, a prospective population-based cohort from fetal life onward, which included 9,778 pregnant women living in Rotterdam, the Netherlands. The aim of the Generation R project is to identify early environmental and genetic factors that affect health and development.[24] For the current study, data were used from two time points. At age 6 years, 6,209 children with behavioral data were included in the analyses and comprised fewer children of ethnic minorities and lower socioeconomic status than would be expected from regional demographic statistics. At age 10 years, 4,724 children were included. Children who participated at follow-up were more often of Dutch nationality, had lower Child Behavior Checklist (CBCL) total problems scores ($p < .001$), and had older and more highly educated mothers.[24] Study protocols were approved by the local ethics committee.

Independent replications were performed in the Netherlands Twin Registry (NTR; n = 4,402) and the Swedish Twin Study of Child and Adolescent Development (TCHAD; n = 1,089) cohorts. Both are twin cohorts, nationally representative with respect to socioeconomic status and ethnicities,[25,26] which aim to explore the genetic and environmental influences on cognitive function, psychopathology, and well-being during development. From each twin pair, one twin was randomly selected.

Additional replication was conducted in a clinical sample of children aged 6 to 11 years of age (n = 1,933) who were referred to one of three child and adolescent mental health services in the greater metropolitan area of Rotterdam, the Netherlands. Sampling took place in 2011 for a period of 9 months. This sample is representative of the clinical population in this study base, and has previously been used for clinical validation of the Dutch CBCL.[27]

Measures

DBPs were consistently assessed with the CBCL in all samples, a widely used reliable and valid measure for behavioral problems.[28] The CBCL was completed by the primary caregiver, principally the mother (Table 1). In the Generation R sample, the CBCL/1.5-5 was used at the first time point when most children (58%) were less than 6 years of age, whereas the remaining children were either 6 (38%) or 7 (3%) years of age. In the next examination, the CBCL/6-18 was used, which was also used in the NTR and clinical samples. TCHAD used the CBCL/4-18, an earlier version of the CBCL.

The items included in the FMM analyses were part of the Aggressive Behavior scale of the CBCL/1.5-5, and the Aggressive Behavior and Rule-Breaking Behavior scales of the CBCL/6-18. Items were selected on clinical relevance for measuring DBPs using the following three predefined criteria. Items were not included if (a) they did not reflect problem behavior (e.g., "prefers being with older kids"); (b) were more indicative of behavior problems or disorders other than DBPs (e.g., "can't stand waiting, wants

TABLE 1 Demographic Characteristics of the Population-Based Samples

Characteristic	Generation R 6 Years, n = 6,209	Generation R 10 Years, n = 4,724	NTR, n = 4,402	TCHAD, n = 1,089
Child characteristics				
Gender, % girls	49.7	50.3	49.6	49.2
Age, y, mean (SD)	6.03 (0.42)	9.70 (0.23)	9.83 (0.42)	8.68 (0.47)
Ethnicity, %				
Western	70.4	74.9	96.3	97.8
Nonwestern	29.9	25.1	3.8	2.1
CBCL, % completed by mother	92.3	97.5	100.0	100.0
DBPs score, median (IQR)	4.00 (6.00)	2.00 (5.00)	3.00 (7.00)	3.00 (6.00)
Maternal characteristics				
Educational level, %				
High	51.8	55.2	31.4	38.0
Medium	41.1	39.8	65.4	41.9
Low	7.1	5.0	3.3	19.5

Note: CBCL = Child Behavior Checklist; DBPs = disruptive behavior problems; IQR = interquartile range; NTR = Netherlands Twin Registry; TCHAD = Swedish Twin Study of Child and Adolescent Development.

everything now" or "demands must be met immediately," which are attention-deficit/hyperactivity disorder [ADHD] symptoms); or (c) were endorsed infrequently due to the child's young age (e.g., "drinks alcohol without parents' approval"). The items included in the analyses are listed in Table S1 (available online). They reflect the whole severity range of DBPs (e.g. "hurting animals/people" on the severe end, and "disobedience" on the less severe end of the spectrum), which is a requirement for FMMs.[22]

At age 10 years, we included a validated mother-reported measure for callous traits[29] to incorporate into the cross-lagged model in the Generation R sample. The items are listed in Table S1 (available online).

Statistical Analyses

In principle, two major techniques describe the variability among observed correlated symptoms in terms of unobserved (latent) variables. First, categorical latent variables explain the variability in terms of differences between two or more discrete subgroups (classes) of individuals. Second, dimensional latent variables explain the variability in terms of one or more dimensions on a linear severity scale. The first are used in latent class analysis, and the second in factor analysis. FMMs allow for the presence of both categorical and dimensional latent variables to describe symptoms patterns. This approach is therefore most appropriate for examining whether DBPs occur as dimensional phenomena or as distinct subgroups.[22,23,30]

Before fitting the FMM, it is necessary to conduct exploratory factor analyses (EFA) to empirically assess the dimensional structures of DBPs. This was done in the Generation R sample at ages 6 and 10 years. Because of the ordinal nature of the data, the weighted least-squares means and variances (WLSMV) estimator and the geomin rotation were used. To determine the number of factors, a clear and pure loading pattern had to be present, between-factor correlations must not indicate large overlap between factors (factor correlations larger than 0.9), and fit indices must indicate good model fit. Hence, exploratory factor solutions were compared with an increasing number of factors. Factor loading cut-offs were at 0.30, and items were assigned to the factor with the highest factor loading. In the replication samples, confirmatory factor analyses (CFA) were conducted using the factor structure found in the Generation R sample at 10 years. Model fit for all factor analyses was judged to be good when the following three fit metrics were met: root mean square error of approximation (RMSEA) <0.05; Comparative Fit Index (CFI) >0.95; and Tucker Lewis Index (TLI) >0.95. Fit was deemed acceptable when RMSEA was <0.08, CFI >0.90, and TLI >0.90.

Subsequently, FMMs were fitted on the item selections derived from the EFA to represent dimensions of DBPs. This was done as opposed to fitting a single FMM of all items allowing multiple factors within each class, as mixture models with a complex multi-factor structure within class are often highly unstable.[22] FMMs with an increasing number of classes were fitted to examine whether subgroups of DBPs could be empirically distinguished. Conversely, first fitting categorical and subsequently dimensional latent variables yields unstable results.[22] We allowed factor variances, factor loadings, and item thresholds to vary across classes because it is unlikely that variances are equal for high-scoring and low-scoring children.[30] The primary fit index was the Bayesian Information Criterion (BIC), as it has been shown to be the most reliable fit index for mixture models.[31] A lower BIC indicates a better model fit. The entropy statistic was given to measure the accuracy with which each individual can be classified into a latent class, with values above 0.80 indicating adequate classification.[32] This FMM approach conforms with methodological recommendations[22,30] and is similar to those in other recent studies that have examined the dimensional and/or categorical nature of psychotic symptoms,[23] autistic traits,[33] and ADHD.[34]

Next, we conducted a cross-lagged model in the Generation R sample. By regressing each variable in the model on all of the variables that precede it in time, longitudinal associations between the DBPs dimensions could be estimated. Thus, all associations are mutually adjusted. All analyses were performed in Mplus version 7.[35]

RESULTS

Demographic Characteristics

All three population-based samples were comparable in terms of demographic characteristics (Table 1). However, the Generation R sample included more children of nonwestern ethnic backgrounds. From the clinical sample, 1,933 children were included (mean age 8.35 ± 1.64 years, 34.8% female).

Dimensions and Subgroups of DBPs

EFA were fitted in the Generation R sample. At age 6 years, a three-factor model provided the best solution (Table 2). Fit indices supported this model, with RMSEA = 0.031, CFI = 0.991, and TLI = 0.984. The three factors were physical aggression, irritability, and oppositional behavior, respectively.

At age 10 years (using the CBCL/6-18), a four-factor solution had the clearest loading pattern (Table 3). Fit indices supported this model, with RMSEA = 0.016, CFI = 0.994, and TLI = 0.990. In this wave, the factors were physical aggression, irritability, disobedient behavior, and rule-breaking behavior, respectively.

The dimensional structure found at age 10 years in the Generation R sample was tested in the population-based and clinical samples using CFA. In all three replication samples, the four-factor model fitted the data well. The fit indices for the four-factor solution in the NTR sample were RMSEA = 0.029, CFI = 0.981, and TLI = 0.978. In the TCHAD sample, the fit indices were RMSEA = 0.026, CFI = 0.985, and TLI = 0.983. In the clinical sample, the fit indices for the four-factor solution were RMSEA = 0.056, CFI = 0.965, and TLI = 0.961.

Next, FMM were fitted; fit indices are presented in Tables S2 through S6 (available online). In the Generation R sample at age 6 years, the one-class solution had the best fit for the physical aggression dimension (Table S2, available online). Fitting higher-class models of this dimension produced unreliable results, which was likely because resolution in the data was too low to reliably detect multiple classes in this sample. For both the irritability and oppositional behavior dimensions, two-class models had the best fit. However, in both cases, the entropy statistic was low (0.42 and 0.21, respectively), which indicated substantial overlap between the two classes.

TABLE 2 Factor Loadings for the Three-Factor Solution Exploratory Factor Analysis in the Generation R Sample at Age 6 Years

Behavior	Physical Aggression	Irritability	Oppositional Behavior
Destroys things of others	0.41	−0.03	0.39
Fights	0.78	−0.01	0.02
Hits	0.83	0.01	0.06
Hurts animals/people	0.47	0.06	0.19
Physically attacks	0.75	0.18	−0.01
Easily frustrated	0.01	0.48	0.22
Angry moods	0.03	0.69	0.18
Screams	0.21	0.43	0.22
Stubborn or irritable	−0.07	0.62	0.34
Temper tantrums	0.17	0.73	−0.01
Defiant	0.14	0.05	0.55
Disobedient	−0.01	0.01	0.74
Doesn't feel guilty	0.13	−0.05	0.65
Punishment doesn't change behavior	0.10	0.04	0.70
Uncooperative	0.00	0.23	0.60

Note: The shading illustrates for which particular factor an item has the highest factor loading.

In the Generation R sample at age 10 years, BIC and entropy values did not support the need for a second class; therefore, the one-class solutions were selected as the preferred model for all four DBPs dimensions (Table S3, available online). The difference in BIC between one-class and two-class solutions for disobedient behavior was small, indicating that a second class was not needed to describe the data. This was consistent with the low entropy of the two-class solution (0.63), thus showing substantial overlap between the two classes. Therefore, the one-class solution was determined as the preferred model. In sensitivity analyses, we repeated our EFA and FMM analyses in the Generation R Study using all items of the CBCL Externalizing Problems scales and obtained identical results (Supplemental Material; Tables S7–S9, available online).

Subsequently, FMMs were fitted in the replication samples using the dimensional structure derived from CFA. In both population-based samples, the FMM analyses yielded results similar to those obtained in the Generation R sample. In the NTR sample, one-class solutions had the best model fit for physical aggression, disobedient behavior, and rule-breaking behavior (Table S4, available online). For irritability, a two-class solution had a better fit. However, the entropy statistic was 0.33 and the difference in BIC between one and two classes was small. In the TCHAD sample, one-class solutions had the best model fit for all four DBPs dimensions (Table S5, available online).

In the clinical sample, one-class solutions fit the data best for physical aggression and rule-breaking behavior (Table S6, available online). For disobedient behavior, a two-class solution would have been selected based on the lowest BIC, but the entropy was low at 0.52. For irritability, the two-class solution had the best model fit and the entropy statistic was acceptable (0.85), with one class containing 3.73% of children. This class included children with, on average, higher scores on all irritability items (Figure S1, available online).

Cross-Lagged Model of DBPs Dimensions

Using the Generation R sample, we examined cross-sectional and longitudinal associations between DBPs dimensions in a cross-lagged model (Figure 1). The cross-sectional standardized correlations were all substantial, ranging from 0.49 to 0.63 at age 6 years and from 0.18 to 0.59 at age 10 years. The cross-sectional associations between rule-breaking

TABLE 3 Factor Loadings for the Four-Factor Solution Exploratory Factor Analysis in the Generation R Sample at Age 10 Years

Behavior	Physical Aggression	Irritability	Disobedient Behavior	Rule-Breaking Behavior
Argues	0.62	0.32	0.04	−0.12
Mean to others	0.87	−0.01	−0.01	−0.04
Doesn't feel guilty	0.32	0.14	0.26	0.11
Fights	0.67	0.06	0.05	0.02
Physically attacks	0.59	0.23	−0.15	0.15
Swears	0.39	0.22	0.15	0.05
Teases	0.85	−0.07	0.06	−0.01
Threatens	0.74	−0.06	0.02	0.15
Screams	0.40	0.52	0.01	−0.01
Stubborn or irritable	0.04	0.69	0.25	−0.03
Mood changes	−0.02	0.73	0.09	0.09
Temper tantrums	0.30	0.62	0.00	0.06
Disobedient at home	−0.03	0.25	0.80	−0.03
Disobedient at school	0.15	−0.07	0.73	0.04
Breaks rules[a]	0.10	0.10	0.69	0.09
Lying or cheating	0.19	0.06	0.36	0.25
Destroys his/her own things	0.01	0.29	−0.03	0.78
Destroys things of others	0.09	0.23	−0.02	0.78
Runs away	−0.05	0.24	0.05	0.50
Sets fires	0.02	−0.15	0.19	0.55
Steals at home	0.02	−0.01	0.19	0.69
Steals outside home	−0.01	−0.13	0.14	0.68

Note: The shading illustrates for which particular factor an item has the highest factor loading.
[a]*This item was not part of the Child Behavior Checklist (CBCL)/4-18 (which was used in the Swedish Twin Study on Child and Adolescent Development [TCHAD]), an older version of the CBCL/6-18.*

behavior and other DBPs dimensions at 10 years were all significantly lower (range, 0.18–0.36) than other cross-sectional associations, as can be seen from the nonoverlapping confidence intervals. For example, the association between physical aggression and irritability was 0.59 (95% CI = 0.57–0.62).

Longitudinally, a significant within-trait across-time association was observed for irritability (path coefficient = 0.24, 95% CI = 0.20–0.28). Similarly, age 6 oppositional behavior was associated with disobedient behavior at age 10 (path coefficient = 0.26, 95% CI = 0.22–0.30). All age 6 DBPs dimensions were significantly associated with physical aggression at 10 years (range, 0.08–0.16). Both irritability and oppositional behavior, but not physical aggression, were predictive of later irritability problems (range, 0.09–0.24). Only physical aggression and oppositional behavior, and not irritability, predicted rule-breaking behavior (range, 0.07–0.09) and callous traits (range, 0.04–0.16) at age 10.

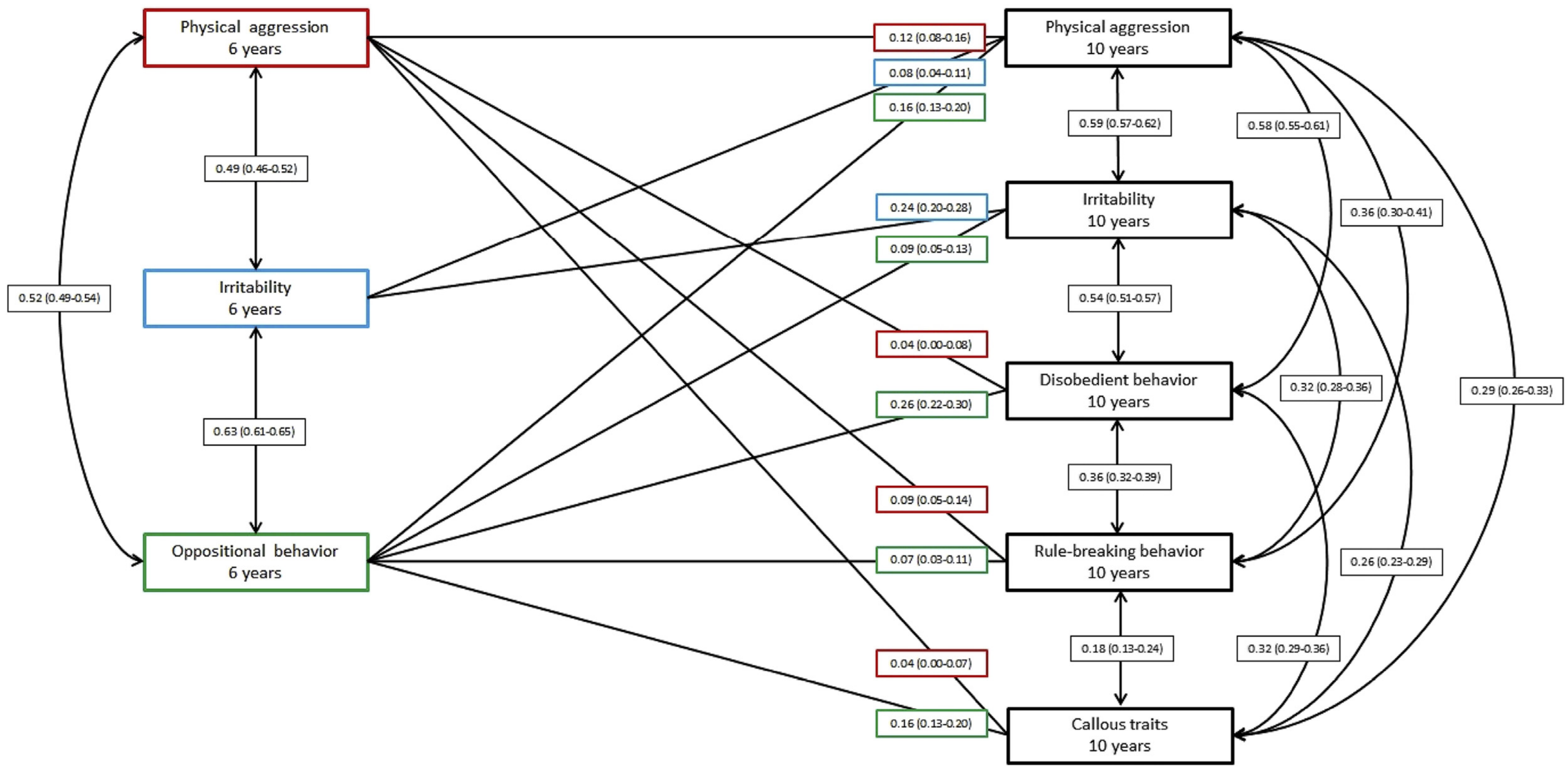
Post hoc cross-lagged analyses separate for boys and girls demonstrated that in both sexes oppositional behavior was predictive of all DBPs dimensions at age 10 years, including callous traits (Figures S2–S3, available online). Similar across both sexes, irritability did not predict rule-breaking behavior. Irritability, however, was associated with later physical aggression in both girls and boys.

DISCUSSION

The present findings suggest that childhood DBPs can best be understood as multiple correlated dimensions rather than as a mixture of distinct categorical subgroups, as indicated by the FMM results in both the population-based and clinical samples. The observed dimensions included physical aggression, irritability, oppositional/disobedient behavior, and rule-breaking behavior. All dimensions demonstrated significant developmental continuity. Interestingly, we also found longitudinal interrelations between most dimensions; for example, all DBPs dimensions at age 6 years, including irritability, predicted later physical aggression symptoms.

These findings add to the existing literature on the multidimensionality of CD and ODD, where particularly the ODD subdomains of irritability and oppositionality have received considerable recognition in the past years.[13,14] Instead of providing another new approach to the classification of DBPs, the present study expands these investigations by assessing symptoms across the spectrum of

FIGURE 1 Cross-lagged model of cross-sectional and longitudinal associations between disruptive behavior problems (DBPs) dimensions in the Generation R sample. *Note:* Significant coefficients with 95% CIs are displayed.



childhood DBPs in both population-based and clinically referred samples. Notably, we found that irritability was a distinct dimension on the broader DBPs spectrum with strong cross-sectional and longitudinal associations with other DBPs dimensions. We therefore argue that irritability is not only a key component of ODD, but must be defined and assessed in all children with DBPs. This is in line with some previous studies that suggested that emotion regulation problems are involved in the development of other aspects of DBPs, such as physical aggression,[20] inattention/impulsivity,[36,37] and rule-breaking.[38] Other dimensions observed in this study include physical aggression, oppositional/disobedient behavior, and rule-breaking behavior, which are comparable to previous findings.[6,7] The importance of recognizing these distinct dimensions of DBPs has recently been highlighted in a review by Tremblay,[21] and our study offers strong empirical support for improved diagnostic accuracy of DBPs along these dimensions. It is interesting to observe that in the present study, these above-mentioned dimensions were all considerably correlated with the irritability dimension, underscoring the value of recognizing emotion regulation problems in youth presenting with a broad range of DBPs.[36,38]

The present cross-lagged model demonstrated significant developmental continuity for all DBPs dimensions, confirming other study findings.[7,13,15,18] However, heterotypic associations (behavioral manifestations that change over time) were also found. Physical aggression at age 6 years predicted disobedience, rule-breaking behavior, and callousness at age 10 years, predominantly in boys, which is well known from previous research.[7,9,13] Interestingly, we found that in both girls and boys, irritability predicted later physical aggression and not rule-breaking behavior, and that oppositional behavior predicted all later DBPs dimensions, including irritability. A previous study found evidence for distinct developmental pathways of irritable and oppositional behavior,[15] whereas others have shown that irritability is associated with conduct problems, both cross-sectionally[39] and longitudinally.[18] The present longitudinal results suggest that irritability is a distinct but not separate dimension in the network of DBPs.[20] More longitudinal replication studies across childhood to adolescence are certainly needed. Finally, the path coefficients from the age 6 dimensions to rule-breaking behavior at age 10 were relatively small, which was the case for both sexes. This probably reflects that rule-breaking behavior is not very prevalent at younger ages. Indeed, studies have shown that rule-breaking behavior is strongly associated with adolescent-onset disruptive behavior,[7,21] and might therefore not be predicted well by childhood characteristics.[9,12]

This study used an FMM approach, which is appropriate for assessing heterogeneity of complex psychiatric traits.[22,23,30] Our findings suggested that one-class models were the best solution for all DBP dimensions, and this was replicated in all population-based samples, which is crucial for maximum-likelihood–based methods such as FMMs.[22,30] These results were similar in the clinical sample. However, fit indices demonstrated that a small class might be distinguished within the clinical sample, containing 3.73% of children with higher irritability scores. It is possible that irritability presents more frequently and more severely in clinical than in population-based samples.[17,36] Taken together, we found no empirical support for the existence of distinct subgroups of DBPs. Instead, DBPs can be best understood as a dimensional construct on a continuum in the population.[6,8,11,21,38,40] This has important clinical consequences, as empirical subgroups do not seem to be a useful basis for diagnosis of DBPs; hence, treatment decisions will need to be made based on a conventional cut-off on the dimensional severity scale of DBPs. Thus, future studies exploring the development and heterogeneity of DBPs should move beyond categorical approaches (e.g., latent class analysis) and integrate the known multidimensionality of DBPs.

This study extends more than 30 years of research on the heterogeneity of DBPs. Important progress has been made with various approaches, such as the work on life-course–persistent DBPs by Moffitt *et al.*,[9,12] developmental continuities of DBPs using data from the Great Smoky Mountains Study,[10,18] or the role of temperamental antecedents of DBPs.[20] Taken together, these findings will help us to better understand the heterogeneity of DBPs and hence to refine psychiatric taxonomies and to improve diagnosis.[40] Another important task for our diagnostic systems would be to address the etiological heterogeneity present in individuals with the same diagnosis, and efforts such as the Research Domain Criteria initiative are increasingly applied to DBPs.[41] For example, it would be interesting to disentangle the neurobiological etiology of DBPs based on its multidimensionality, such as irritability or physical aggression. This would extend recent work demonstrating that psychopathic traits were associated with different neurodevelopmental dysfunctions compared to CD, such as poorer affect reactivity versus impaired reward-based decision making.[42] Likewise, large-scale genome-wide association projects could benefit from taking into account the multidimensionality of DBPs, as unique genetic factors might be in play for different dimensions.[43] Arguably, treatment studies targeting specific DBPs dimensions are even more needed. It is surprising that these are lacking, as irritability could be a potential target for both psychological and pharmacological treatments, which may be of high priority given that treatment success for DBPs has remained limited.[38,44]

This study benefitted from the use of data from three prospective population-based cohorts and a large clinical sample, and the broad and developmentally sensitive assessment of DBPs. However, our study had several limitations. First, DBPs were maternally reported using the CBCL, which might have introduced single reporter bias, and some of the findings could partly be explained by common-method variance. Second, the CBCL lacks items that measure callous-unemotional traits, which could possibly explain why a distinct dimension of these symptoms was not found. To address this limitation, we included a separate measure on callous traits in the cross-lagged model to assess the longitudinal associations between DBP dimensions and callousness. Third, there was attrition in the longitudinal cohort studies. However, as attrition might affect prevalence estimates, it is less likely to influence associations between traits.[45]

In conclusion, this study builds upon a large body of literature investigating the heterogeneity of DBPs. The present findings provide little evidence for the presence of distinct subgroups, but instead support the multidimensionality of childhood DBPs. Notably, we found that irritability was a distinct dimension on the broader spectrum of DBPs, and that irritability predicted later physical aggression symptoms. Future studies should investigate the shared and unique mechanisms underlying the complex multidimensionality of DBPs, and specifically the role of irritability in the development of other DBPs. Herein lies an important opportunity for genetic and neuroimaging studies. &

CG Clinical Guidance

- Childhood disruptive behavior problems can best be conceptualized as a complex multidimensional phenotype. This well-powered study provides no empirical support for the existence of distinct subgroups of children with specific patterns of disruptive behavior problems.
- Clinical classification of disruptive behavior problems will need to be based on cut-offs that are determined by (data-driven) convention, as empirical classification strategies will likely remain unsuccessful.
- This study has corroborated many previous study findings by showing the existence of discrete dimensions of physical aggression and rule-breaking behavior.
- Irritability is a distinct dimension on the broader spectrum of disruptive behavior predicting later physical aggression. This underscores the value of recognizing emotion regulation problems in youth presenting with a broad range of disruptive behavior problems.

Accepted May 31, 2017.

Dr. Bolhuis and Mr. Van der Ende and Drs. Jaddoe, Verhulst, and Tiemeier are with Erasmus Medical Center-Sophia Children's Hospital, Rotterdam, the Netherlands. Dr. Lubke is with University of Notre Dame, Notre Dame, IN. Drs. Lubke, Bartels, van Beijsterveldt and Boomsma are with Vrije Universiteit, Amsterdam, the Netherlands. Drs. Lichtenstein and Larsson are with Karolinska Institutet, Stockholm, Sweden. Dr. Larsson is also with Örebro University, Örebro, Sweden. Dr. Kushner is with Erasmus Medical Center, Rotterdam, the Netherlands.

This study has received funding from the European Union Seventh Framework Program (FP7/2007-2013): ACTION: Aggression in Children: Unravelling gene-environment interplay to inform Treatment and InterventiON strategies (grant number 602768). The Generation R Study data collection is made possible by financial support from the Erasmus Medical Centre, Rotterdam; Erasmus University Rotterdam; The Netherlands Organization for Health Research and Development (ZonMw). The Netherlands Twin Registry (NTR) is supported by the Netherlands Organisation for Scientific Research (NWO 480-04-004; ZonMw 912-10-020; NWO 480-15-001). The Swedish Twin Study on Child and Adolescent Development (TCHAD) is supported by the Swedish Research Council for Health, Working Life and Welfare and the Swedish Research Council. Further support was obtained from the Netherlands Organization for Scientific Research: NWO-grant 016.VICI.170.200 to H.T.

Dr. Lubke and Mr. van der Ende served as the statistical experts for this research.

The Generation R Study is conducted by the Erasmus Medical Centre (Rotterdam) in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam; the Municipal Health Service Rotterdam area, Rotterdam; the Rotterdam Homecare Foundation, Rotterdam; and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond, Rotterdam. The authors gratefully acknowledge the contribution of all children and parents, general practitioners, hospitals, midwives, and pharmacies involved in the different cohorts used in the present study.

Disclosure: Dr. Larsson has served as a speaker for Eli Lilly and Co. and Shire and has received research grants from Shire. Dr. Verhulst has received remuneration as distributor of the Dutch translations of the Achenbach System of Empirically Based Assessment (ASEBA). Drs. Bolhuis, Lubke, Bartels, van Beijsterveldt, Lichtenstein, Jaddoe, Kushner, Boomsma, Tiemeier, and Mr. van der Ende report no biomedical financial interests or potential conflicts of interest.

Correspondence to Henning Tiemeier, MD, PhD, Department of Child and Adolescent Psychiatry, Erasmus MC, P.O. Box 2060, 3000CB Rotterdam, the Netherlands; e-mail: h.tiemeier@erasmusmc.nl

http://dx.doi.org/10.1016/j.jaac.2017.05.019

REFERENCES

1. Rutter M, Shaffer D. DSM-III. A step forward or back in terms of the classification of child psychiatric disorders? J Am Acad Child Psychiatry. 1980;19:371-394.
2. Achenbach TM. DSM-III in light of empirical research on the classification of child psychopathology. J Am Acad Child Psychiatry. 1980;19:395-412.
3. Lahey BB, Loeber R, Quay HC, Frick PJ, Grimm J. Oppositional defiant and conduct disorders: issues to be resolved for DSM-IV. J Am Acad Child Adolesc Psychiatry. 1992;31:539-546.
4. Moffitt TE, Arseneault L, Jaffee SR, *et al*. Research review: DSM-V conduct disorder: research needs for an evidence base. J Child Psychol Psychiatry. 2008;49:3-33.
5. Pardini DA, Frick PJ, Moffitt TE. Building an evidence base for DSM-5 conceptualizations of oppositional defiant disorder and conduct disorder: introduction to the special section. J Abnorm Psychol. 2010;119:683-688.
6. Frick PJ, Lahey BB, Loeber R, *et al*. Oppositional defiant disorder and conduct disorder—a meta-analytic review of factor-analyses and cross-validation in a clinic sample. Clin Psychol Rev. 1993;13:319-340.
7. Bongers IL, Koot HM, van der Ende J, Verhulst FC. Developmental trajectories of externalizing behaviors in childhood and adolescence. Child Dev. 2004;75:1523-1537.
8. Lahey BB, Waldman ID. Annual research review: phenotypic and causal structure of conduct disorder in the broader context of prevalent forms of psychopathology. J Child Psychol Psychiatry. 2012;53:536-557.
9. Odgers CL, Moffitt TE, Broadbent JM, *et al*. Female and male antisocial trajectories: from childhood origins to adult outcomes. Dev Psychopathol. 2008;20:673-716.
10. Rowe R, Maughan B, Pickles A, Costello EJ, Angold A. The relationship between DSM-IV oppositional defiant disorder and conduct disorder: findings from the Great Smoky Mountains Study. J Child Psychol Psychiatry. 2002;43:365-373.
11. Bezdjian S, Krueger RF, Derringer J, Malone S, McGue M, Iacono WG. The structure of DSM-IV ADHD, ODD, and CD criteria in adolescent boys: a hierarchical approach. Psychiatry Res. 2011;188:411-421.
12. Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. Psychol Rev. 1993;100:674-701.

13. Stringaris A, Goodman R. Longitudinal outcome of youth oppositionality: irritable, headstrong, and hurtful behaviors have distinctive predictions. J Am Acad Child Adolesc Psychiatry. 2009;48:404-412.
14. Herzhoff K, Tackett JL. Subfactors of oppositional defiant disorder: converging evidence from structural and latent class analyses. J Child Psychol Psychiatry. 2016;57:18-29.
15. Whelan YM, Stringaris A, Maughan B, Barker ED. Developmental continuity of oppositional defiant disorder subdimensions at ages 8, 10, and 13 years and their distinct psychiatric outcomes at age 16 years. J Am Acad Child Adolesc Psychiatry. 2013;52:961-969.
16. Stringaris A, Zavos H, Leibenluft E, Maughan B, Eley TC. Adolescent irritability: phenotypic associations and genetic links with depressed mood. Am J Psychiatry. 2012;169:47-54.
17. Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E, Stringaris A. The status of irritability in psychiatry: a conceptual and quantitative review. J Am Acad Child Adolesc Psychiatry. 2016;55:556-570.
18. Rowe R, Costello EJ, Angold A, Copeland WE, Maughan B. Developmental pathways in oppositional defiant disorder and conduct disorder. J Abnorm Psychol. 2010;119:726-738.
19. Copeland WE, Shanahan L, Costello EJ, Angold A. Childhood and adolescent psychiatric disorders as predictors of young adult disorders. Arch Gen Psychiatry. 2009;66:764-772.
20. Wakschlag LS, Tolan PH, Leventhal BL. Research Review: 'Ain't misbehavin': towards a developmentally-specified nosology for preschool disruptive behavior. J Child Psychol Psychiatry. 2010;51:3-22.
21. Tremblay RE. Developmental origins of disruptive behaviour problems: the 'original sin' hypothesis, epigenetics and their consequences for prevention. J Child Psychol Psychiatry. 2010;51:341-367.
22. Lubke GH, Miller PJ. Does nature have joints worth carving? A discussion of taxometrics, model-based clustering and latent variable mixture modeling. Psychol Med. 2015;45:705-715.
23. Miettunen J, Nordstrom T, Kaakinen M, Ahmed AO. Latent variable mixture modeling in psychiatric research—a review and application. Psychol Med. 2016;46:457-467.
24. Kooijman MN, Kruithof CJ, van Duijn CM, *et al*. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31:1243-1264.
25. Lichtenstein P, Tuvblad C, Larsson H, Carlstrom E. The Swedish Twin study of CHild and Adolescent Development: the TCHAD study. Twin Res Hum Genet. 2007;10:67-73.
26. van Beijsterveldt CE, Groen-Blokhuis M, Hottenga JJ, *et al*. The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Res Hum Genet. 2013;16:252-267.
27. Verhulst FC, Van der Ende J. Handleiding ASEBA Vragenlijsten voor leeftijden 6 tot en met 18 jaar. Rotterdam, Netherlands: ASEBA Nederland; 2013.
28. Achenbach TM, Rescorla LA. Manual for the ASEBA preschool forms and profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families; 2000.
29. Pardini D, Obradovic J, Loeber R. Interpersonal callousness, hyperactivity/impulsivity, inattention, and conduct problems as precursors to delinquency persistence in boys: a comparison of three grade-based cohorts. J Clin Child Adolesc Psychol. 2006;35:46-59.
30. Lubke GH, Muthen B. Investigating population heterogeneity with factor mixture models. Psychol Methods. 2005;10:21-39.
31. Nylund KL, Asparoutiov T, Muthen BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equat Modeling. 2007;14:535-569.
32. Clark SL, Muthén B. Relating latent class analysis results to variables not included in the analysis. Available at: https://www.statmodel.com/download/relatinglca.pdf2009. Accessed May 5, 2017.
33. Grove R, Baillie A, Allison C, Baron-Cohen S, Hoekstra RA. Exploring the quantitative nature of empathy, systemising and autistic traits using factor mixture modelling. Br J Psychiatry. 2015;207:400-406.
34. Lubke GH, Hudziak JJ, Derks EM, van Bijsterveldt TC, Boomsma DI. Maternal ratings of attention problems in ADHD: evidence for the existence of a continuum. J Am Acad Child Adolesc Psychiatry. 2009;48:1085-1093.
35. Muthén BO, Muthén LK. Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén and Muthén; 1998-2012.
36. Wakschlag LS, Choi SW, Carter AS, *et al*. Defining the developmental parameters of temper loss in early childhood: implications for developmental psychopathology. J Child Psychol Psychiatry. 2012;53:1099-1108.
37. Shaw P, Stringaris A, Nigg J, Leibenluft E. Emotion dysregulation in attention deficit hyperactivity disorder. Am J Psychiatry. 2014;171:276-293.
38. Hill J, Maughan B. Conceptual issues and empirical challenges in the disruptive behavior disorders. In: Thapar A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor EA, eds. Rutter's Child and Adolescent Psychiatry, Sixth ed. Chichester, West Sussex: John Wiley and Sons; 2015:41-52.
39. Stringaris A, Goodman R. Three dimensions of oppositionality in youth. J Child Psychol Psychiatry. 2009;50:216-223.
40. Coghill D, Sonuga-Barke EJS. Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders—implications of recent empirical study. J Child Psychol Psychiatry. 2012;53:469-489.
41. Blair RJ, White SF, Meffert H, Hwang S. Disruptive behavior disorders: taking an RDoC(ish) approach. Curr Top Behav Neurosci. 2014;16:319-336.
42. Alegria AA, Radua J, Rubia K. Meta-analysis of fMRI studies of disruptive behavior disorders. Am J Psychiatry. 2016;173:1119-1130.
43. Viding E, Frick PJ, Plomin R. Aetiology of the relationship between callous-unemotional traits and conduct problems in childhood. Br J Psychiatry Suppl. 2007;190:S33-S38.
44. Bakker MJ, Greven CU, Buitelaar JK, Glennon JC. Practitioner review: psychological treatments for children and adolescents with conduct disorder problems—a systematic review and meta-analysis. J Child Psychol Psychiatry. 2017;58:4-18.
45. Wolke D, Waylen A, Samara M, *et al*. Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. Br J Psychiatry. 2009;195:249-256.