Evaluation of the Strengths and Difficulties Questionnaire - Dysregulation Profile (SDQ-DP)

Marike H. F. Deutz, MSc[1,2], **Qinxin Shi, MEd**[2], **Helen G. M. Vossen, PhD**[1], **Jorg Huijding, PhD**[1], **Peter Prinzie, PhD**[3], **Maja Dekovi , PhD**[1], **Anneloes L. van Baar, PhD**[1], and **Steven Woltering, PhD**[2]

[1]Department of Child and Adolescent Studies, Utrecht University, the Netherlands [2]Educational Psychology, Texas A&M University, College Station, TX, USA [3]Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, the Netherlands

Abstract

The Dysregulation Profile (DP) has emerged as a measure of concurrent affective, behavioral and cognitive dysregulation, associated with severe psychopathology and poor adjustment. While originally developed with the Child Behavior Checklist, more recently the DP has also been defined on the Strengths and Difficulties Questionnaire (SDQ), mostly with a 5-item, but also a 15-item, SDQ-DP measure. This study evaluated the SDQ-DP by examining its factor structure, measurement invariance, and construct validity. Different SDQ-DP operationalizations were compared. In a US longitudinal community sample ($N = 768$), a bifactor model consisting of a general Dysregulation factor and three specific factors of Emotional Symptoms, Conduct Problems, and Hyperactivity-Inattention fitted best, across three different developmental periods (early childhood, middle childhood, adolescence) and across three different reporters (parents, teachers, youth). Measurement invariance across reporter, gender, and developmental period was demonstrated. These findings indicate that the SDQ-DP, like the CBCL-DP, reflects a broad syndrome of dysregulation that exists in addition to specific syndromes of emotional symptoms, conduct problems, and hyperactivity-inattention. SDQ-DP bifactor scores were strongly related with scores on the 5- and 15-item SDQ-DP measures and similarly concurrently associated with two markers of self-regulation, ego-resiliency and effortful control, and longitudinally with antisocial behavior and disciplinary measures. As reliability, validity and stability was weaker for the SDQ-DP 5-item measure, use of all 15 items is recommended. Advantages of using a bifactor approach are discussed as well as the potential of the SDQ-DP as an easy screening measure of children at risk for developing serious psychopathology.

Keywords

children's emotional and behavioral problems; confirmatory factor analysis; measurement invariance; behavioral assessment; emotional dysregulation; Strengths and Difficulties Questionnaire (SDQ)

Correspondence to: Marike Deutz, Department of Child and Adolescent Studies, Utrecht University, P.O. Box 80.140, 3508 TC Utrecht, the Netherlands. Tel: +31-30-2534804, m.h.f.deutz@uu.nl.

A considerable number of children and adolescents referred for clinical treatment present a complex picture of co-occurrence of affective, behavioral, and cognitive dysregulation, causing significant diagnostic and therapeutic challenges for clinicians. The Dysregulation Profile (DP), based on the Child Behavior Checklist (hence the term CBCL-DP), has emerged as a reliable, valid and relatively simple dimensional measure of this complex phenotype of dysregulation (Althoff, Ayer, Rettew, & Hudziak, 2010; Ayer et al., 2009). The CBCL-DP is not linked to a specific disorder such as attention-deficit/hyperactivity disorder (ADHD), or juvenile bipolar disorder (Diler et al., 2009; McGough et al., 2008). Rather, the DP broadly characterizes dysregulation, which is presupposed to be underlain by deficits in self-regulation as self-regulation is thought to be impaired in all the psychopathological symptom domains measured with the DP (Althoff et al., 2010; Althoff, Verhulst, Rettew, Hudziak, & van der Ende, 2010; Ayer, et al., 2009).

More recently, the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), has also been validated as a measure to capture the dysregulation phenotype (Holtmann, Becker, Banaschewski, Rothenberger, & Roessner, 2011). The SDQ, like the CBCL, is a behavioral screening questionnaire with equivalent forms for parents, teachers and youth self-reports. Although the CBCL and SDQ have been found to correspond well (Goodman & Scott, 1999; Stone, Otten, Engels, Vermulst, & Janssens, 2010), they also differ. The SDQ, in contrast to the CBCL, is freely available online and significantly shorter (25 versus 113 items). The brevity of the SDQ might make it more practical for quick screening or regular short-term monitoring of children's emotional and behavioral problems

Given that the SDQ is increasingly being used in research, and the exponential growth in the past decade of research on childhood emotional and behavioral dysregulation using the CBCL-Dysregulation Profile (see Bellani, Negri, & Brambilla, 2012, and Caro-Cañizares, García-Nieto, & Carballo, 2015, for reviews), it can be expected that the SDQ-Dysregulation Profile will be used in many more studies to come. Research evaluating the structure and psychometric properties of the SDQ-DP however is lacking, leading to a scarcity of evidence for the use of the SDQ-DP as a screening measure either to identify at-risk children in the general population or to identify subgroups of high-risk patients with greater clinical severity (Carballo et al., 2014). The aim of this study is therefore to examine the factor structure, reliability, measurement invariance and validity by examining concurrent associations of the SDQ-DP with measures of self-regulation in early childhood as well as longitudinal outcomes of early childhood SDQ-DP. Furthermore, correspondence between different operationalizations of the SDQ-DP in terms of overlap, reliability and validity will be examined as to advise on the use in research and clinical practice.

Operationalization of the Dysregulation Profile on the CBCL and SDQ

Both the CBCL-DP and SDQ-DP have been operationalized using scores on scales characterized by dysregulation of affect (Anxiety/Depression or Emotional Symptoms), behavior (Aggressive Behavior or Conduct Problems) and cognition (Attention Problems or Hyperactivity-Inattention) (e.g. Althoff et al., 2010; Winsper & Wolke, 2014). For the SDQ-DP, in addition, a short 5-item measure has been developed using stepwise linear discriminant analyses and receiver operating characteristic (ROC) analysis of all 25 SDQ

items (Holtmann et al., 2011). This SDQ-DP 5-item measure is a summed score of 5 items from the Emotional Problems (2 items), Conduct Problems (2 items), and Hyperactivity-Inattention (1 item) scales. It was highly correlated with CBCL-DP scores operationalized as summed *T*-scores of the AAA-scales ($r = .75$; Holtmann et al., 2011). This SDQ-DP 5-item measure has been used in most of the research using the SDQ-DP so far. However low reliability ($a = .52$) has been reported for the 5-item SDQ-DP measure (Holtmann et al., 2011) The use of all 15 items representing emotional symptoms, conduct problems, and hyperactivity-inattention, either summed or within a factor model, might result in a more reliable measure of SDQ-DP. Correspondence between different measures of the SDQ-DP as well as potential differences in validity are therefore examined in this study.

The Factor Structure of the SDQ-Dysregulation Profile

Factor-analytic studies can provide insight into how the DP can best be conceptualized. Recent studies used Confirmatory Factor Analysis (CFA) to examine the factor structure of the CBCL-DP, and showed that a bifactor model consisting of a general DP factor and three specific factors of Anxiety/Depression, Aggression and Attention Problems best described the CBCL-DP (Deutz, Geeraerts, van Baar, Dekoviꞁ, & Prinzie, 2016; Geeraerts et al., 2015). This study aimed to replicate this research as it is important to examine consistency in theoretical conceptualization of DP, regardless of whether the CBCL or the SDQ is used to measure emotional and behavioral symptoms. We tested three competing factor models of SDQ-DP that each conceptualize the DP differently. The simplest model is the one-factor model (Figure 1a), in which symptoms (items) describing emotional, conduct, and hyperactivity-inattention problems all load onto one factor of Dysregulation, representing the idea that the DP is a unidimensional syndrome. In a second-order model (Figure 1b), specific first-order factors represent distinct problems of emotional symptoms, conduct problems and hyperactivity-inattention. A second-order DP factor then accounts for the communalities between the factors, a perspective in line with the idea that dysregulation represents comorbidity (Carlson, 2007). The third model is the bifactor model (Figure 1c), in which symptoms of behavioral and emotional problems load onto one general dysregulation factor, as well as onto specific factors of Emotional Symptoms, Conduct Problems and Hyperactivity-Inattention. While both a one-factor and bifactor model describe dysregulation as one broad syndrome, they differ on whether they suggest that distinguishing between different types of self-regulatory problems is useful (bifactor model) or unnecessary (one-factor model). A bifactor model in general suggests that there might be both shared and nonshared etiological factors and that treatment should be tailored to symptom profile (Martel, von Eye, & Nigg, 2010).

Evaluation of the factor structure of the SDQ-DP is particularly useful because factor scores have several advantages over sum scores. While all items representing symptoms are given equal weight when sum scores are calculated, factor scores are computed based on different weights for different items. Factor scores therefore reflect the fact that some symptoms might be more characteristic of dysregulation than others, corroborating the idea that the DP is more than merely the sum of its components (Boomsma et al., 2006). Factor scores can also reduce measurement error and therefore result in a purer measure of the DP. It is very relevant to examine whether a bifactor approach can also be used to describe the SDQ-DP

given that bifactor models have more recently been rediscovered as an effective approach to examine multidimensionality, i.e. determining whether a set of items (symptoms) reflect a common underlying construct (Reise, 2012). Notwithstanding these advantages, factor scores are less useful in clinical practice as these are based on groups, whereas sum scores can easily be computed for each individual. Furthermore, factor scores are dependent upon characteristics of the dataset whereas individual sum scores are more comparable across studies.

Concurrent Construct and Longitudinal Predictive Validity of the SDQ – DP

The CBCL-DP and SDQ-DP have been associated with a wide range of negative adjustment outcomes such as greater psychiatric comorbidity and functional impairment, reduced psychosocial functioning, worse family functioning, more frequent parental psychiatric history, and sleeping problems (Althoff, et al., 2010; Carballo et al., 2014; Caro-Cañizares, Serrano-Drozdowskyj, Pfang, Baca-García, & Carballo, 2017; Legenbauer, Heiler, Holtmann, Fricke-Oekermann, & Lehmkuhl, 2012; Holtmann et al., 2011). Consequently, both the CBCL- and SDQ-Dysregulation Profile have been described as indices of overall psychological severity and functional impairment (Bellani et al., 2012; Carballo et al., 2014). However, little research used a bifactor approach to examine such relationships of the Dysregulation Profile, while bifactor models come with the major advantage of parsing out specific and overlapping risk factors and outcomes for general dysregulation versus more specific forms of psychopathology. Bifactor models therefore provide a more refined way to examine predictors and outcomes, and can give purer estimates of constructs. This would aid our understanding of the construct of dysregulation. In this study, construct validity was evaluated by examining associations between the DP and two markers of self-regulation: ego-resiliency and effortful control. Whereas the DP has consistently been described as resulting from self-regulatory deficits, only a handful of studies have examined associations between the DP and measures indicative of self-regulation such as inhibition (Geeraerts et al., 2015) and emotion regulation (Legenbauer et al., 2016). The concept of ego-resiliency has its roots in the Ego-Control/Ego-Resiliency Model (Block & Block, 1980), a theoretical model of self-regulation. Ego-control refers to the inhibition/expression of impulses whereas ego-resiliency describes the ability to modulate these impulses flexibly and adaptively. Effortful control can be defined as children's ability to inhibit predominant responses to activate subdominant ones (Kochanska, Murray, & Harlan, 2000). Longitudinal or predictive validity was evaluated by examining associations with negative adjustment outcomes in adolescents, namely antisocial behavior and disciplinary measures.

The Current Study

The objective of this study is to evaluate the Strengths and Difficulties Questionnaire - Dysregulation Profile (SDQ-DP) by: (a) examining the factor structure of the SDQ-DP, (b) as a prerequisite for valid comparisons, examining measurement invariance of the best-fitting model to determine whether the SDQ-DP is similarly defined across reporters (parents, teachers, youth), gender and developmental period (early childhood, middle childhood, adolescence), (c) examining correspondence between the SDQ-DP best-fitting factor model, the 5-item SDQ-DP measure (Holtmann et al., 2011) and the 15-item summed

SDQ-DP score (Winsper & Wolke, 2014), and (d) investigating construct and predictive validity by examining relations with markers of self-regulation and negative adjustment outcomes. Given that a bifactor model best described the CBCL-DP (Deutz et al., 2016; Geeraerts et al., 2015), we expect that a bifactor model also represents SDQ-DP best. This hypothesis is further supported by research demonstrating that a bifactor model can adequately describe the structure of the full SDQ including prosocial and peer problems factors (Caci, Morin & Tran, 2015; Kóbor, Takács, & Urbán, 2013), although no research as yet has been conducted on the factor structure of the SDQ-DP specifically.

Method

Procedure and Participants

The present study is part of a 12-year longitudinal cohort study (2000–2012) called 'Project Achieve' (see Hill & Hughes, 2007), aimed at examining relations between grade retention and academic achievement. Project Achieve was approved by the Research Ethics Board of Texas A&M University. Originally, 1374 children from three different school districts (one urban, two small city districts) in Texas, USA, were targeted, based on several inclusion criteria: (a) scoring below the median of the school district on a state-approved district-administered literacy test, (b) not having received special education, and (c) not having been retained at first grade. The project's focus on children with relatively low academic readiness skills was because these children are known to be at increased risk for the development of emotional and behavioral problems and therefore represent a population of concern. Out of those 1374 eligible participants, parents of 784 children (65%) provided consent across two sequential cohorts in 2000 and 2001 (449 for cohort 1 and 335 for cohort 2). Children with and without consent did not differ on a broad array of variables such as age, gender, ethnicity and socioeconomic status (Hill & Hughes, 2007). Parents, teachers (all waves) and children (only in adolescence) received a monetary reward of $25 for their participation at each wave of data collection.

This study included participants for whom SDQ data (either parent- or teacher-report) was available for at least one of the included waves. This resulted in a final sample of 768 children (52.5% male) that was ethnically diverse (37.5% Hispanic, 34.4% Caucasian, 22.8% African-American, and 5.3% other: e.g. Asian). 476 children (62.5%) were classified as economically disadvantaged based on children's eligibility to receive free or reduced school lunch.

For this study, we used data from three measurement waves representing three distinct developmental periods namely: early childhood (T1, M_{age}= 6.57, SD = 0.39), middle childhood, (T2, M_{age} = 9.57, SD = .039), and adolescence (T3, M_{age} = 13.57, SD = 0.39). At T1, SDQ-data were available for 496 parents (35.4% missing), and 678 teachers (11.7% missing). For 451 children (58.7%) both parent- and teacher-reported data was available at T1, while for 45 children (5.9%) no parent- or teacher-reported SDQ-data was available at T1. At T2, data were available for 446 parents (41.9% missing) and 528 teachers (31.3% missing). For 359 children (46.7%) both teacher- and parent-reported data was available. At T3, data were available for 352 parents (54.2% missing), 437 teachers (43.1% missing), and 505 adolescents (34.2% missing), with 272 children (35.4%) having data for all reporters at

T3. Participants for whom SDQ-data was available at all time points (for at least one reporter per time point: 458 children, 40.4%) did not differ statistically from participants who had missing data on any of the waves (310 children, 59.6%) on gender, SES, or T1 parent- and teacher-reported scores on the Emotional Symptoms, Conduct Problems, and Hyperactivity-Inattention scales of the SDQ as well as the 5- and 15-item SDQ-DP measures. They only differed on ethnicity with Hispanic children being slightly more likely to have missing data on any of the waves.

Instruments

Strengths and Difficulties Questionnaire—The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a brief 25-item behavioral screening questionnaire asking to what extent both positive and negative psychological attributes of the child were true in the past six months, using a 3-point scale (0 = *not true*, 1 = *somewhat true*, 2 = *certainly true*). The SDQ has good psychometric properties (Stone et al., 2010), and has parallel forms for parents, teachers, and children aged 11 and over. In this study, parents and teachers completed age-equivalent forms of the SDQ at all three time points, while youth reported on the SDQ in adolescence (T3) only.

The SDQ consists of five subscales, each consisting of five items: Prosocial Behavior, Hyperactivity-Inattention, Emotional Symptoms, Conduct Problems, and Peer Relationships. In this study, we used items from three subscales to estimate the Dysregulation Profile (DP) factor models: Emotional Symptoms, Conduct Problems and Hyperactivity-Inattention. After reverse coding of positively worded items, Cronbach's α on these subscales ranged from .70 to .80 (mean α = .74) across reporters and time points.

In addition, scores for the SDQ-DP 5-item measure were computed according to the study of Holtmann et al. (2011), by summing scores on five items: two items from the Emotional Symptoms scale (13: often unhappy, down-hearted, or tearful; 8: many worries, often seems worried), two items from the Conduct Problems scale (12: often fights with other children or bullies them, 22: steals from home, school or elsewhere) and one item from the Hyperactivity-Inattention subscale (2: restless, overactive, cannot stay still for long). Cronbach's α's were low, ranging from .54 to .65 across reporters and time points (mean α = .59). A cut-off point of 5 is suggested to identify children exhibiting clinical levels of problems (Holtmann et al., 2011). Using this cut-off, on average 9.40% of children in the study met this criterion (averaged across reporters and developmental periods). The SDQ-DP 15-item measure as used by Winsper and Wolke (2014) was computed by summing scores on the 15 items of the Emotional Symptoms, Conduct Problems and Hyperactivity-Inattention, with Cronbach's α ranging from .80 to .87 across reporters and time points (mean α = .85)

Validity Measures

Measures of ego-resiliency and inhibitory control were only assessed at T1; therefore, examination of construct validity was limited to T1 only. Longitudinal outcomes were assessed at T3.

Ego-resiliency—The measure of ego-resiliency (the ability to express and modulate impulses effectively and adaptively) was derived from a selection of items of the California Child Q-Set (Caspi et al, 1990), filled out by teachers at T1 using a 1–5 Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). In a previous study on this dataset (Kwok, Hughes, & Luo, 2007) exploratory and confirmatory factor analyses resulted in the development of a 7-item ego-resiliency scale consisting of 4 items describing ego-resiliency (resourceful in initiating activities; curious, eager to learn, open; self-reliant, confident; persistent, does not give up easily) and 3 items describing ego-brittleness (becomes rigidly repetitive; falls to pieces under stress; rapid mood shifts, emotionally labile). After reverse-coding the ego-brittleness items, we computed a mean score of these 7 items with higher scores representing higher ego-resiliency. The measure showed good internal consistency (Cronbach's alpha a = .85).

Effortful control—At T1, trained research assistants individually administered tasks from a behavioral battery designed to assess behavioral self-regulation by tapping into the ability to deliberately slow down motor activity (Kochanska, Murray, & Coy, 1997). This study used data from three tasks: Telephone Poles, Stars, and Walk-a-Line. In the Walk-a-Line task, children were asked to walk along a ribbon that was taped onto the floor. In the Telephone poles tasks, children were asked to draw wires (i.e. straight lines) to connect telephone poles for the squirrels to sit on using a ruler. In the Star task children were given a picture of a star and were asked to draw the shape staying in the lines. In each of the tasks children participated in three trials: a baseline trial with no instructions regarding speed, a fast trial in which children were asked to be as fast as possible and the third trial was a slow trial, in which children were asked to slow or inhibit their (gross or fine motor) behaviors. As effortful control is defined as the ability to inhibit a predominant response to perform a subdominant response (Kochanska et al., 2000), differences between the fast trial and the slow trial (in which children had to deliberately slow down motor behavior after preceding instructions to be as fast as possible) were averaged across the three tasks (a = .75) to create the effortful control score. Higher scores indicate higher effortful control.

Longitudinal Construct Validity

Antisocial involvement—At T3, students were interviewed individually and were asked to briefly report by saying yes or no on whether they have been involved in each of four antisocial activities (i.e. been caught by the police, taken part in a fight, destroyed things, and skipped school) during the past year. This 4-item Antisocial Involvement Questionnaire is adapted from an 8-item measure used by Mahoney and Stattin (2000). Data were available for 505 adolescents.

Disciplinary actions—At T3, teachers reported whether any of the five following disciplinary actions occurred for the student: sent to the office for disciplinary reasons, assigned to in-school-suspension, assigned to disciplinary alternative education, judicial placement outside school district and expelled from school (all answered as yes or no). At each time point, scores on the five disciplinary actions were summed. Data were available for 437 adolescents.

Statistical Analyses

All analyses were conducted in Mplus 7.4 (Muthén & Muthén, 2012) using the Weighted Least Squares Means and Variances adjusted estimator (WLSMV) with Delta parameterization to account for categorical symptom ratings and resulting non-normality. First, three competing SDQ-DP factor models (bifactor, second-order, one-factor, see Figure 1a, 1b, 1c) were compared in three developmental periods (early childhood, middle childhood, and adolescence) and across three different reporters (parents, teachers, youth) using Confirmatory Factor Analysis (CFA). The least restricted model in terms of degrees of freedom is the bifactor model (see Figure 1c), in which the 15 items loaded both onto one of three orthogonal first-order factors (Emotional Symptoms, Conduct Problems, Hyperactivity-Inattention), as well as on a general factor of Dysregulation. This model was restricted into a second-order model (Figure 1b), in which the 15 items loaded onto one of three first-order factors which in turn loaded onto a second-order Dysregulation factor. The second-order model is statistically indistinguishable from a three-factor correlated model. Finally, the second-order model was restricted into a one-factor model (Figure 1a), in which all items loaded only onto one Dysregulation factor.

Second, measurement invariance (MI) of the best-fitting model across reporter, gender, and developmental period was tested following recommendations of Muthén and Muthén (2012) for testing MI with categorical (ordinal) indicators using the WLSMV estimator and Delta parameterization. This procedure consists of testing two models. Model 1 was the least restrictive model that tested for configural invariance, in which factor loadings and thresholds were freely estimated, scale factors were fixed at one, and factor means were fixed at zero. Model 2 was the more restricted scalar invariance model constraining factor loadings (metric invariance) and thresholds (scalar invariance) jointly, and in which scale factors were fixed at one in one group and free in the other, and factor means were fixed at zero in one group and free in the other. Because in bifactor models factor indicators load on more than one factor (specific and general factor), and we set the metric of the factors by fixing factor variances to one, testing the metric model separately was not allowed (see Mplus User Guide version 7 page 486, or version 8 page 544).

Finally, correlations between the Dysregulation factor scores of the best-fitting model, and scores on the 5-item and 15-item SDQ-DP measures were computed, to examine overlap. Construct and predictive validity was examined with regression analyses in Mplus, controlling for gender, socio-economic status and three dummy-variables for ethnicity (African-American vs White, Hispanic vs White, Other vs White).

Model fit was evaluated using three primary and widely used fit indices: the Root Mean Square Error of Approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI), at the following thresholds for good fit: RMSEA .05, and CFI/TLI .95 (Cheung & Rensvold, 2002, Hu & Bentler, 1999). Chi-square is reported, but not interpreted as it is nearly always significant in larger samples and/or complex models (Kline, 2006). Changes in RMSEA (RMSEA) and CFI (CFI) are used as the main criterion to define invariance as they are much less sensitive to sample size and more sensitive to a lack of invariance than chi-square-based tests of MI (Meade, Johnson, & Braddy, 2008).

Measurement invariance holds if the changes in fit statistics between Model 1 and 2 are . 015 for RMSEA .01 for CFI (Chen, 2007).

Results

Factor Structure

Factor scores for the early childhood models (both parent- and teacher-reported) could not be saved because in each model, R-Square for one item (10, 18 or 25) could not be computed. This warning was checked and in each model the factor loading of the specific item was restricted to the value closest to the original factor loading until the warning disappeared (e.g. for the parent-reported model this meant that the factor loading of item 25, which was .785, was constrained to .70 thus adding one degree of freedom).

Table 1 presents the results for the CFA analyses. Values for CFI and TLI indicated that the bifactor models generally showed a very good fit whereas RMSEA values were acceptable across developmental period and reporters, while the second-order models fit adequately and the one-factor models showed poor fit overall. Chi Square difference testing for WLSMV estimator was used to statistically compare the nested models, with significant values indicating worse model fit of the more restricted model. From these analyses, it can be concluded that the bifactor SDQ-DP models statistically described the data better than the second-order and one-factor models across developmental periods and regardless of informant.

Factor Loadings

To further evaluate the quality and content of the models, factor loadings for the SDQ-DP bifactor models were inspected. Table 2 presents factor loadings for the bifactor SDQ-DP model at T1 (early childhood), T2 (middle childhood) and T3 (adolescence) using different reporters. Inspection of the standardized factor loadings showed that the DP factor loadings were all positive and statistically significant (all $p < .001$) and moderately high with an average loading of .58. For Emotional Symptoms, all scale-specific loadings were positive and statistically significant ($p < .001$), and moderately high (average = .59). The scale-specific loadings of Conduct Problems and Hyperactivity-Inattention were generally lower than the scales-specific loadings for Emotional Problems, and both scales also had several nonsignificant scale-specific loadings as well as some negative loadings (of which three were statistically significant at $p < .05$).

Within the bifactor model, the five items used to calculate the SDQ-DP 5-item measure (Holtmann et al., 2011) had moderate to high DP-loadings averaging .56 across developmental periods and reporters, with averages ranging from .33 ('many worries') to .76 ('restless, overactive'). However, these items also had fairly high scale-specific loadings (average .53) compared to other items. For example, 'many worries' had an average (across developmental periods and reporters) scale-specific loading of .69 and an average DP-loading of .33, suggesting that this item better describes specific emotional symptoms than general dysregulation.

Table 3 presents the factor loadings of the one-factor, second-order and bifactor models in early childhood (parent-report) side by side, to demonstrate changes in the salience of items from a one-factor model in which all items load on one factor, to grouping the items into specific factors (second-order model) and then adding a general Dysregulation factor (bifactor model). In the second-order model almost all loadings were high and significant, but in the bifactor model, especially the loadings of Conduct Problems and Hyperactivity-Inattention decreased in size. Two of the Hyperactivity-Inattention items ('restless, overactive', 'constantly fidgeting or squirming') and one Conduct Problems item ('generally obedient', reverse-coded) now had nonsignificant loadings, while they had moderate to high DP factor loadings, suggesting that these items more directly predict the underlying general factor of dysregulation.

Measurement Invariance across Reporter, Gender, and Developmental Period

A series of measurement invariance analyses were conducted, examining configural versus scalar invariance in line with recommended procedures for testing MI with categorical indicators using WLSMV estimation and Delta parametrization (Muthén & Muthén, 2012). When needed (because or errors), factor loadings were constrained to values previously identified when testing the factor models. Results for these analyses are reported in Tables s1, s2, and s3 in the online supplemental material. Values of RMSEA and CFI indicated that measurement invariance across reporter, gender and developmental period was demonstrated.

Relations among SDQ-DP Measures

At each time point and for each reporter, (saved) SDQ-DP bifactor scores were most highly correlated with the SDQ-DP 15-item sum scores (mean $r = .92$, range $= .90 – .96$), while they were lower, but still highly correlated with the SDQ-DP 5-item score (mean $r = .78$, range $= .68 – .83$). Stability of all measures was moderate across an 8-year period, and highest for the SDQ-DP bifactor scores ($r = .57 / .38$ for parent- and teacher-report respectively), followed by the SDQ-DP 15-item measure ($r = .55 / .40$) and the SDQ-DP 5-item measure ($r = .39 / .35$). Interrater agreement for all measures was moderate and again highest for the DP bifactor scores ($r = .36$), followed by $r = .35$ for the SDQ-DP 15-item measure and $r = .25$ for the SDQ-DP 5-item measure. Boys consistently had significantly higher DP scores (regardless of operationalization), except for youth self-report when no significant differences emerged (tables for these analyses can be requested from the first author).

Construct and Longitudinal Validity

Construct validity results are presented in Table 4. For all SDQ-DP measures, lower ego-resiliency and effortful control similarly predicted higher DP, for both parent- and teacher-reports. In addition, ego-resiliency also predicted lower teacher-reported specific Emotional Symptoms, Conduct Problems and higher Hyperactivity-Inattention, whereas effortful control predicted lower Emotional Symptoms also.

Longitudinal validity was examined by regressing all factors of the T1 bifactor model, 5-item SDQ-DP and 15-item SDQ-DP (in separate models) on disciplinary measures and

antisocial behavior measured at T3 (for results see also Table 4). Higher levels of DP predicted more disciplinary measures and antisocial behavior, for all operationalizations of the DP and for both parent- and teacher-reports. For the bifactor model, significant associations between the specific factors and the outcomes emerged. For parent-reports, lower Emotional Symptoms predicted higher antisocial behavior, and higher Conduct Problems predicted both antisocial behavior and disciplinary measures. For teacher-reports, lower Emotional Symptoms and Hyperactivity-Inattention predicted lower antisocial behavior, whereas Conduct Problems predicted higher disciplinary measures.

Discussion

The objective of this study was to evaluate the Strengths and Difficulties Questionnaire - Dysregulation Profile (SDQ-DP), measuring a broad syndrome of child and adolescent difficulties in regulating affect, behavior and cognition that can reflect overall psychopathology severity and functional impairment (Holtmann et al., 2011). Specifically, we examined the factor structure, reliability, measurement invariance, and construct and predictive validity of the SDQ-DP. We compared the best-fitting factor model with the previously developed 5 and 15-item SDQ-DP 5-item measures (Holtmann et al., 2011; Winsper & Wolke, 2014) in order to recommend when to use which operationalization.

The results of this study replicate findings of previous studies on the Child Behavior Checklist - Dysregulation Profile (Deutz et al., 2015; Geeraerts et al., 2016), by demonstrating that a bifactor model described the SDQ-DP better than a one-factor or second-order model. This bifactor SDQ-DP model suggests that the DP reflects a broad syndrome of dysregulation that exists in addition to specific syndromes of emotional symptoms, conduct problems, and hyperactivity-inattention. Importantly, the bifactor model best described the data across three different developmental periods (early childhood, middle childhood, adolescence) and across different reporters (parents, teachers, youth). Measurement invariance across reporters, gender and developmental period was also demonstrated, showing that the DP bifactor model is constructed similarly regardless of reporter, gender or developmental period, adding to the generalizability of the findings.

Factor Loadings of the SDQ-DP Bifactor Models

Examination of the factor loadings gives further insight into the meaning of the underlying general Dysregulation factor. Symptoms of hyperactivity and inattention (e.g. 'restless, overactive', 'constantly fidgeting or squirming') seemed to contribute most directly to the underlying factor, while in research on CBCL-DP items concerning mood dysregulation (e.g. 'Stubborn, sullen, or irritable', 'Sudden changes in mood or feelings') seem to describe dysregulation most directly (Geeraerts et al., 2015). The brevity of the SDQ might explain these differences, as the SDQ includes only five key symptoms for emotional, conduct and attention problems each, while symptoms of mood dysregulation and irritability are considered to be transdiagnostic (e.g. Kring, 2008). The only item directly describing emotional lability in the SDQ is the item 'Often has temper tantrums or hot tempers', which loaded fairly high on the Dysregulation bifactor. One might also wonder whether symptoms of hyperactivity and inattention truly form a unique factor. However, research has shown that

Attention Hyperactivity Disorder (ADHD) is under unique genetic influence (Dick, Viken, Kaprio, Pulkkinen, & Rose, 2005), and post-hoc analyses we conducted showed that models in which items of hyperactivity-inattention did not load on a specific factor but rather only directly on the Dysregulation factor, did not fit better. Furthermore, symptoms of hyperactivity and inattention have been previously found to strongly contribute to DP, next to symptoms of mood lability (Althoff et al., 2010a; Geeraerts et al., 2015), showing that the DP reflects not only affective dysregulation, but also behavioral and cognitive dysregulation.

The SDQ-DP bifactor scores were most strongly related with SDQ-DP 15-item scores, and less strong, but still highly related with scores on the most often used SDQ-DP 5-item measure (Holtmann et al., 2011), regardless of reporter and developmental period. Stability and interrater agreement were moderate and highest for the SDQ-DP bifactor scores and lowest for the SDQ-DP 5-item scores. Furthermore, reliability of the SDQ-DP 5-item measure was low in our and previous studies (Holtmann et al., 2011). Inspection of factor loadings showed that the 5 items that make up the SDQ-DP 5-item measure did not necessarily load very highly on the Dysregulation bifactor and/or contribute most directly to the Dysregulation bifactor (as would be indicated by low scale-specific factor loadings), further questioning the validity of the SDQ-DP 5-item measure.

Construct and Longitudinal Predictive Validity

The SDQ-DP was concurrently associated with lower ego-resiliency and lower effortful control, for both teacher- and parent-reports and regardless of operationalization of the SDQ-DP. This evidence for construct validity was rather robust since the self-regulation measures were quite distinct. Effortful control was assessed with behavioral tasks tapping into more cognitive aspects of self-regulation, primarily slowing down motor control. Ego-resiliency was teacher-reported and tapping into emotional self-regulation with items such as 'Rapid mood shifts, emotionally labile'. The DP furthermore predicted more disciplinary measures and antisocial behavior seven years later, behaviors of which dimensions of emotion and self-regulation are thought to be disrupted (Hyde, Shaw, & Hairiri, 2013), demonstrating longitudinal validity (again for both parent- and teacher-report and for all operationalizations of the DP). This study is one of the few so far providing empirical evidence for the notion that the Dysregulation Profile is indeed related to self- and emotion regulation difficulties. Concurrent construct and longitudinal predictive validity of the SDQ-DP bifactor, 5-item and 15-item measures was generally comparable across operationalizations, but associations between the specific factors and correlates were controlled for in the bifactor model, possibly attenuating the strength of these associations. In addition, significant associations between the specific factors of the SDQ-DP bifactor model and the external correlates emerged, with especially the specific Conduct Problems factor predicting additional variance in the longitudinal outcomes. Unexpectedly, ego-resiliency positively predicted specific Hyperactivity-Inattention, which could be the result of cross-over suppression (Paulhus et al., 2004). In the early childhood teacher-reported SDQ-DP bifactor model, items representing symptoms of hyperactivity-inattention contributed most directly to dysregulation, with several scale-specific factor loadings being non-significant. Thus, while ego-resiliency and hyperactivity-inattention are generally negatively related (e.g. Martel & Nigg, 2006), after accounting for general dysregulation, specific hyperactivity-inattention

(for T1 teacher-report primarily defined by being restless and fidgety), might no longer be positively related with ego-resiliency.

Implications and Conclusions

A bifactor model best described the SDQ-DP, conceptualizing the DP as a broad syndrome of dysregulation that exists next to specific problems of emotional symptoms, conduct problems and hyperactivity-inattention. This research adds to the growing notion that different types of emotional and behavioral problems can be (largely) explained by one underlying factor of dysregulation (or general psychopathology, see e.g. Caspi et al., 2014; Patalay et al., 2015). This is consistent with recent findings showing that different psychiatric conditions partly share the same genetic origin (Pettersson, Larsson, & Lichtenstein, 2016). Given increasing consensus on the presence of a general factor of dysregulation/psychopathology already at an early age (e.g., Geeraerts et al., 2015; Olino, Dougherty, Bufferd, Carlson, & Klein, 2014), the availability of the SDQ as a short and freely available well-validated measure of core symptoms of psychopathology, including those that are central to the CBCL-DP, across developmental age-groups is highly important, for both research and clinical purposes.

The consistency of our results with previous work on the CBCL-DP furthermore shows that the Dysregulation Profile is constructed similarly regardless of which questionnaire is used to measure behavioral and emotional symptoms. This suggests that the DP is indeed a unique phenotype that can probably be established with any broad behavioral screening measure. Future research could examine whether Severe Mood Dysregulation (SMD) and the Disruptive Mood Dysregulation Disorder (DMDD), aimed at capturing patterns of severe dysregulation as expressed in symptoms such as hyperarousal, mood instability, temper outbursts, and chronic irritability (Althoff, 2010; Leibenluft, 2011; Zepf & Holtmann, 2012), represent the extreme end of the dimensional spectrum of the DP, as the DP is thought to capture symptoms broadly overlapping with these clinical presentations (Dougherty et al., 2014; Legenbauer et al., 2016; Zepf & Holtmann, 2012).

While the three different SDQ-DP operationalizations were more alike than different, the bifactor SDQ-DP scores were consistently best in terms of stability, reliability and validity, while the 5-item measure persistently performed poorest. Use of all 15 items representing emotional symptoms, conduct problems and hyperactivity-inattention within a bifactor model, or using summed scores (Winsper & Wolke, 2014) is thus recommended as this might result in a more reliable and stable assessment of the DP. When possible in sufficiently large samples, we advise using bifactor modeling. Bifactor models are highly useful in psychopathology research (Snyder & Hankin, 2016), as etiology factors and outcomes can be examined in a more refined way. The SDQ-DP 5-item measure might, however, be a simple and efficient screening measure of dysregulation in practice, which could aid in early identification of children at risk for serious psychopathology, after which a more in-depth clinical assessment can be done. Also, the SDQ-DP 5-item measure could be used as an identifier in clinical populations (e.g. children with ADHD) to identify children at greater risk for difficulties with treatment adherence and recovery (Caro-Cañizares et al.,

2017). It would be useful to examine whether a shorter set of CBCL-items could similarly be developed and used as a screening measure of dysregulation.

Limitations and Future Directions

Strengths of this study include the systematic approach adopted to evaluating the SDQ-DP by examining competing factor models across different developmental periods and different reporters in a longitudinal sample, as well as testing statistically whether the best-fitting model was equivalent across gender, reporters and developmental period by examining measurement invariance. These factors strengthen the generalizability of the results. However, several limitations must also be noted. First, no direct comparison was possible with the CBCL - Dysregulation Profile as the CBCL was not assessed in the study from which the data were derived. However, previous research showed high overlap between the CBCL-DP and SDQ-DP (albeit not operationalized within a bifactor model), and good correspondence between the CBCL and SDQ generally (Goodman & Scott, 1999; Stone et al., 2010). For determination of the Dysregulation Profile specifically it might however be relevant to examine the impact of the presence of items describing more transdiagnostic emotional regulatory problems (e.g. 'Stubborn, sullen, or irritable'), that are known to strongly contribute to the DP (e.g., Geeraerts et al., 2015) in the CBCL. A systematic comparison of the CBCL-DP and the SDQ-DP in the same dataset would thus be useful.

Second, although we validated the SDQ-DP with measures of self-regulation and antisocial outcomes, future research should also validate the SDQ-DP bifactor model using clinical diagnostic measures of psychopathology such as in-depth diagnostic interviews. Such research is needed to determine the usefulness of the SDQ-DP bifactor model in clinical research and to further validate the specific factors. Third, children in the study were selected based on their below-median performance on literacy. In that sense, this was not a true community study, which likely explains that the prevalence rate of children exceeding the recommended clinical threshold of the SDQ-DP 5-item measure was around 10%, which is higher than previously reported numbers of 2.6% for SDQ-DP (Holtmann et al., 2011), and around 1% for CBCL-DP (Holtmann et al., 2007; Hudziak, Althoff, Derks, Faraone, & Boomsma, 2005). Children scoring below the median of a statewide literacy measure however do still form a large portion of the children in schools, but nonetheless, replication of our findings in large epidemiological samples as well as clinical samples is desirable.

The DP might have its roots in infant and toddler regulatory problems as expressed in sleeping and feeding problems and excessive crying (Winsper & Wolke, 2012). Future research could examine these early predictors of dysregulation to prevent maladaptive pathways resulting in dysregulation and eventually psychopathology. Given that the SDQ is brief and can be filled out in about five minutes, this measure has great potential to determine a Dysregulation Profile.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Althoff RR. Dysregulated children reconsidered. Journal of the American Academy of Child and Adolescent Psychiatry. 2010; 49:302–305. DOI: 10.1097/00004583-201004000-00004 [PubMed: 20410722]

Althoff RR, Ayer LA, Rettew DC, Hudziak JJ. Assessment of dysregulated children using the Child Behavior Checklist: A receiver operating characteristic curve analysis. Psychological Assessment. 2010a; 22:609–617. DOI: 10.1037/a0019699 [PubMed: 20822273]

Althoff RR, Verhulst FC, Rettew DC, Hudziak JJ, van der Ende J. Adult outcomes of childhood dysregulation: A 14-year follow-up study. Journal of the American Academy of Child & Adolescent Psychiatry. 2010b; 49:1105–1116. DOI: 10.1097/00004583-201011000-00004 [PubMed: 20970698]

Ayer L, Althoff R, Ivanova M, Rettew D, Waxler E, Sulman J, Hudziak J. Child Behavior Checklist Juvenile Bipolar Disorder (CBCL-JBD) and CBCL Posttraumatic Stress Problems (CBCL-PTSP) scales are measures of a single dysregulatory syndrome. Journal of Child Psychology and Psychiatry. 2009; 50:1291–1300. DOI: 10.1111/j.1469-7610.2009.02089.x [PubMed: 19486226]

Bellani M, Negri GAL, Brambilla P. The dysregulation profile in children and adolescents: A potential index for major psychopathology? Epidemiology and Psychiatric Sciences. 2012; 21:155–159. DOI: 10.1017/S2045796011000849 [PubMed: 22789163]

Block JH, Block J. The role of ego-control and ego-resiliency in the organization of behavior. In: Collins WA, editorMinnesota symposia on child psychology. Vol. 13. Hillsdale, NJ: Erlbaum; 1980. 39–101.

Boomsma DI, Rebollo I, Derks EM, van Beijsterveldt TCEM, Althoff RR, Rettew DC, Hudziak JJ. Longitudinal stability of the CBCL-juvenile bipolar disorder phenotype: A study in Dutch twins. Biological Psychiatry. 2006; 60:912–920. DOI: 10.1016/j.biopsych.2006.02.028 [PubMed: 16735031]

Caci H, Morin AJ, Tran A. Investigation of a bifactor model of the Strengths and Difficulties Questionnaire. European Child & Adolescent Psychiatry. 2015; 24:1291–1301. DOI: 10.1007/s00787-015-0679-3 [PubMed: 25631690]

Carballo JJ, Serrano-Drozdowskyj E, García Nieto R, Díaz de Neira-Hernando M, Pérez-Fominaya M, Molina-Pizarro CA, … Baca-García E. Prevalence and correlates of psychopathology in children and adolescents evaluated with the Strengths and Difficulties Questionnaire Dysregulation Profile in a clinical setting. Psychopathology. 2014; 47:303–311. DOI: 10.1159/000360822 [PubMed: 24819241]

Carlson GA. Who are the children with severe mood dysregulation, a.k.a. "rages"? The American Journal of Psychiatry. 2007; 164:1140–1142. DOI: 10.1176/appi.ajp.2007.07050830 [PubMed: 17671272]

Caro-Cañizares I, García-Nieto R, Carballo JJ. Biological and environmental predictors of the dysregulation profile in children and adolescents: The story so far. International Journal of Adolescent Medicine and Health. 2015; 27:135–141. DOI: 10.1515/ijamh-2015-5004 [PubMed: 25411984]

Caro-Cañizares I, Serrano-Drozdowskyj E, Pfang B, Baca-García E, Carballo JJ. SDQ Dysregulation Profile and its relation to the severity of psychopathology and psychosocial functioning in a sample of children and adolescents with ADHD. Journal of Attention Disorders. 2017; Advance online publication. doi: 10.1177/1087054717691829

Caspi A, Block J, Block JH, Klopp B, Lynam D, Moffit T, Stouthamer-Loeber M. A "common-language" version of the California Child Q-Set for personality assessment. Psychological Assessment. 1990; 4:512–523. DOI: 10.1037//1040-3590.4.4.512

Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, … Moffitt TE. The p factor: One general psychopathology factor in the structure of psychiatric disorders? Clinical Psychological Science. 2014; 2:119–137. DOI: 10.1177/2167702613497473 [PubMed: 25360393]

Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling: A Multidisciplinary Journal. 2007; 14:464–504. DOI: 10.1080/10705510701301834

Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling. 2002; 9:233–255. DOI: 10.1207/S15328007SEM0902_5

Deutz MHF, Geeraerts SB, van Baar AL, Dekovi M, Prinzie P. The Dysregulation Profile in middle childhood and adolescence across reporters: factor structure, measurement invariance, and links with self-harm and suicidal ideation. European Child & Adolescent Psychiatry. 2016; 25:431–442. DOI: 10.1007/s00787-015-0745-x [PubMed: 26226917]

Dick DM, Viken RJ, Kaprio J, Pulkkinen L, Rose RJ. Understanding the covariation among childhood externalizing symptoms: Genetic and environmental influences on conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder symptoms. Journal of Abnormal Child Psychology. 2005; 33:219–229. DOI: 10.1007/s10802-005-1829-8 [PubMed: 15839499]

Diler RS, Birmaher B, Axelson D, Goldstein B, Gill M, Strober M, … Keller MB. The Child Behavior Checklist (CBCL) and the CBCL-bipolar phenotype are not useful in diagnosing pediatric bipolar disorder. Journal of Child and Adolescent Psychopharmacology. 2009; 19:23–30. DOI: 10.1089/cap.2008.067 [PubMed: 19232020]

Dougherty LR, Smith VC, Bufferd SJ, Carlson GA, Stringaris A, Leibenluft E, Klein DN. DSM-5 disruptive mood dysregulation disorder: Correlates and predictors in young children. Psychological Medicine. 2014; 44:2339–2350. DOI: 10.1017/S0033291713003115 [PubMed: 24443797]

Geeraerts SB, Deutz MHF, Dekovi M, Bunte T, Schoemaker K, Espy KA, … Matthys W. The Child Behavior Checklist Dysregulation Profile in preschool children: A broad dysregulation syndrome. Journal of the American Academy of Child and Adolescent Psychiatry. 2015; 54:595–602. e2. DOI: 10.1016/j.jaac.2015.04.012 [PubMed: 26088665]

Goodman R. The Strengths and Difficulties Questionnaire: A research note. Journal of Child Psychology and Psychiatry. 1997; 38:581–586. DOI: 10.1111/j.1469-7610.1997.tb01545.x [PubMed: 9255702]

Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? Journal of Abnormal Child Psychology. 1999; 27:17–24. DOI: 10.1023/A:1022658222914 [PubMed: 10197403]

Hill CR, Hughes JN. An examination of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. School Psychology Quarterly. 2007; 22:380–406. DOI: 10.1037/1045-3830.22.3.380 [PubMed: 18843384]

Holtmann M, Becker A, Banaschewski T, Rothenberger A, Roessner V. Psychometric validity of the Strengths and Difficulties Questionnaire-Dysregulation Profile. Psychopathology. 2011; 44:53–59. DOI: 10.1159/000318164 [PubMed: 21072000]

Holtmann M, Bölte S, Goth K, Döpfner M, Plück J, Huss M, … Poustka F. Prevalence of the Child Behavior Checklist-pediatric bipolar disorder phenotype in a German general population sample. Bipolar Disorders. 2007; 9:895–900. DOI: 10.1111/j.1399-5618.2007.00463.x [PubMed: 18076540]

Holtmann M, Buchmann AF, Esser G, Schmidt MH, Banaschewski T, Laucht M. The Child Behavior Checklist-Dysregulation Profile predicts substance use, suicidality, and functional impairment: A longitudinal analysis. Journal of Child Psychology and Psychiatry. 2011; 52:139–147. DOI: 10.1111/j.1469-7610.2010.02309.x [PubMed: 20854363]

Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55. DOI: 10.1080/10705519909540118

Hudziak JJ, Althoff RR, Derks EM, Faraone SV, Boomsma DI. Prevalence and genetic architecture of Child Behavior Checklist-juvenile bipolar disorder. Biological Psychiatry. 2005; 58:562–568. DOI: 10.1016/j.biopsych.2005.03.024 [PubMed: 16239161]

Hyde LW, Shaw DS, Hariri AR. Understanding youth antisocial behavior using neuroscience through a developmental psychopathology lens: Review, integration, and directions for research. Developmental Review. 2013; 33:168–223. DOI: 10.1016/j.dr.2013.06.001

Kline RB. Principles and practice of structural equation modeling. 2. New York: Guilford Press; 2006.

Kóbor A, Takács Á, Urbán R. The bifactor model of the Strengths and Difficulties Questionnaire. European Journal of Psychological Assessment. 2013; 29:299–307. DOI: 10.1027/1015-5759/a000160

Kochanska G, Murray K, Coy KC. Inhibitory control as a contributor to conscience in childhood: from toddler to early school age. Child Development. 1997; 68:263–277. DOI: 10.1111/j.1467-8624.1997.tb01939.x [PubMed: 9180001]

Kochanska G, Murray KT, Harlan ET. Effortful control in early childhood: continuity and change, antecedents, and implications for social development. Developmental Psychology. 2000; 36:220–232. DOI: 10.1037//0012-1649.36.2.220 [PubMed: 10749079]

Kring AM. Emotion disturbances as transdiagnostic processes in psychopathology. In: Lewis M, Haviland-Jones JM, Barret LF, editorsHandbook of emotion. 3. New York: Guilford Press; 2008. 691–705.

Kwok O, Hughes JN, Luo W. Role of resilient personality on lower achieving first grade students' current and future achievement. Journal of School Psychology. 2007; 45:61–82. DOI: 10.1016/j.jsp.2006.07.002 [PubMed: 18084626]

Legenbauer T, Heiler S, Holtmann M, Fricke-Oerkermann L, Lehmkuhl G. The affective storms of school children during night time: Do affective dysregulated school children show a specific pattern of sleep disturbances? Journal of Neural Transmission. 2012; 119:989–998. DOI: 10.1007/s00702-012-0837-4 [PubMed: 22684420]

Legenbauer T, Hübner J, Pinnow M, Ball A, Pniewski B, Holtmann M. Proper emotion recognition, dysfunctional emotion regulation: The mystery of affective dysregulation in adolescent psychiatric inpatients. Zeitschrift Für Kinder- Und Jugendpsychiatrie Und Psychotherapie. 2016; 44:1–10. DOI: 10.1024/1422-4917/a000479

Leibenluft E. Severe mood dysregulation, irritability, and the diagnostic boundaries of bipolar disorder in youths. The American Journal of Psychiatry. 2011; 168:129–142. DOI: 10.1176/appi.ajp.2010.10050766 [PubMed: 21123313]

Mahoney JL, Stattin H. Leisure activities and adolescent antisocial behavior: The role of structure and social context. Journal of Adolescence. 2000; 23:113–127. DOI: 10.1006/jado.2000.0302 [PubMed: 10831137]

Martel MM, Nigg JT. Child ADHD and personality/temperament traits of reactive and effortful control, resiliency, and emotionality. Journal of Child Psychology and Psychiatry. 2006; 47:1175–1183. DOI: 10.1111/j.1469-7610.2006.01629.x [PubMed: 17076757]

Martel MM, von Eye A, Nigg JT. Revisiting the latent structure of ADHD: is there a "g" factor? Journal of Child Psychology and Psychiatry. 2010; 51:905–914. DOI: 10.1111/j.1469-7610.2010.02232.x [PubMed: 20331490]

McGough JJ, Loo SK, McCracken JT, Dang J, Clark S, Nelson SF, Smalley SL. CBCL pediatric bipolar disorder profile and ADHD: Comorbidity and quantitative trait loci analysis. Journal of the American Academy of Child and Adolescent Psychiatry. 2008; 47:1151–1157. DOI: 10.1097/CHI.0b013e3181825a68 [PubMed: 18724256]

Meade AW, Johnson EC, Braddy PW. Power and sensitivity of alternative fit indices in tests of measurement invariance. The Journal of Applied Psychology. 2008; 93:568–592. DOI: 10.1037/0021-9010.93.3.568 [PubMed: 18457487]

Muthen LK, Muthen BO. Mplus User's Guide. 7. Los Angeles, CA: Muthén & Muthén; 2012.

Olino TM, Dougherty LR, Bufferd SJ, Carlson GA, Klein DN. Testing models of psychopathology in preschool-aged children using a structured interview-based assessment. Journal of Abnormal Child Psychology. 2014; 42:1201–1211. DOI: 10.1007/s10802-014-9865-x [PubMed: 24652485]

Patalay P, Fonagy P, Deighton J, Belsky J, Vostanis P, Wolpert M. A general psychopathology factor in early adolescence. The British Journal of Psychiatry. 2015; 207:15–22. DOI: 10.1192/bjp.bp.114.149591 [PubMed: 25906794]

Paulhus DL, Robins RW, Trzesniewski KH, Tracy JL. Two replicable suppressor situations in personality research. Multivariate Behavioral Research. 2004; 39:303–328. DOI: 10.1207/s15327906mbr3902_7 [PubMed: 26804578]

Pettersson E, Larsson H, Lichtenstein P. Common psychiatric disorders share the same genetic origin: A multivariate sibling study of the Swedish population. Molecular Psychiatry. 2016; 21:717–721. DOI: 10.1038/mp.2015.116 [PubMed: 26303662]

Reise SP. The rediscovery of bifactor measurement models. Multivariate Behavioral Research. 2012; 47:667–696. DOI: 10.1080/00273171.2012.715555 [PubMed: 24049214]

Snyder HR, Hankin BL. All models are wrong, but the p factor model is useful: Reply to Widiger and Oltmanns (2017), and Bonifay, Lane, and Reise (2017). Clinical Psychological Science. 2016; 5:187–189. DOI: 10.1177/2167702616659389 [PubMed: 28286706]

Stone LL, Otten R, Engels RC, Vermulst AA, Janssens JM. Psychometric properties of the parent and teacher versions of the Strengths And Difficulties Questionnaire for 4-to 12-year-olds: A review. Clinical Child And Family Psychology Review. 2010; 13:254–274. DOI: 10.1007/s10567-010-0071-2 [PubMed: 20589428]

Winsper C, Wolke D. Infant and toddler crying, sleeping and feeding problems and trajectories of dysregulated behavior across childhood. Journal of Abnormal Child Psychology. 2014; 42:831–843. DOI: 10.1007/s10802-013-9813-1 [PubMed: 24122374]

Zepf FD, Holtmann M. Disruptive Mood Dysregulation Disorder. In: Rey JM, editorIACAPAP e-Textbook of Child and Adolescent Mental Health. Geneva: International Association for Child and Adolescent Psychiatry and Allied Professionals; 2012. Retrieved from http://www.iacapap.org/wp-content/uploads/E.3-MOOD-DYSREGULATION-072012.pdf

Public Significance Statement

The Strengths and Difficulties Questionnaire can be used to measure a broad syndrome of affective, behavioral, and cognitive dysregulation: the SDQ – Dysregulation Profile. Preferably a bifactor operationalization is used. The 15-item SDQ-DP measure is preferred over the SDQ-DP 5-item measure given greater reliability, validity, and stability.
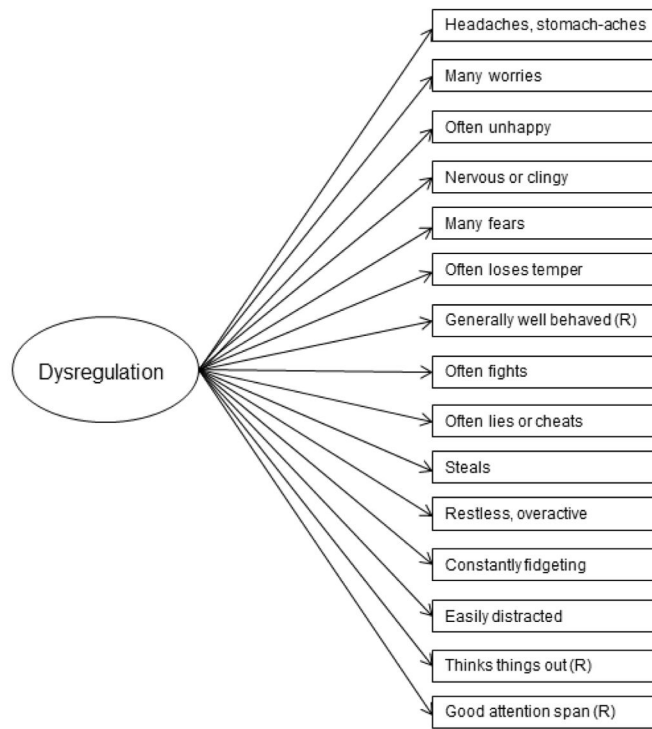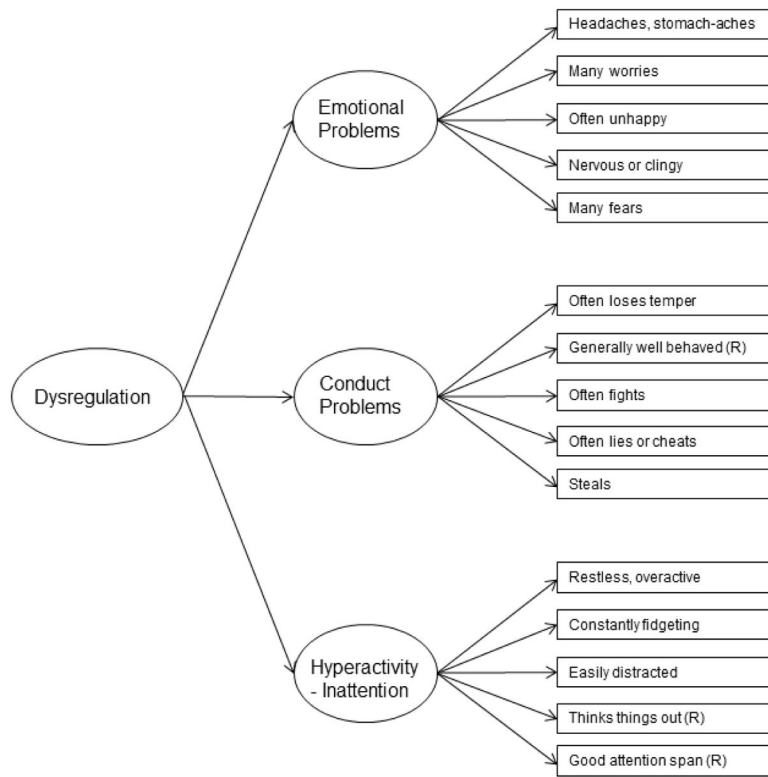
Figure 1a

Figure 1b

Figure 1c

Figure 1.
Figure 1a. One-factor SDQ-DP model
Figure 1b. Second-order factor SDQ-DP model
Figure 1c. Bifactor SDQ-DP model

Table 1

Fit Indices for the One-Factor, Second-Order, and Bifactor SDQ-DP Models in Three Developmental Periods across Three Different Reporters

Reporter	Developmental Period	Model	N	χ^2	df	RMSEA	RMSEA 90% CI	CFI	TLI	χ^2
Parents	Early Childhood	Bifactor	498	231.901	76	.064	[.055 – .074]	.954	.937	
		Second-Order	498	315.327	87	.073	[.064 – .081]	.933	.919	2 vs. 1 (11) = 79.451, $p < .001$
		One-Factor	498	697.866	90	.116	[.108 – .125]	.822	.793	3 vs. 2 (3) = 177.426, $p < .001$
	Middle Childhood	Bifactor	446	223.789	76	.066	[.056 – .076]	.953	.935	
		Second-Order	446	280.757	87	.071	[.062 – .080]	.938	.925	2 vs. 1 (11) = 55.004, $p < .001$
		One-Factor	446	634.314	90	.116	[.108 – .125]	.826	.797	3 vs. 2 (3) = 163.070, $p < .001$
	Adolescence	Bifactor	352	144.659	76	.051	[.038 – .063]	.975	.966	
		Second-Order	352	182.130	87	.056	[.044 – .067]	.966	.959	2 vs. 1 (11) = 38.286, $p < .001$
		One-Factor	352	315.673	90	.084	[.074 – .095]	.919	.905	3 vs. 2 (3) = 63.568, $p < .001$
Teachers	Early Childhood	Bifactor	678	327.900	76	.070	[.062 – .078]	.982	.975	
		Second-Order	678	504.522	87	.084	[.077 – .091]	.969	.963	2 vs. 1 (11) = 125.433, $p < .001$
		One-Factor	678	1771.837	90	.166	[.159 – .173]	.877	.857	3 vs. 2 (3) = 427.096, $p < .001$
	Middle Childhood	Bifactor	528	277.347	75	.071	[.063 – .081]	.974	.964	
		Second-Order	528	565.903	87	.102	[.094 – .110]	.939	.927	2 vs. 1 (12) = 196.821, $p < .001$
		One-Factor	528	1274.537	90	.158	[.150 – .166]	.850	.825	3 vs. 2 (3) = 252.397, $p < .001$
	Adolescence	Bifactor	437	210.401	75	.064	[.054 – .075]	.978	.969	
		Second-Order	437	353.954	87	.084	[.075 – .093]	.957	.948	2 vs. 1 (12) = 114.941, $p < .001$
		One-Factor	437	752.857	90	.130	[.121 – .138]	.892	.875	3 vs. 2 (3) = 149.534, $p < .001$
Children	Adolescence	Bifactor	505	232.328	75	.064	[.055 – .074]	.944	.922	
		Second-Order	505	345.062	87	.077	[.068 – .085]	.908	.889	2 vs. 1 (12) = 97.898, $p < .001$
		One-Factor	505	647.086	90	.111	[.103 – .119]	.801	.768	3 vs. 2 (3) = 164.440, $p < .001$

Note. Degrees of freedom can differ due to restriction of factor loadings because of errors (see main text).

Table 2

Standardized Factor loadings for the SDQ-DP Bifactor Models across Three Reporters and Developmental Periods

| | T1 (Early Childhood) | | | | T2 (Middle Childhood) | | | | T3 (Adolescence) | | | | | |
| | Parents | | Teacher | | Parents | | Teacher | | Parents | | Teachers | | Youth | |
	S-FL	DP-FL	S-FL	DP-FL	S-FL	DP-FL	S-FL	DP-FL	S-FL	DP-FL	S-FL	DP-FL	S-FL	DP-FL
Emotional Problems														
3. Often complains of headaches	.524	.224	.406	.230	.401	.305	.477	.411	.340	.520	.482	.467	.414	.333
8. Many worries	.652	.330	.802	.176	.573	.333	.757	.277	.526	.544	.768	.301	.726	.330
13. Often unhappy, downhearted	.581	.468	.642	.475	.365	.608	.644	.487	.417	.594	.604	.455	.666	.490
16. Nervous or clingy in new situations	.487	.452	.611	.334	.525	.443	.745	.435	.345	.467	.669	.420	.483	.267
24. Many fears	.659	.361	.820	.212	.692	.310	.813	.269	.527	.540	.815	.294	.565	.256
Conduct Problems														
5. Often has temper tantrums or hot tempers	.429	.492	.477	.614	.162	.726	.514	.627	-.012	.732	.525	.732	.316	.609
7. Generally obedient (R)	.131	.554	.373	.798	.103	.653	.234	.793	.128	.674	.200	.833	.096	.556
12. Often fights with other children	.520	.582	.587	.678	.285	.741	.498	.688	.272	.775	.606	.737	.664	.488
18. Often lies or cheats	.589	.558	.546	.615	.700	.643	.654	.628	.618	.666	.155	.827	.341	.593
22. Steals from home, school, or elsewhere	.568	.510	.684	.527	.534	.526	.656	.575	.500	.581	.269	.628	.412	.521
Hyperactivity-Inattention														
2. Restless, overactive	-.109	.830	.363	.868	.426	.688	.616	.720	.168	.801	.707	.708	.490	.685
10. Constantly fidgeting or squirming	-.116	.850	.500	.855	.390	.708	.608	.727	.210	.758	.580	.669	.477	.591
15. Easily distracted, concentrating wanders	.151	.796	.063	.855	.557	.566	.214	.826	.374	.683	.390	.740	.262	.768
21. Thinks things out before acting (R)	.411	.512	-.074	.815	.468	.373	-.161	.856	.307	.596	-.029	.816	-.489	.710
25. Sees tasks through to the end (R)	.700	.665	-.083	.864	.658	.462	.048	.856	.800	.581	.091	.871	-.150	.598

Note. Factor loadings in bold are significant at p<.001 level, underlined at p<.01 level, and italicized at p<.05.

Note. S-FL stands for scale-specific factor loadings, DP-FL stands for factor loadings on the general DP bifactor.

Note. Items followed by (R) are reverse-coded

Table 3

Standardized Factor Loadings for the SDQ-DP One-Factor, Second-Order and Bifactor Model for Parent-Reports in Early Childhood

	One-factor Model	Second-order Model		Bifactor Model	
	DP-FL	S-FL	S-O FL	S-FL	DP-FL
Emotional Problems			.582		
3. Often complains of headaches	.350	.505		.524	.224
8. Many worries	.486	.672		.652	.330
13. Often unhappy, downhearted	.599	.794		.581	.468
16. Nervous or clingy in new situations	.547	.728		.487	.452
24. Many fears	.516	.709		.659	.361
Conduct Problems			.858		
5. Often has temper tantrums or hot tempers	.548	.644		.429	.492
7. Generally obedient (R)	.541	.631		.131	.554
12. Often fights with other children	.661	.769		.520	.582
18. Often lies of cheats	.650	.756		.589	.558
22. Steals from home, school, or elsewhere	.628	.716		.568	.510
Hyperactivity			.844		
2. Restless, overactive	.778	.822		−.109	.830
10. Constantly fidgeting or squirming	.800	.833		−.116	.850
15. Easily distracted, concentrating wanders	.790	.823		.151	.796
21. Thinks things out before acting (R)	.537	.588		.411	.512
25. Sees tasks through to the end (R)	.679	.732		.700	.665

Note. Factor loadings in bold are significant at p<.001 level, and italicized at p<.05.

Note. S-FL stands for scale-specific factor loadings, DP-FL stands for factor loadings, DP-FL stands for factor loadings on the general DP bifactor, and S-O FL stands for second-order factor loadings.

Note. Items followed by (R) are reverse-coded

Table 4

Regression Coefficients for Construct Validity Analyses

		Parent-Reports						Teacher-Reports					
		DP Bifactor Model						DP Bifactor Model					
		5-item DP	15-item DP	DP	ES	CP	H-I	5-item DP	15-item DP	DP	ES	CP	H-I
Construct	Ego-resiliency	**-.309**	**-.368**	**-.366**	-.032	-.084	-.099	**-.541**	**-.624**	**-.553**	**-.418**	**-.200**	**.251**
Validity	Effortful Control	_-.164_	_-.125_	_-.132_	-.043	-.095	.068	_-.103_	_-.125_	_-.106_	_-.107_	-.056	.038
Longitudinal	Disciplinary Measures	**.231**	**.274**	**.332**	_-.152_	_.231_	-.137	**.231**	**.286**	**.273**	-.095	**.270**	.009
Validity	Antisocial Behavior	**.249**	**.260**	**.286**	_-.151_	_.228_	-.045	**.138**	**.144**	**.262**	_-.177_	-.004	_-.169_

Note. ES = Emotional Symptoms, CP = Conduct Problems, H-I = Hyperactivity-Inattention

Note. 5-item DP is the SDQ-DP 5-item measure, 15-item DP is a summed score of the items from the ES, CP and H-I scales

Note. Estimates (STDYX standardized beta's) in bold are significant at p<.001 level, underlined at p<.01 level, and italicized at p<.05.