Dimensions and Latent Classes of Episodic Mania-Like Symptoms in Youth: An Empirical Enquiry

Argyris Stringaris · Daniel Stahl · Paramala Santosh · Robert Goodman

Abstract The dramatic increase in diagnostic rates of bipolar disorder in children and adolescents in the USA has led to an intense interest in the phenomenology of the disorder. Here we present data from a newly-developed instrument to assess episodic mania-like symptoms in youth in a large population-based sample (N=5326) using parent- and self-report. We found that a substantial proportion of children screened positive for having episodes of "going high" and were at an increased risk for morbidity and impairment. Using factor analysis, we identified that episodic mania-like symptoms comprised two dimensions: An *under-controlled* dimension that was associated with significant impairment, and a low-risk *exuberant* dimension. Using latent class analysis, we identified a small group of children scoring high on a range of manic symptoms and suffering from severe psychosocial impairment and morbidity. Our results carry implications for the nosology and psychosocial impairment associated with episodic mood changes in young people.

A. Stringaris (✉) · R. Goodman
Department of Child and Adolescent Psychiatry,
King's College London, Institute of Psychiatry,
Box 85, De Crespigny Park, Denmark Hill,
London SE5 8AF, UK
e-mail: argyris.stringaris@kcl.ac.uk

D. Stahl
Department of Biostatistics,
King's College London, Institute of Psychiatry,
London, UK

P. Santosh
Department of Psychological Medicine,
Great Ormond Street Hospital London,
London, UK

Introduction

Rates of diagnosis of bipolar disorder (BP) in children and adolescents have increased dramatically over the last decade in the USA (Blader and Carlson 2007; Moreno et al. 2007). This has raised questions about the phenomenology of BP (Carlson and Meyer 2006) and its diagnostic boundaries (Carlson and Glovinsky 2009; Youngstrom et al. 2008a) leading to extensive research on the validity of existing and suggested DSM-IV BP categories for children and adolescents. This study uses an alternative approach to the question of youth BP starting at the level of individual symptoms that occur during an episode of elated mood. In a first step, we examine whether young people screening positive for episodes of elated mood are increased risk for psychiatric morbidity and social impairment. We then use multivariate statistical methodology to help identify dimensions and classes of mania-like symptomatology that occur as part of distinct episodes of changes in mood. This is a means of assessing whether such mania-like symptoms form meaningful dimensions or clusters, independently of prior theoretical expectations. Therefore, in a second step, the study examines whether meaningful dimensions of mania-like symptoms can be derived. As a starting point for the validation of these symptom dimensions occurring during discrete episodes, we examine how they relate to other common childhood psychopathology in a community sample. This is important, given that "gold standard" cases

of BP are too rare in community samples (Costello et al. 1996) to use for validation purposes. In addition, we further assess the clinical relevance of such dimensions by estimating their relationship to social impairment. In a third step, we examine whether meaningful groups of individuals with mania like symptoms can be empirically derived and how they relate to other psychopathology and social impairment.

A significant part of existing research into "pediatric bipolar disorder" has focused on how the DSM-IV categories of BP-I and BP-II apply to children and adolescents and to possible modifications of these categories. For example, Birmaher and colleagues in their large clinic-referred sample used a longitudinal design to study the clinical importance of BP not otherwise specified (BP-NOS) (Birmaher et al. 2006), showing that approximately one third of such cases show a transition to either BP-I or BP-II over a 3 year follow up. These findings suggested that BP-NOS–defined by having episode durations of less than 4 days—is important and may be on a spectrum with BP-I and BP-II. Similarly, in a recent community-based study, Stringaris et al., showed that episodes that meet the symptom and impairment criteria of DSM-IV BP-I or BP-II but are of shorter duration (also termed BP-NOS), were common and led to functional impairment beyond what could be accounted for by other DSM-IV diagnoses (Stringaris et al. 2010b). However, the particularly poor agreement between parent- and child-reported episodes and the fact that they did not increase in duration with advancing age did cast some doubt on the validity of BP-NOS and whether it was on a continuum with BP-I and BP-II (Stringaris et al. 2010b).

Leibenluft and colleagues devised an *ad hoc* category called severe mood dysregulation (Leibenluft et al. 2003) to study children presenting with extreme non-episodic irritability and hyper-arousal, as this had previously been suggested to be an early developmental manifestation of BP-I or BP-II (Biederman 2006). In a recent study, it was shown that severe mood dysregulation is unlikely to convert to BP over a 28-month follow up (Stringaris et al. 2010a). Previous studies have shown differences as well as overlap in the neuropsychological (Guyer et al. 2007; Rich et al. 2008) and pathophysiological (Brotman et al. 2010) mechanisms between severe mood dysregulation and BP.

The generation of categories based on prior theory or ad hoc modifications of existing disorders has important advantages in the research on BP phenotypes. Firstly, such attempts use the DSM-IV as a point of reference and, thus, allow communication and data comparison across groups. Secondly, such approaches can be used to test important clinical questions directly—for example establishing whether brief episodes that meet all the criteria for mania except for duration are predictive of future BP-I or BP-II. However, patterns of symptoms may exist that do not

conform to existing categories or that are difficult to predict on the basis of prior theory.

An alternative to using categories generated on the basis of historical or theoretical considerations is the empirical derivation of symptom dimensions and classes. A prominent example of this approach has been the application of latent class analysis on the items of the Child Behavior Checklist (CBCL) to identify groups of children with potential bipolar disorder. Initial analyses using the CBCL identified a group of children who had a high probability of endorsing symptoms from the aggression, attention/hyperactivity problems, and anxiety-depression domains (Hudziak et al. 2005). This profile of symptom endorsement, which was shown to have considerable heritability, was initially termed "juvenile bipolar disorder" (Faraone et al. 2005). However, subsequent work cast doubt on the notion that this symptom profile corresponded to a manifestation of bipolar disorder. Instead, its associations with PTSD, suicidality and other mood problems were identified (Holtmann et al. 2007; Meyer et al. 2009; Volk and Todd 2007), leading to the re-labelling of this group as "dysregulation profile"(Ayer et al. 2009; Holtmann et al. 2010). The progress made using an empirical approach to the CBCL profile provides an important example of how multivariate methodology can prove useful in testing the boundaries of BP in youth. However, a limitation of using the CBCL is that it was not designed to ascertain symptoms that occur during episodes of changes in mood—an important limitation since episodic changes in mood are integral to the definition of bipolar illness (Leibenluft et al. 2003). This centrality is reflected in the DSM-IV "Criterion A" for a manic episode requiring "a distinct period of abnormally and persistently elevated, expansive, or irritable mood." (APA 2000). Another example of a multivariate approach has been the empirical derivation of a 10-item mania scale from the Parent General Behavior Inventory for children and adolescents (Youngstrom et al. 2008b). The scale shows good psychometric properties, including discrimination between bipolar and unipolar depression, and between mood disorders and ADHD (Youngstrom et al. 2008b). However, the primary aim of this study was to develop a screening instrument to be used in a sample referred to a mood clinic, rather than examine the structure of a broad range of symptoms in a community-based non-referred sample.

This study uses a new interview focusing on a broad range of mania-like symptoms that were reported by parents or young people themselves in a large population-based sample of children and adolescents. As our focus is a better understanding of possible developmental variants or phenocopies of bipolar disorder, we seek to ascertain symptoms that occur as part of discrete episodes, rather than symptoms that are chronically present. It is important

to note that some authors have suggested that episodicity may not be characteristic of early life mania and that chronic (non episodic) irritability may be the hallmark of early bipolar presentations (Wozniak et al. 1995). However, subsequent studies have not found evidence to support such a view (Leibenluft 2011; Potegal, et al. 2009; Stringaris et al. 2010a). In light of the current state of evidence, this study anchors its examination of mania-like symptoms on the *a priori* assumption that bipolar illness is characterized by episodicity (APA 2000).

The first question we address is whether children screening positive for episodic changes in mood are at an increased risk for morbidity and impairment. This is particularly important to establish given concerns that measures or concepts that do not pay enough attention to life stage may label as pathological some episodic mood changes that are normative in childhood, e.g., short periods of exuberance at a party. In the literature about bipolar disorder in adults, it is reported that short episodes of elated mood (less than 4 days) have similar characteristics to typical (hypo-) manic episodes (Angst et al. 2003). In addition, we have previously shown that even short-lived episodes of elated mood during which symptom and impairment criteria for mania are fulfilled, are associated with significant social impairment (Stringaris et al. 2010b). Here we examine whether even the single screening question about episodic changes in mood is a predictor of morbidity and social impairment.

Our second question is whether it is possible to identify meaningful dimensions of mania-like symptoms. An important subsidiary question for any such dimension is whether that dimension is associated with distress or social impairment ("impact"), particularly when adjusting for other dimensions of mania-like symptoms and for comorbid psychopathology. A dimension that is strongly associated with impact may be a form of psychopathology, whereas a dimension without apparent adverse consequences may better be seen as normal variation, perhaps related to temperament or personality.

The third question of the paper is whether empirical approaches can identify particular classes of individuals based on their profile of mania-like symptoms. If so, an important subsidiary issue is whether different classes are differentially predictive of social impairment and comorbidity.

As mentioned above, "gold standard" cases, that is, patients who fulfil DSM-IV criteria for manic or hypomanic episodes (including the duration criterion), are rare in community samples (Costello et al. 1996; Stringaris et al. 2010b). It is therefore not possible to use such cases as validators of the dimensions and latent classes derived from this study. Obviously, using cases of a possible broader phenotype (Leibenluft et al. 2003) of bipolar disorder is also impossible, since it is precisely the validity of such

phenotypes that is at stake. Instead, we provide preliminary validation of empirical categories and dimensions by assessing associated comorbidity and social impairment.

Methods

Population

The field work for the 2004 British Child and Adolescent Mental Health Survey (B-CAMHS04) was conducted in 2004. The sample was a representative group of 5–16 year olds (*N*=7977). The design of this survey has been previously described in detail (Green et al. 2005). Briefly, the study used "Child benefit" to generate a sampling frame. This refers to a universal state benefit payable in Great Britain for each child in a family. This was used to develop a sampling frame of 5–16 year olds in the different postal sectors in England, Wales, and Scotland; after excluding families with no recorded postal code, it was estimated that this represented 90% of all British children. Out of the 12,294 recruited, there were 1,085 who opted out and 713 who were non eligible or had moved without trace, leaving 10,496 who were approached. Of those, 7,977 participated (65% of those selected; 76% of those approached). Three years after the original field work (i.e., in 2007 (Parry-Langdon 2008)), families were approached once more unless they had previously opted out or the child was known to have died. Of the original 7,977 participants, 5,326 (67%) participated in the detailed follow-up (Parry-Langdon 2008).

Measures

The Strengths and Difficulties Questionnaire (SDQ) is a 25-item scale that has been shown to possess robust psychometric properties (Bourdon et al. 2005; Goodman 1997, 2001). It was administered to parents and youth and generated an SDQ total symptoms score (reflecting hyperactivity, inattention, behaviour problems, emotional symptoms and peer problems). There are no symptoms of elated or expansive mood in the SDQ symptom items.

The SDQ total impact score, a measure of overall distress and social impairment due to all mental health problems (Goodman and Scott 1999), was also used. It specifically includes impairment in the domains of family, school, learning and leisure. Hereafter, we refer to this score as "social impairment".

The Development and Well-Being Assessment (DAWBA) was used in the survey (Ford et al. 2003; Goodman et al. 2000). It is a structured interview administered by lay interviewers with questions that are closely related to the diagnostic criteria of the DSM-IV (American Psychiatric

2000). It focuses on current rather than life-time problems. The κ-statistic for chance-corrected agreement between two raters was 0.86 for any disorder (SE 0.04), 0.57 for internalizing disorders (SE 0.11), and 0.98 for externalizing disorders (SE 0.02) (Ford et al. 2003). Children were given a diagnosis only if the symptoms reported were causing significant distress or social impairment. The DAWBA interview was administered to all parents and to all youth aged 11 or more, while teachers were administered an abbreviated questionnaire. In this, as in previous studies using the DAWBA, DSM-IV diagnoses are assigned by integrating information from the parent, youth and teacher reports on symptoms and impairment (Ford et al. 2003). Further information on the DAWBA is available from http://www.dawba.info, including on-line and downloadable versions of the measures and demonstrations of the clinical rating process. For the purposes of the analyses in this paper, diagnoses of anxiety and depression were treated as part of an over-riding category of "emotional problems".

As described previously (Stringaris et al. 2010b), the 2007 survey incorporated questions on elated mood and symptoms of mania for parents and youth (but not teachers). The complete bipolar section of the interview can be seen at http://www.dawba.info/Bipolar/. Throughout this paper, the findings for parent-report and youth report on episodes of elated mood and mania like symptoms are presented separately. In summary, the parents of 8–19 year olds and the 11–19 year olds themselves were presented with the following preamble: "*Some young people have episodes of going abnormally high. During these episodes they can be unusually cheerful, full of energy, speeded up, talking fast, doing a lot, joking around, and needing less sleep. These episodes stand out because the young person is different from their normal self.*" They were then asked: "*Do you [Does X] ever go abnormally high?*", to which they had the options of answering: *No, A little, A lot*. For this screening question, those answering "A little" had significantly more comorbidity and social impairment than those answering "No". In the interest of having as broad as possible a representation of subjects, we chose to include as "screen positive" those who answered "A little" as well as those who answered "A lot"—enquiring whether 26 specific symptoms of mania (including those stipulated by DSM-IV(APA 2000)) occurred during such episodes of going high. For each individual symptom, the participants had the option of choosing between one of the following answers: *No, A little, A lot*. Given that the answers *No* and *A little* did not differ with respect to their prediction of overall impairment, we dichotomised each item by merging *No* and *A little*. Re-running analyses using the trichotomous item coding did not alter the key findings of this paper.

Statistical Analyses

The DAWBA bipolar module was completed by 93% (583/627) of participating parents and 96% (872/913) of participating youth. There were no statistically significant differences in rates of overall psychopathology between those who completed the module and those who did not ($\chi^2=0.42$, $p=0.52$, by parent report; $\chi^2=3.31$, $p=0.07$, by self report). By parent report, incomplete information was associated with a significantly higher level of social impairment ($t=2.3$, $p=0.03$), but this was not replicated in the self report ($t=1.4$, $p=0.15$). Given the low attrition rate and the weak evidence than incomplete responders differed from complete responders, we used listwise deletion to deal with missing data, including in the analysis only those cases with complete data on all 26 items of the DAWBA bipolar module.

The first question of the paper was whether children screening positive for episodes of going high were at increased risk of comorbidity and social impairment. To address this we: a) estimated how often children screened positive for episodic changes in mood; b) we compared the prevalence of psychopathology in those screening positive and negative using χ^2 tests; and c) we used linear regression models to test the association between screening positive for episodic mood changes and impairment.

The second question of the paper was whether it was possible to identify meaningful dimensions of mania-like symptoms emprically. To address this, we first performed exploratory factor analyses (EFA) using weighted least square estimators for binary data, followed by orthogonal rotation. As with all analyses of mania-like symptoms in this paper, parent-reported symptoms were analysed separately from youth-reported symptoms, i.e., no attempt was made to combine parent and youth reported symptoms into "combined" dimensions or categories. Choice between factor solutions was based on inspecting eigenvalue (scree) plots and on parsimony. We summed items with factor loadings of 0.5 or more to create sub-scales corresponding to the extracted factors. In a second step, we performed logistic regression models to test the association between a given psychiatric diagnosis (the outcome) and the sum scores of the extracted factors (predictors) while controlling for age and gender. Finally, in linear regression models we tested the association between the SDQ impact score (the outcome) and the sum scores of the extracted factors (predictors), controlling for age, gender and the presence of other diagnoses. Analyses for each reporting source were performed separately. For the analyses of parent-reported dimensions, we used parent-reported SDQ as the outcome; whereas for the analyses involving self-reported dimensions, we used self-reported SDQ as the outcome.

The third question of the paper was whether it was possible to identify classes of individuals based on their profile of mania-like symptoms. To address this, latent class analyses (LCA) were performed for binary data using maximum likelihood estimation with robust standard errors. A latent class analysis (LCA) was conducted to explain associations between observed manifest indicator variables (clinical observations) through hypothesized underlying unobserved latent variables. LCA is a model based cluster analysis method used to identify subtypes of related cases (latent classes or cluster) from categorical data (Lazarsfeld and Henry 1968; Muthen 2001). The method assumes k latent clusters or latent classes underlying the data set and that each case belongs to only one group. The number of classes and their sizes are not known a priori. LCA uses maximum likelihood estimation methods to minimize association (i.e., satisfy mutual independence) among the responses across multiple observed variables within each latent class (Agresti 2002). By using LCA for cluster analysis, the clustering problem becomes that of estimating the parameters of the assumed mixture and then using the estimated parameters to calculate the posterior probabilities of cluster membership. A case is assigned to the cluster with the highest posterior probability. It recognizes that there is some degree of uncertainty in the classification by assigning each case a posterior probability of belonging to each cluster. As with all analyses of mania-like symptoms in this paper, LCA for the parent-reported symptoms was performed separately to those for the youth-reported symptoms. We compared models with between one and six classes. Deciding on the number of classes for latent class modes is not completely resolved (Nylund et al. 2007). Using information criteria, such as the Bayesian information criterion (BIC), favoured solutions with increasing numbers of small classes that were theoretically implausible; we encountered the same problem when using bootstrap likelihood ratios as an alternative. Therefore, following recommendations, we plotted log likelihood values (Nylund et al. 2007) and BIC values against number of classes as a descriptive method on which to base our decision. To assess the agreement between raters for the latent classes, we estimated kappa inter-rater agreement statistics. To explore the association of the latent classes with age and gender we used ANOVA tests (with post-hoc Tukey tests) and χ^2 tests, respectively. To examine the validity of the latent classes using external validators (psychiatric diagnoses and social impairment), we used χ^2 tests and analysis of variance (ANOVA) approaches. In addition, the latent classes were used as the outcome in a multinomial logistic regression model in which social impairment as measured by the SDQ was entered as a predictor with or without further covariates such as other psychiatric diagnoses. Relative risk ratios (RR) with 95%

confidence intervals were calculated. Multinomial models pair each response category with a baseline category. Here one of the latent classes served as the baseline and each of the other categories was compared to that baseline. Thus the RR refers to the change in relative risk ratio for a one point increase in SDQ impairment scores for one of the classes *relative* to the baseline category, adjusting for the other variables in the model. For the analyses involving parent-reported latent classes, we used parent-reported SDQ as the outcome; for the analyses involving self-reported latent classes, we used self-reported SDQ as the outcome. To correct for multiple comparisons, a protected significance level of $p < 0.001$ was used for the regression models predicting from parent- or self- reported dimensions to psychopathology and impairment. As stated above, in all the analyses performed in this paper, parent- and self-reported mania-like symptoms were analysed in different models and data from the two reporters were never entered as simultaneous predictors in the models.

EFA and LCA were performed in MPlus Version 5 (Muthen and Muthen 2007).

Results

Screening Positive for Episodic Mood Changes

The initial screening question asked about distinct episodes of going high. On the parent report, 10.5% replied "a little" to the screening question and a further 2.2% replied "a lot" to the screening question. On the youth report, 23% replied "a little" to the screening question and a further 5% replied "a lot". The agreement between parent and youth reports for such episodes was low but significant ($\kappa=0.14$, $p < 0.001$). Those screening positive for episodes of elated mood (answering "a little" or "a lot" as opposed to "no") had a higher rate of psychopathology: 27.6% compared with 6.5% for those screening positive ($N=4627$) by parent report ($\chi^2=300.3$. $p < 0.001$); and 12.7% compared with 6.4% for those screening positive by self report ($\chi^2=35.6$, $p < 0.001$). Screening positive for episodes of elated mood by parent report was significantly associated with a raised SDQ impact score ($b=0.48$, 95% CI 0.41–0.55) even after adjusting for depressive and anxiety disorders, ADHD, and CD/ODD as covariates; similarly, screening positive for episodes of elated mood by self report was significantly associated with a raised SDQ impact score ($b=0.25$, 95% CI 0.21–0.30) even after adjusting for depressive and anxiety disorders, ADHD, and CD/ODD.

Of those screening positive for episodes of elated mood by parent report, 85% reported a typical duration of less than a day, 13.6% a duration of 1–3 days, 0.6% between 4–6 days, and 0.8% of 1 week or more.

Of those screening positive for episodes of elated mood by self report, 80% reported a typical duration of less than a day, 16% a duration of 1–3 days, 2.1% between 4–6 days, and 1.2% of 1 week or more.

Symptom Dimensions

The internal reliability of the DAWBA scale for parent-reported mania-like symptoms (generated by summing the individual items) was excellent, with a Cronbach's alpha of 0.90. The internal reliability of the DAWBA scale for self-report symptoms (generated by summing the individual items) was almost as high, with a Cronbach's alpha of 0.88.

For both parent and self reported mania-like symptoms, a two factor model solution to the exploratory factor analysis emerged from inspection of scree plots and on the basis of parsimony. As shown in Tables 1 and 2, the factor structure was very similar for both informants. On the basis of this analysis, we generated sub-scales (by summing the individual items) designating the dimensions

as *episodic under-control* (Cronbach's alpha = 0.86) and *episodic exuberance* (Cronbach's alpha = 0.84) respectively. The correlation between the two sum-score sub-scales was moderate: 0.54 for parents and 0.56 for self report.

Tables 3 and 4 show the results of logistic regression models using parent and self report respectively. Episodic exuberance was only ever predictive of *psychopathology* when entered in models without episodic under-control. When the two DAWBA scales were entered simultaneously, only episodic under-control was predictive of psychopathology.

The association of the two parent-reported sub-scales with *impact* was tested in linear regression models where both sub-scales were entered simultaneously: only episodic under-control ($b=0.26$; 95% CI 0.20–0.32) was predictive of impact, while parent-reported episodic exuberance was not ($b=0.00$; 95% CI −0.06–0.07). Episodic under-control continued to predict impact even after adjusting for the presence of DSM-IV disorders or dimensional measures of psychopathology (see supplementary results). The results were similar for the association between self-reported sub-

Table 1 Frequencies of individual items and their factor structure. Parent report

	Parent Report			
	Frequency (%)	Exuberant	Under-controlled	Difference
irritable	18.01	0.159	**0.786**	−0.627
distractible	21.27	0.226	**0.846**	−0.62
risk taking	10.81	0.210	**0.822**	−0.612
less self control	15.44	0.249	**0.825**	−0.576
poor concentration	27.79	0.351	**0.822**	−0.471
invades space	18.87	0.289	**0.750**	−0.461
bossy	19.04	0.209	**0.665**	−0.456
not concerned	20.24	0.283	**0.737**	−0.454
over-sexed	4.12	0.236	**0.600**	−0.364
constant changes	13.55	0.299	**0.634**	−0.335
unkempt	7.72	0.189	0.481	−0.292
flight of ideas	20.58	0.308	**0.592**	−0.284
talks to strangers	10.46	0.280	**0.561**	−0.281
overconfident	22.3	0.370	**0.581**	−0.211
overspends	26.59	0.228	0.369	−0.141
sleeps less	13.21	0.360	0.436	−0.076
restlessness	28.82	**0.527**	0.563	−0.036
exciteable	37.91	**0.644**	0.498	0.146
full of energy	58.83	**0.583**	0.379	0.204
gets more done	14.92	**0.532**	0.308	0.224
noisier	57.46	**0.675**	0.447	0.228
fast talk	34.48	**0.681**	0.370	0.311
active	43.74	**0.745**	0.362	0.383
outgoing	35.51	**0.838**	0.202	0.636
jokes	42.02	**0.865**	0.191	0.674
cheerful	35.68	**0.740**	0.056	0.684

Table 2 Frequencies of individual items and their factor structure. Self report

	Self Report			
	Frequency (%)	Exuberant	Under-controlled	Difference
irritable	4.23	0.118	**0.685**	−0.567
risk taking	6.64	0.229	**0.725**	−0.496
bossy	3.89	0.212	**0.680**	−0.468
less self control	9.38	0.317	**0.755**	−0.438
poor concentration	20.59	0.322	**0.759**	−0.437
distractible	22.31	0.308	**0.654**	−0.346
not concerned	20.02	0.332	**0.590**	−0.258
overconfident	7.67	0.360	**0.586**	−0.226
invades space	5.03	0.365	**0.498**	−0.133
restlessness	27.46	0.433	**0.566**	−0.133
flight of ideas	16.93	0.395	**0.520**	−0.125
unkempt	10.64	0.334	0.438	−0.104
overspends	18.65	0.334	0.412	−0.078
constant changes	9.38	0.468	0.484	−0.016
sleeps less	18.99	0.407	0.405	0.002
fast talk	20.02	0.499	0.385	0.114
talks to strangers	16.7	0.468	0.325	0.143
over-sexed	3.66	**0.524**	0.365	0.159
noisier	46.91	**0.577**	0.414	0.163
full of energy	58.35	**0.664**	0.374	0.290
gets more done	15.1	0.480	0.172	0.308
active	40.73	**0.714**	0.309	0.405
exciteable	36.84	**0.738**	0.310	0.428
jokes	47.25	**0.774**	0.289	0.485
outgoing	42.68	**0.724**	0.229	0.495
cheerful	45.08	**0.692**	0.154	0.538

scales and impact: Only episodic under-control ($b=0.19$; 95% CI 0.14–0.23), but not episodic exuberance ($b=−0.02$; 95% CI −0.05–0.02) was a significant predictor of impact. The predictive power of episodic under-control remained significant even after adjusting for the presence of DSM-IV disorders or dimensional measures of psychopathology (see supplementary results).

Latent Classes of Individuals Based on Mania-Like Symptom Profiles

For both parent and self report, a 3-class solution of the LCA models was chosen on the basis of fit indices (see supplementary Tables 1 and 2) and parsimony. The results of the two LCA models are presented in Fig. 1a and b for

Table 3 Associations of the two parent-reported sub-scales with DSM-IV disorders

	Sub-Scale	Any Disorder	Emotional	CD/ODD	ADHD
Univariate	Under-control	**1.41** (1.31–1.52)	**1.15** (1.06–1.25)	**1.46** (1.35–1.57)	**1.45** (1.31–1.61)
	Exuberance	**1.19** (1.11–1.27)	1.08 (0.99–1.18)	**1.21** (1.12–1.31)	**1.28** (1.12–1.45)
Multivariate	Under-control	**1.44** (1.32–1.57)	1.10 (1.00–1.21)	**1.54** (1.40–1.71)	**1.50** (1.32–1.70)
	Exuberance	0.96 (0.88–1.05)	1.11 (0.99–1.25)	0.89 (0.79–1.01)	0.93 (0.79–1.11)

Odds ratios and 95% confidence intervals from logistic regression models are shown. Univariate refers to models were only one of the sub-scales (either under-control *or* exuberance) were entered and multivariate refers to models were both (under-control *and* exuberance) were simultaneously entered as predictors in the model. All models co-varied for age and gender. Any disorder stands for any DSM-IV disorder. CD/ODD = conduct disorder or oppositional defiant disorder; ADHD = attention deficit/hyperactivity disorder. Results that are statistically significant at the protected level ($p<0.001$) are shown in bold

Table 4 Associations of the two self-reported sub-scales with psychopathology

	Sub-Scale	Any Disorder	Emotional	CD/ODD	ADHD
Univariate	Under-control	**1.36** (1.25–1.49)	**1.32** (1.20–1.47)	**1.36** (1.22–1.52)	1.35 (1.1–1.64)
	Exuberance	**1.12** (1.03–1.22)	1.13 (1.02–1.26)	1.14 (1.01–1.29)	1.21 (0.95–0.54)
Multivariate	Under-control	**1.46** (1.30–1.63)	**1.38** (1.20–1.58)	**1.44** (1.24–1.66)	1.36 (1.04–1.77)
	Exuberance	0.90 (0.81–1.01)	0.94 (0.82–1.08)	0.91 (0.78–1.07)	0.99 (0.72–1.35)

Odds ratios and 95% confidence intervals from logistic regression models are shown. Univariate refers to models were only one of the sub-scales (either under-control *or* exuberance) were entered and multivariate refers to models were both (under-control *and* exuberance) were simultaneously entered as predictors in the model. All models co-varied for age and gender. Any disorder stands for any DSM-IV disorder. CD/ODD = conduct disorder or oppositional defiant disorder; ADHD = attention deficit/hyperactivity disorder. Results that are statistically significant at the protected level ($p<0.001$) are shown in bold

parent and self reported-items respectively. For ease of interpretation, the horizontal axis lists mania-like symptoms according to their relative loading on the two previously identified factors, i.e., based on each item's loading on episodic under-control minus its loading on episodic exuberance. Thus the items on the left are primarily exuberant and those on the right primarily under-controlled, with a gradient in between. On the vertical-axis, we plotted for each of the identified latent classes the adjusted probability of scoring "A lot" on each listed item to take account of the fact some items were more common than others. We did this by expressing the probabilities of

Fig. 1 **a** Latent classes for bipolar symptoms in the community parent report. **b** Latent classes for bipolar symptoms in the community self report. The items are presented on the horizontal axis ordered according to whether they were predominantly loading on the under-control or exuberance factor (based on the difference score between the loadings). The vertical axis represents the probability of scoring "a lot" (normalised by the average probability acrosss classes) for a given item conditional upon membership to one of there classes (each indicated by different colours). "interm" is an abbreviation for intermediate

a Latent Classes for Bipolar Symptoms in the Community parent report
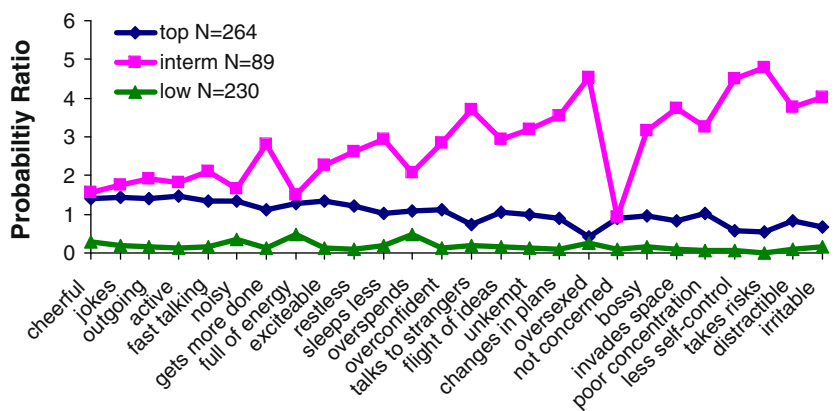
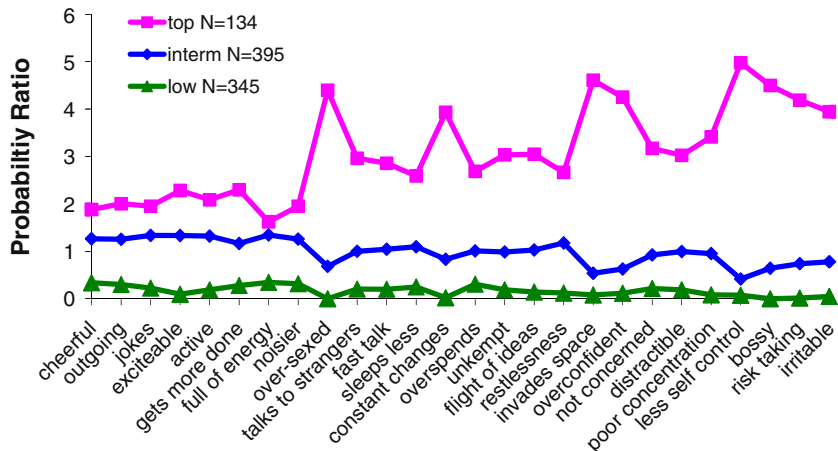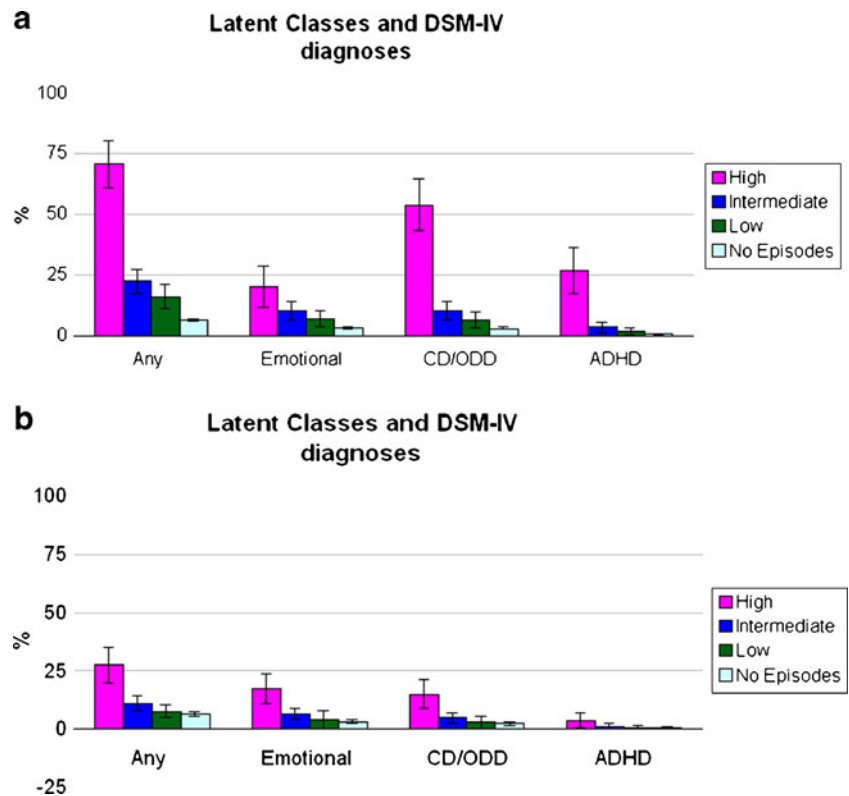**b** Latent Classes for Bipolar Symptoms in the Community self report

Fig. 2 **a** Parent reported latent classes and their association with psychiatric disorders. **b** Self reported latent classes and their association with psychiatric disorders


each item as ratios to the average probability for that item in the sample as a whole.

As shown in Fig. 1a and b, reflecting both reporting sources, a small class of people emerged with high probability of scoring "A lot" on all items across the episodic under-control and exuberance sub-scales; we termed this class "top". A larger class of subjects scoring low on all items emerged by both reporting sources—we designated this group "low". The largest class, which we designated "intermediate" was made up of individuals with probability ratios that lay between

the top and bottom classes. By parent report, 1.7% of the total sample of 8–19 year-olds were in the top latent class. The corresponding figure for self report was 2.5%.

The agreement between parent and self reported classes was low but significant ($\kappa=0.19$, $p<0.0001$). By parent report, the ANOVA for age was significant overall ($F=4.5$, df=2; $p<0.05$); children in the "low" class (13.6 years, sd=3.3) were older compared to those in the "top" (12.6 years, sd=2.8, Tukey HSD = 3.9)) but not those in the intermediate (12.9 years, sd=3.1,

Fig. 3 A four-classes solution for the LCA by parent report.
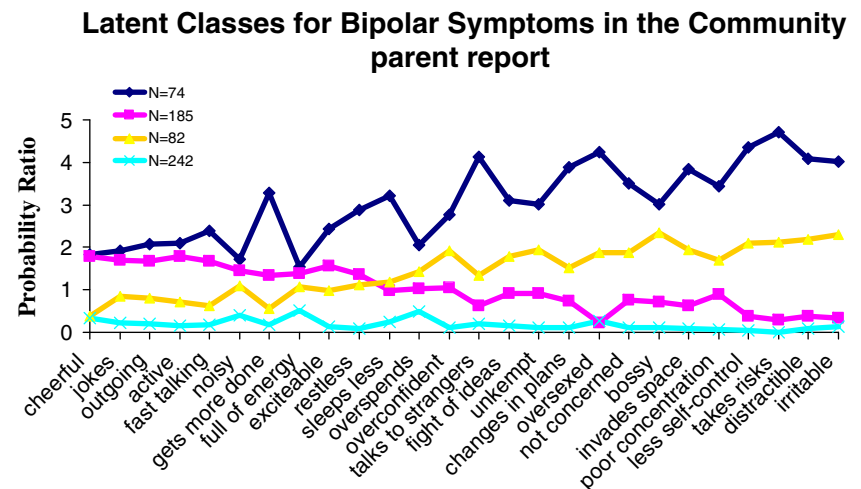
Tukey HSD = 1.1) classes, and the intermediate and low classes did not differ significantly from each other (Tukey HSD = 2.76). By self report, there were no significant age differences (overall ANOVA: $F=1.2$, df=2; $p>0.05$) between classes (top (14.6 years, sd=2.3), intermediate (14.9 years, sd=2.3) and low (14.8 years, sd=2.3)).

By parent report, the top class contained significantly more boys (72%) compared to the intermediate (48%; $\chi^2=15.2$, df=1, $p<0.001$) and low class (50%; $\chi^2=13.0$, df=1, $p<0.001$), but the difference between the intermediate and low classes was not significant ($\chi^2=0.1$, df=1, $p=0.74$). By self report, there were no significant differences in the proportion of boys between the top (40%), intermediate (47%) and low (43%) classes (all $\chi^2<2$, $p>0.05$).

External Validation of the Latent Classes

The relationship of each class to comorbid psychopathology is shown in Fig. 2a and b, for parent and self report respectively. By both reporting sources, "top" showed the strongest associations with psychopathology. For both reporting sources, the top category generally differed significantly from both other categories, whereas there was no significant difference between "low" and "intermediate" with regards to relationships with psychopathology. The only exception to this rule was that the higher rate of ADHD in the top self-report category was not statistically significant.

Similar results for both reporting sources were obtained when predicting to social impairment rather than comorbidity. For parent report the ANOVA for the mean parent rated SDQ-impact score across classes showed significant differences (overall ANOVA: $F=54.1$, df=2, $p<0.001$); in post-hoc tests, the top class had significantly higher scores (3.0, sd=2.6) compared to the intermediate (1.1, sd=1.8, Tukey HSD = 13.6) and low (0.7, sd=1.5, Tukey HSD = 15.8) classes, while the intermediate and low classes did not differ significantly from each other in mean SDQ-impact score (Tukey HSD = 2.2). For self report, the ANOVA for the self-reported mean SDQ-impact score across classes showed significant differences (overall ANOVA: $F=29.6$, df=2, $p<0.001$); in post-hoc tests the mean SDQ-impact score of the top class was significantly higher (1.1, sd=1.6) compared to the intermediate (0.5, sd=0.1; Tukey HSD = 26.5) and low (0.3, sd=0.8, Tukey HSD = 11.9) classes and the difference was significant; in this instance the self-reported intermediate and low classes differed significantly from each other(Tukey HSD = 38.4). Similar results for both reporting sources were obtained when using multinomial logistic regression models to test the association to impact: The top class was associated with significantly higher social impairment compared to the other two classes

even after adjusting for the presence of psychiatric disorders (see supplementary results).

In our factor analyses we identified an episodic exuberance score which did not seem to have independent associations with psychopathology. We tried to establish whether other solutions of the LCA would identify a group of children who would score highly only on symptoms of exuberance. We found that by parent report (but not self report) there was a 4-class solution in which one class of children scored high predominantly on symptoms of episodic exuberance, as opposed to episodic under-control (Fig. 3). However, this "pure exuberance" class did not differ from the low class in terms of comorbidity or impact. Given this further evidence that identifying episodic exuberance does not add to predictive power, we retained our focus on the 3-class LCA.

Discussion

Answering this paper's first question, our results show that using a single screening question for episodes of elated mood identifies a group of children with high rates of psychiatric disorder. By parent report, this group has a fourfold higher rate than those screening negative; by youth report the increase is twofold over those screening negative. Moreover, a positive response to the screening question predicts impact over and above that attributable to other psychopathology, and this was true for both reporting sources.

The second question of this paper was whether it would be possible to identify meaningful groupings of symptoms occurring within episodes and whether such groupings would have differential predictions to psychopathology and social impairment. By either reporting source, exploratory factor analysis identified two dimensions that we termed episodic under-control and episodic exuberance. Although each of these dimensions was a significant univariate predictor of psychopathology for most outcomes, episodic exuberance was no longer a significant predictor once adjustment had been made for the dimension of episodic under-control. By contrast, episodic under-control remained a significant predictor of psychopathology after adjusting for episodic exuberance, i.e., only episodic under-control was independently associated with psychiatric disorders. Moreover, only episodic under-control predicted social impairment over and above other psychopathology. The pattern of the findings was similar across informants.

The third question of the paper was whether it would be possible empirically to derive groups defined by symptom patterns. Latent class analysis identified three groups among those screening positive for episodes of elated mood. These groups, which we termed top, intermediate, and low, appeared to differ on a severity continuum, in that

they showed a parallel shift on the probability scale and no class was characterised by a differentially high probability for specific items. The three classes also differed in their associations with markers of severity. The associations between the top group and psychiatric disorder were striking: By parent report over 70% and by self report over 25% of those in the top groups had another psychiatric diagnosis. By either reporting source, the top group was significantly associated with social impairment, even adjusting for psychiatric disorders.

What are we to make of the symptoms that we have termed episodic under-control? Several of the symptoms in this dimension are present, in their non-episodic form, in externalising disorders (distractibility, intrusiveness, risk taking, irritability). We see two possible interpretations: According to the first, the episodic under-control merely reflects symptoms of the underlying externalising disorders, rather than bipolar disorders; the alternative interpretation is that episodic under-control reflects genuine manic symptoms, with the children in the top group suffering from bipolar disorder.

It is important to reflect carefully on episodicity when trying to decide between these alternatives. The DSM-IV (APA 2000) and the ICD-10 (World Health 1994) criteria for a (hypo) manic episode require the presence of "a distinct period of abnormally and persistently elevated, expansive, or irritable mood". Moreover, the need for attention to episodicity has been particularly emphasised in conceptual reviews of bipolar disorder in children and adolescents (Leibenluft et al. 2003). By contrast, episodicity does not feature as a diagnostic criterion for some of the most characteristic conditions that child psychiatrists deal with, such as ASD, CD/ODD, ADHD, and anxiety disorders. Indeed, persistent symptoms since early life is a requirement for ASD (onset in early development) and ADHD (onset before age 7), while a diagnosis of ODD or CD requires that behavioural problems have been present for a minimum of 6 months. In this study, we asked parents and children about symptoms that occurred during distinct episodes of "going high". Therefore, our findings would seem to fit more readily within the framework of the distinct episodes stipulated by the criteria for (hypo) mania than within the framework of the persistent symptomatology stipulated for externalizing disorders.

However, a possible interpretation is that what parents or young people describe when they are asked about an episode of going high are simply exacerbations in the severity of the symptoms of a chronic externalizing disorder. Indeed, the lack of a requirement for episodicity does not mean that there is no fluctuation over time in the symptom severity of, say ADHD or CD/ODD. Parents, teachers and clinicians are all well aware of children with behavioural problems whose symptoms flare up for brief periods of time, often adding to the turmoil experienced by

the child and those around the child,. For example, increased demands at school or at home may lead to a transient exacerbation of irritable behaviours. Following this interpretation, children with episodic increases in their symptoms of externalizing disorders would not be suffering from BP. This is not to say that they would be immune to progression to BP—a progression that is reportedly more likely for any child with behavioural problems in general (Kim-Cohen et al. 2003).

A related, but not identical, interpretation of fluctuating externalizing symptoms is that there are subtypes of behavioural disorders characterised by higher symptom variability. For example, while one of the hallmarks of classical ODD may be chronic irritability (Stringaris and Goodman 2009a, 2009b), rarer subforms may be characterised by fluctuations superimposed on the chronic course. Proneness to variability in symptoms would be specific to this subgroup, perhaps due to genes shared in common with bipolar disorder. Accordingly, the risk for progression to bipolar disorder would be higher than the rate for non-variable behavioural problems.

While these interpretations favour assimilating episodic under-control into the externalizing disorders, a plausible alternative is that parents and children are indeed able to report episodes of distinct mood that are different from fluctuations in the symptoms of coexisting externalizing disorders. As previously described (Stringaris et al. 2010b), there were only a few cases of BPI or BPII in this sample, in keeping with other epidemiologic studies in children (Costello et al. 1996). This low prevalence is in contrast to the more common occurrence of these conditions in adults; for example, in a recently conducted US community-based study (Merikangas et al. 2007) the 12-month prevalence of BP-I was 0.8% and that for BP-II 1.4%. It should, however, also be noted that factors other than episode duration, such as number of symptoms (Lewinsohn et al. 2003) or the endorsement of irritability as a Criterion A (Geller et al. 2007) can affect prevalence rates of mania in children. Moreover, rates seem to be higher in adolescents than in children and a recent meta-analytic review suggests that community prevalences of pediatric bipolar disorder may have been under-appreciated (Van Meter et al. 2011). In any case, the low prevalence of classical bipolar in community samples precludes using typical cases to validate the episodic under-control dimension or the latent classes against these few cases. The vast majority of those in the latent classes would fall short of the duration criterion and therefore not qualify for BPI or BPII. However, previous research in referred samples has shown that young people experiencing short (<4 days) episode durations, termed BP-NOS, progressed to BPI or BPII at a rate of 38% over 4-year follow up (Birmaher et al. 2009). Moreover, evidence from the literature on adults (Angst et al. 2003) suggests

that short episodes (1–3 days) of mania-like symptoms do not differ significantly from episodes of typical duration (>4 days) on a number of external validators. Indeed, it has been pointed out (Dunner 1998) that in the DSM-IV "the duration of a hypomanic episode was *arbitrarily* (added italics) set at 4 days or longer" (page 190). Another potentially relevant finding from the work of Angst and colleagues (Angst et al. 2003), is that their group of "pure hypomania" (which does not conform to typical DSM-IV criteria for BP), showed the highest rates of conduct problems across all subgroups of mood disorders. Accordingly, symptoms that appear externalising might be a consequence of risk-taking and disinhibited behaviours occurring during a hypomanic episode.

In conclusion, this study shows that episodic changes in mood occur frequently, and that meaningful symptom dimensions and latent classes can be identified during such episodes. Furthermore, if our findings are replicated, then operationalized diagnostic criteria in future might need to focus on undercontrolled symptoms since these, and not exuberant symptoms, were the independent predictors of impairment in our study. Future studies, validating the proposed dimensions and clusters derived in this study against longitudinal outcomes and genetic markers would be particularly valuable.

References

Agresti, A. (2002). *Categorical data analysis*. Hoboken: Wiley Interscience.

American Psychiatric, A. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington: American Psychiatric Press.

Angst, J., Gamma, A., Benazzi, F., Ajdacic, V., Eich, D., & Rossler, W. (2003). Toward a re-definition of subthreshold bipolarity: epidemiology and proposed criteria for bipolar-II, minor bipolar disorders and hypomania. *Journal of Affective Disorders, 73*(1–2), 133–146.

APA. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington: American Psychiatric Press.

Ayer, L., Althoff, R., Ivanova, M., Rettew, D., Waxler, E., Sulman, J., et al. (2009). 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, 50*(10), 1291–1300. doi:10.1111/j.1469-7610.2009.02089.x.

Biederman, J. (2006). The evolving face of pediatric mania. *Biological Psychiatry, 60*(9), 901.

Birmaher, B., Axelson, D., Strober, M., Gill, M. K., Valeri, S., Chiappetta, L., et al. (2006). Clinical course of children and adolescents with bipolar spectrum disorders. *Archives of General Psychiatry, 63*(2), 175.

Birmaher, B., Axelson, D., Goldstein, B., Strober, M., Gill, M. K., Hunt, J., et al. (2009). Four-year longitudinal course of children and adolescents with bipolar spectrum disorders: the course and outcome of bipolar youth (COBY) study. *The American Journal of Psychiatry, 166*(7), 795–804. doi:10.1176/appi.ajp.2009.08101569.

Blader, J. C., & Carlson, G. A. (2007). Increased rates of bipolar disorder diagnoses among U.S. child, adolescent, and adult inpatients, 1996–2004. *Biological Psychiatry, 62*(2), 107.

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The strengths and difficulties questionnaire: U.S. normative data and psychometric properties. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(6), 557.

Brotman, M. A., Rich, B. A., Guyer, A. E., Lunsford, J. R., Horsey, S. E., Reising, M. M., et al. (2010). Amygdala activation during emotion processing of neutral faces in children with severe mood dysregulation versus ADHD or bipolar disorder. *The American Journal of Psychiatry, 167*(1), 61–69. doi:10.1176/appi.ajp.2009.09010043.

Carlson, G. A., & Meyer, S. E. (2006). Phenomenology and diagnosis of bipolar disorder in children, adolescents, and adults: complexities and developmental issues. *Development and Psychopathology, 18*(4), 939–969.

Carlson, G. A., & Glovinsky, I. (2009). The concept of bipolar disorder in children: a history of the bipolar controversy. *Child Adolesc Psychiatr Clin N Am, 18*(2), 257–271. doi:10.1016/j.chc.2008.11.003. vii.

Costello, E. J., Angold, A., Burns, B. J., Stangl, D. K., Tweed, D. L., Erkanli, A., et al. (1996). The great smoky mountains study of youth. goals, design, methods, and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry, 53*(12), 1129.

Dunner, D. L. (1998). Bipolar disorders in DSM-IV: impact of inclusion of rapid cycling as a course modifier. *Neuropsychopharmacology, 19*(3), 189–193. doi:10.1016/S0893-133X(98)00024-4.

Faraone, S. V., Althoff, R. R., Hudziak, J. J., Monuteaux, M., & Biederman, J. (2005). The CBCL predicts DSM bipolar disorder in children: a receiver operating characteristic curve analysis. *Bipolar Disorders, 7*(6), 518–524. doi:10.1111/j.1399-5618.2005.00271.x.

Ford, T., Goodman, R., & Meltzer, H. (2003). The British child and adolescent mental health survey 1999: the prevalence of DSM-IV disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(10), 1203.

Geller, B., Tillman, R., & Bolhofner, K. (2007). Proposed definitions of bipolar I disorder episodes and daily rapid cycling phenomena in preschoolers, school-aged children, adolescents, and adults. *Journal of Child and Adolescent Psychopharmacology, 17*(2), 217.

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337.

Goodman, R., & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: is small beautiful? *Journal of Abnormal Child Psychology, 27*(1), 17.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 41*(5), 645–655.

Green, H., McGinnity, A., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in great Britain, 2004*. London: The Stationery Office.

Guyer, A. E., McClure, E. B., Adler, A. D., Brotman, M. A., Rich, B. A., Kimes, A. S., et al. (2007). Specificity of facial expression labeling deficits in childhood psychopathology. *Journal of Child Psychology and Psychiatry, 48*(9), 863–871. doi:10.1111/j.1469-7610.2007.01758.x.

Holtmann, M., Bolte, S., Goth, K., Dopfner, M., Pluck, J., Huss, M., et al. (2007). Prevalence of the child behavior checklist-pediatric bipolar disorder phenotype in a German general population sample. *Bipolar Disorders, 9*(8), 895–900.

Holtmann, M., Buchmann, A. F., Esser, G., Schmidt, M. H., Banaschewski, T., & Laucht, M. (2010). The Child Behavior Checklist-Dysregulation Profile predicts substance use, suicidality, and functional impairment: a longitudinal analysis. *J Child Psychol Psychiatry.*10.1111/j.1469-7610.2010.02309.x

Hudziak, J. J., Althoff, R. R., Derks, E. M., Faraone, S. V., & Boomsma, D. I. (2005). Prevalence and genetic architecture of child behavior checklist-juvenile bipolar disorder. *Biological Psychiatry, 58*(7), 562–568. doi:10.1016/j.biopsych.2005.03.024.

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H., Milne, B. J., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry, 60*(7), 709–717.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. New York: Houghton Mifflin.

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

Leibenluft, E., Charney, D. S., Towbin, K. E., Bhangoo, R. K., & Pine, D. S. (2003). Defining clinical phenotypes of juvenile mania. *The American Journal of Psychiatry, 160*(3), 430.

Lewinsohn, P. M., Seeley, J. R., & Klein, D. N. (2003). Bipolar disorder in adolescents: epidemiology and suicidal behavior. In B. Geller & M. P. DelBello (Eds.), *Bipolar disorder in childhood and early adolescence*. New York: Guilford Press.

Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M., Petukhova, M., et al. (2007). Lifetime and 12-month prevalence of bipolar spectrum disorder in the national comorbidity survey replication. *Archives of General Psychiatry, 64*(5), 543–552. doi:10.1001/archpsyc.64.5.543.

Meyer, S. E., Carlson, G. A., Youngstrom, E., Ronsaville, D. S., Martinez, P. E., Gold, P. W., et al. (2009). Long-term outcomes of youth who manifested the CBCL-pediatric bipolar disorder phenotype during childhood and/or adolescence. *Journal of Affective Disorders, 113*(3), 227–235. doi:10.1016/j.jad.2008.05.024.

Moreno, C., Laje, G., Blanco, C., Jiang, H., Schmidt, A. B., & Olfson, M. (2007). National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. *Archives of General Psychiatry, 64*(9), 1032–1039.

Muthen, B. (2001). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumaker (Eds.), *New developments and techniques in structural equation modeling*. Hillsdale: Lawrence Erlbaum Associates.

Muthen, L. K., & Muthen, B. O. (2007). *Mplust User's guide (5th* (editionth ed.). Los Angeles: Muthen & Muthen.

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling, 14*, 535–569.

Parry-Langdon, N. (2008). Three years on: Survey of the emotional development and well-being of children and young people. *Office for National Statistics*. Newport: The Stationery Office.

Potegal, M., Carlson, G. A., Margulies, D., Basile, J., Gutkovich, Z. A., & Wall, M. (2009). The behavioral organization, temporal characteristics, and diagnostic concomitants of rage outbursts in child psychiatric inpatients. *Current Psychiatry Reports, 11*(2), 127–133.

Rich, B. A., Grimley, M. E., Schmajuk, M., Blair, K. S., Blair, R. J., & Leibenluft, E. (2008). Face emotion labeling deficits in children with bipolar disorder and severe mood dysregulation. *Development and Psychopathology, 20*(2), 529–546. doi:10.1017/S0954579408000266.

Stringaris, A., & Goodman, R. (2009a). Longitudinal outcome of youth oppositionality: irritable, headstrong, and hurtful behaviors have distinctive predictions. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 404–412. doi:10.1097/CHI.0b013e3181984f30.

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

Stringaris, A., Baroni, A., Haimm, C., Brotman, M., Lowe, C. H., Myers, F., et al. (2010a). Pediatric bipolar disorder versus severe mood dysregulation: risk for manic episodes on follow-Up. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*, 397–405.

Stringaris, A., Santosh, P., Leibenluft, E., & Goodman, R. (2010b). Youth meeting symptom and impairment criteria for mania-like episodes lasting less than four days: an epidemiological enquiry. *Journal of Child Psychology and Psychiatry, 51*, 31–38.

Van Meter, A. R., Moreira, A.L., Youngstrom,E.A. (2011). Meta-analysis of epidemiological studies of pediatric bipolar disorder. *J Clin Psychiatry, in press*

Volk, H. E., & Todd, R. D. (2007). Does the child behavior checklist juvenile bipolar disorder phenotype identify bipolar disorder? *Biological Psychiatry, 62*(2), 115–120.

World Health, O. (1994). *ICD-10: classification of mental and behavioural disorders*. Geneva: World Health Organization.

Wozniak, J., Biederman, J., Kiely, K., Ablon, J. S., Faraone, S. V., Mundy, E., et al. (1995). Mania-like symptoms suggestive of childhood-onset bipolar disorder in clinically referred children. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*(7), 867–876.

Youngstrom, E. A., Birmaher, B., & Findling, R. L. (2008a). Pediatric bipolar disorder: validity, phenomenology, and recommendations for diagnosis. *Bipolar Disorders, 10*(1 Pt 2), 194–214. doi:10.1111/j.1399-5618.2007.00563.x.

Youngstrom, E. A., Frazier, T. W., Demeter, C., Calabrese, J. R., & Findling, R. L. (2008b). Developing a 10-item mania scale from the parent general behavior inventory for children and adolescents. *The Journal of Clinical Psychiatry, 69*(5), 831–839.