



Published in final edited form as:

J Clin Child Adolesc Psychol. 2012 May ; 41(3): 261–274. doi:10.1080/15374416.2012.654465.

Understanding the Phenotypic Structure of Adult Retrospective ADHD Symptoms During Childhood in the United States

Krista W. Ranby,

Center for Child and Family Policy, Duke University

Marcella H. Boynton,

Center for Child and Family Policy, Duke University

Scott H. Kollins,

Department of Psychiatry and Behavioral Sciences, Duke University

F. Joseph McClernon,

Department of Psychiatry and Behavioral Sciences, Duke University Medical Center and Durham VAMC Mental Illness Research, Education, and Clinical Center

Chongming Yang, and

Center for Child and Family Policy, Duke University

Bernard F. Fuemmeler

Department of Community and Family Medicine, Duke University Medical Center and Department of Psychology and Neuroscience, Duke University

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a highly heterogeneous disorder, and the phenotypic structure comprising inattentive and hyperactive-impulsive type symptoms has been the focus of a growing body of recent research. Methodological studies are needed to better characterize phenotypes to advance research as well as clinical practice. A large U.S. population-based sample of young adults ($N = 14,307$, aged 17–28 years, 52.8% female) retrospectively reported their experiences of childhood ADHD symptoms. Factor analysis, latent class analysis, and factor mixture modeling of ADHD symptoms were compared to determine which underlying structure best fit the data. Fit statistics as well as substantive criteria compared models within and across model subtypes. Analyses supported a two-factor two-class structure for both male and female subjects. The two latent factors represented inattentive and hyperactive-impulsive symptom dimensions. The two latent classes divided people into a smaller affected class and a larger unaffected class. Individuals who reported having been diagnosed with ADHD were more likely to be in the affected class (*OR* male subjects = 4.03, 95% CI [2.65, 6.13]; *OR* female subjects = 5.65, 95% CI [3.15, 10.10]). This work aids in the understanding of ADHD symptomatology within the population; a majority of people experience very low symptom severity, whereas a minority of people experience high symptom severity. Within this high symptom group, however, variability in symptom experiences exists. Empirical models can be helpful in clarifying ADHD phenotypic structure that has the potential to advance research on the etiology and consequences of ADHD symptoms.

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common childhood disorders, affecting approximately 9% of children in the United States (Froehlich et al., 2007). ADHD persists into adulthood, affecting 4.4% of U.S. adults, and the symptoms can lead to significant impairments throughout the lifespan (Kessler et al., 2006). Those with the disorder are at increased risk for substance use problems, conduct behavior problems, academic and/or occupational underachievement, interpersonal difficulties, automobile accidents, arrests, and incarceration (Barkley & Cox, 2007; Mannuzza, Klein, Bessler, Malloy, & Hynes, 1997; Mannuzza, Klein, Bessler, Malloy, & LaPadula, 1993, 1998; Mannuzza, Klein, & Moulton, 2008). Further, the cost of ADHD to society is substantial, with estimates ranging from \$30 to \$50 billion annually in the United States, in addition to the considerable costs in terms of individual and familial distress (Pelham, Foster, & Robb, 2007).

ADHD is a highly heterogeneous disorder, due in part to the diagnostic criteria used to define the condition. Current nosology defines ADHD on the basis of relative presentation of symptoms in two broad domains: inattention (IN) and hyperactive-impulsive (HI; American Psychiatric Association [APA], 2000). Subtypes of ADHD, defined by these relative symptom presentations, include predominantly inattentive type, predominantly hyperactive-impulsive type, and combined type. In addition to the symptomatic heterogeneity observed across these subtypes, there is evidence that ADHD subtypes exhibit different neurocognitive correlates as well as differential risk for adverse outcomes, such as additional psychopathology. For example, Eiraldi, Power, and Nezu (1997) found that children with ADHD combined subtype were more likely to be diagnosed with oppositional defiant disorder and conduct disorder than were children with ADHD inattentive subtype. Solanto et al. (2007) found that children with ADHD combined subtype were more impulsive than were children with ADHD inattentive subtype. Willcutt, Pennington, Chhabildas, Friedman, and Alexander (1999) found that among children and adolescents, ADHD inattentive and combined subtypes were more strongly associated with experiencing symptoms of depression than was ADHD hyperactive-impulsive subtype.

Whereas the approach of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev. [DSM-IV-TR]; APA, 2000) to defining ADHD subtypes has guided clinical practice for a number of years, considerable recent work has been devoted to better characterizing the underlying phenotypic characteristics of the disorder. For example, theoretically driven work has suggested that individuals with ADHD may be more precisely characterized by deficits in several fundamental processes, such as inhibitory control, delay aversion, or other related executive functions (Barkley, 1997; Castellanos & Tannock, 2002; Sonuga-Barke, Bitsakou, & Thompson, 2010). Whereas these findings provide valuable insights, others have suggested that empirically driven methods that identify the latent structure of symptoms and elucidate phenotypically distinct subtypes of ADHD may also be useful (Hudziak et al., 1998). Recent studies have begun to examine ADHD phenotypes (Lubke, Hudziak, Derks, van Bijsterveldt, & Boomsma, 2009; Lubke et al., 2007); however, it remains unclear within a nationally representative U.S. population-based sample whether quantitative symptom severity, qualitatively distinct subtypes, or a combination of the two best characterize unique experiences of ADHD symptoms.

Traditionally, two empirically driven approaches that have been used to characterize the latent structure of ADHD symptoms and identify heterogeneous subgroups of individuals with ADHD symptoms are factor analysis (FA) and latent class analysis (LCA), respectively. FA is a frequently employed estimation strategy for examining the latent variable structure of measures designed to tap one or more associated constructs and has been used to examine the factor structure of measures of DSM-IV-TR ADHD symptoms (e.g., Amador-Campos, Forns-Santacana, Martorell-Balanza, Guardia-Olmos, & Pero-

Cebollero, 2005, 2006; Fergusson, Horwood, & Lynskey, 1994; Gomez, Burns, Walsh, & Hafetz, 2005; Hardy et al., 2007; Martel, von Eye, & Nigg, 2010; Wolraich, Lambert, Baumgaertel, & Garcia-Tornel, 2003). A recent innovation in this line of work using a FA approach showed that a bifactor model that included both a general and a specific factor was superior to one-, two-, and three-factor models (Martel, et al., 2010). This study, along with others (e.g., Hardy et al., 2007), provides support for the idea that individuals with ADHD share some phenotypic features (i.e., a single general factor) but also demonstrate heterogeneity that usually is described by separate factors of inattention and hyperactive-impulsivity. LCA is another approach for describing the heterogeneity of the ADHD phenotype and has been used to identify distinct classes of individuals with ADHD symptoms (Neuman et al., 2005; Rasmussen et al., 2002, 2004; Todd et al., 2001; Volk, Neuman, & Todd, 2005). For instance, Todd et al. (2001), using LCA, found evidence for three mild and three severe classes of ADHD. The three severe classes overlapped with the *DSM-IV-TR* diagnostic criteria: primarily inattentive, primarily hyperactive-impulsive, and combined. Similarly, Volk et al. (2005) identified seven classes; severe types were represented by inattentive and combined symptoms. The LCA approach has been criticized, however, because it may overestimate the number of classes in a given sample. If systematic severity in symptoms differ within a class, an LCA model with a greater number of classes might be retained because there is no other means (e.g., the use of factors) to model these severity differences (Lubke & Neale, 2006).

A hybrid of the FA and LCA approaches is factor mixture modeling (FMM). FMM models combine latent factors of FA and latent classes of LCA with an assumption that there are ADHD symptom subtypes while allowing for differences in severity within each subtype. These different modeling techniques (FA, LCA, and FMM) have typically been used in isolation without comparing the various results produced by the different approaches. Exceptions have been recent studies conducted by Lubke and colleagues (Lubke et al., 2009; Lubke et al., 2007). These studies demonstrate that the symptoms of ADHD might best be viewed as a dimensional condition with at least two correlated continua (inattentiveness and hyperactive-impulsive) and that the disorder of ADHD falls at the extreme end of these continua. This work, which employed primarily Northern European samples, is useful for better understanding how the full range of ADHD symptoms might be reflected in a given population.

The current work extends earlier research in several critical ways. First, a comparison of FA, LCA, and FMM models of ADHD self-reported symptom data has not been conducted on a U.S. population-based sample. Understanding ADHD symptom prevalence and co-occurrence within a population-based sample is important given evidence that persons with mental disorders who seek services from clinics are not representative of all persons with psychological symptoms in the population. Goodman et al. (1997), for instance, found that a clinic sample was more impaired, more likely to have comorbid disorders, more likely to be non-Hispanic White than other ethnic groups, and more likely to be male. The large, nationally representative U.S. sample that we employ herein allows for a test of the structure of self-reported ADHD symptoms in the United States and provides a comparison with models of population-based Northern European samples used by Lubke and colleagues (Lubke et al., 2009; Lubke et al., 2007). This is important because the symptoms of ADHD may be interpreted differently or the disorder may carry different meaning in different cultures (Bauermeister, Canino, Polancayk, & Rohde, 2010; Singh, 2011). Second, previous studies using similar methodologies have not related the mixture results to reported or confirmed ADHD diagnosis. Such approaches lend support for establishing meaningful subgroup differences. To aid in assessing convergent validity, we examined the extent to which the empirically derived mixture results were related to self-reported history of an ADHD diagnosis. Although ADHD symptoms have been modeled using a FMM approach,

this technique is still quite new and has been employed in a very limited number of studies. Tests of different samples are needed to best understand whether ADHD symptoms in the population exist along a continuum, whether ADHD symptoms are only observed in an “affected” minority, or whether a hybrid of these two possibilities exists. A better understanding of the phenotypic structure is especially helpful for research that examines the etiology and consequences of ADHD symptomatology.

METHODS

Sample

The study population was drawn from 20,747 adolescents in The National Longitudinal Study of Adolescent Health (Add Health), a nationally representative study of adolescents. Participants were recruited for initial data collection in 1995 and stratified by region, urbanicity, school type, ethnic mix, and size to achieve a nationally representative sample. The full study design and data collection procedures have been described in detail (Resnick et al., 1997). ADHD symptom measures used in the current analysis were drawn from an in-home survey collected during the third wave of data collection, conducted in 2001 and 2002. At that time, participants were between 17 and 28 years of age. For the current analyses 14,307 participants (52.8% female) were able to be included because they possessed all of the necessary data, including poststratification weights. One item from the fourth wave of data collection, self-reported ADHD diagnosis, was also included. Wave 4 was collected through in-home interviews between 2007 and 2009.

Measures

Participants were asked in Wave 3 to retrospectively report on their ADHD symptoms as they recall experiencing them between 5 and 12 years of age. Frequency of each symptom was assessed using a 4-point categorical scale with the following labels: “never or rarely,” “sometimes,” “often,” or “very often.” Table 1 displays the symptoms of ADHD assessed in Add Health and the rates of their endorsement in the sample. Our analyses included responses to 17 measured *DSM-IV-TR* ADHD symptoms (nine IN symptoms and eight HI symptoms). One of the *DSM-IV-TR* hyperactive-impulsive items (“Often interrupts or intrudes on others”) was not administered in the Add Health survey. In addition, in Wave 4, participants were asked whether “a doctor, nurse, or other health care provider ever told you that you have or had: attention problems, or ADD or ADHD?” This item was used as the self-reported ADHD diagnosis measure.

Statistical Analyses

Exploratory factor analyses for one, two, three, and four factors were estimated to examine how well models that allowed for quantitative symptom severity across items fit the data. LCA models with two through eight classes were estimated to examine the fit of models that allowed for qualitatively distinct classes among participants. Finally, FMM models containing two factors and one through five classes were estimated to examine how well models that allowed for both quantitative symptom severity differences and qualitatively distinct classes fit the data. The FMM models were specified to have two factors based on the nine inattentive items and eight hyperactive-impulsive items. This two-factor structure was chosen because it is in line with the design of the measures. Further, this factor structure characterization has been most widely supported in previous research (Bauermeister et al., 2010). All analyses were conducted separately for male and female subjects. Male and female subjects are diagnosed with ADHD at markedly different rates, with male subjects substantially more likely to receive a diagnosis (APA, 2000). Further, previous studies examining the structure of ADHD symptoms have analyzed male and female subjects

separately (Lubke et al., 2007); therefore, we wanted our analytic strategy to be consistent to allow for comparison. In all, 16 models were estimated for both male and female subjects.

Symptom items ranging from zero to 3 were treated as ordinal categorical measures, not continuous measures, within all models. Further, within the FMM (and LCA, where appropriate), item thresholds were allowed to vary across the classes, factor variances and covariances were allowed to vary across the classes and the factor means were fixed to zero. These model specifications were included so that the overall fit statistics did not reflect fit (or lack of fit) of these parameters but rather reflected differences in the number and kind of latent variables included in the model (B. Muthén, 2008).

The 16 models for both male and female subjects were compared through the examination of goodness-of-fit indices and substantive criteria. These fit indices included the Akaike Information Criterion (AIC; Akaike, 1973), the Bayesian Information Criterion (BIC; Schwarz, 1978), and the sample size adjusted Bayesian Information Criterion (adjusted BIC; Sclove, 1987). Lower values on all of these indices indicate a better fitting model. Models with more parameters are better able to fit the data, thus indices such as the AIC and BIC that include a penalty for more parameters and protect against unnecessary model complexity were used. Within each estimation method (FA, LCA, and FMM), additional indicators of fit that are not available across all types of models were also considered. All models were run within a mixture model framework specifying latent continuous variables (factors), latent categorical variables (classes), or both, to allow for comparison of fit across all models. Simulation studies have shown that comparing model fit of FA, LCA, and FMM leads to correct model choice in a wide variety of settings (Lubke & Neale, 2006, 2008).

When comparing models across subtypes, we considered all fit statistics presented in Tables 2 and 3. Within LCA and FMM models, the Lo-Mendell-Rubin (L-M-R) statistic (Lo, Mendell, & Rubin, 2001) provided a test of whether a given model fit the data better than a nested model with one fewer class. The most parsimonious model that best represents the data is preferred. A significant p value indicates that the model with fewer classes should be rejected in favor of the model with the greater number of classes. The L-M-R test was used in conjunction with overall fit statistics as well as an examination of the classes produced to ensure that the model determined to best fit the data was interpretable. The BIC has been shown to perform better or equally well when compared with the AIC, the adjusted BIC, and the L-M-R test (Nylund, Asparouhov, & Muthén, 2007). Therefore, when there was a discrepancy between the L-M-R tests and the other fit statistics, the BIC was given greater consideration. After deciding on a best fitting model, that model was rerun with the inclusion of the ADHD diagnosis variable as a test of the validity of the produced classes.

All analyses were conducted with *Mplus* 5.21 software (L. Muthén & Muthén, 1998–2009). Full information maximum likelihood was employed to include all available data. Less than 1% of participants had missing data on any one symptom. Survey design effects of individuals clustered within sampling unit (school), and stratification of geographic region were accounted for in the models. In addition, poststratification weights for the longitudinal sample were included in analyses to generalize the results to young adults in the United States and account for nonresponse bias.

RESULTS

Fit statistics and the proportion of the sample assigned to each class as produced by the 16 estimated models are presented in Tables 2 and 3 for male and female subjects, respectively.

Exploratory Factor Analysis

Models with one, two, three, and four factors were specified. As shown by the fit statistics reported in Tables 2 (for male subjects) and 3 (for female subjects), with the addition of factors, fit improved. The largest improvement in fit (as shown by the relative change in AIC, BIC, and adjusted BIC values), however, came with the addition of a second factor over the first initial factor. In contrast, the improvement in fit between the two-factor and three-factor solution was relatively small. For both male and female subjects, the first two eigenvalues exceeded 1, whereas the third eigenvalue for both solutions did not exceed 1. Therefore, both the fit statistics and eigenvalues supported a two-factor solution (Tabachnick & Fidell, 2007). Further, an examination of the two-factor solution showed that for male and female subjects, the estimated correlation between the two latent factors was .58 and .53, respectively. This suggests that, although correlated, they are distinct and further suggests a two-factor solution is a better characterization of the data than a one-factor solution. An examination of the three-factor solution revealed factors that did not have clear interpretation; for both male and female subjects, items exhibited relatively equal cross loadings and seven of the 17 items did not exhibit a loading that exceeded .50 on any of the three factors.

When factor loadings in the two-factor solution were considered, items generally loaded well on one of the two factors; however, several items exhibited relatively equal cross-loadings. All of the nine items that were hypothesized to measure inattentive symptoms loaded on the first factor. Four of the items that were hypothesized to measure hyperactive-impulsive dimensions of ADHD loaded on the second factor (on the go, talked, blurted, and difficulty awaiting your turn). The other four items that were hypothesized to measure hyperactive-impulsive dimensions of ADHD (fidgeted, left your seat, restless, and difficulty doing fun things quietly) loaded on both factors. Although we conclude that of the FA models, the two-factor solution was best supported, the items loadings did not perfectly represent two factors. This lack of a strongly supported FA model may be due to heterogeneous subgroups in the population (Hancock & Samuelsen, 2008; B. Muthén, 2008). For example, there are a large number of respondents in the data set who do not experience any symptoms to a high degree. When this subgroup, or qualitatively distinct class of people, is not accounted for in a model, such as in a FA model, the solution may not fit the data well. For this reason, LCA and FMM models that allow for subgroups with differing factor variances and covariances may better characterize the data.

Latent Class Analysis

Next, LCA models ranging from one to eight classes were fit to the data. For female subjects, the L-M-R test was significant for the two-, three-, and four-class models and nonsignificant for the five-class model. For male subjects, the L-M-R test was significant for the two-, three-, four-, five-, and six-class models and nonsignificant for the seven-class model. Therefore, the L-M-R tests support a four-class model for female subjects and a six-class model for males. For male and female subjects, the AIC, BIC, and adjusted BIC improved as the number of classes increased. This trend continued through the eight-class solution, which was the highest number of classes tested. Although the three statistical information criteria decreased in size as more classes were included, the criteria generally dropped at a diminished rate as more classes were added. For both male and female subjects, there was a large drop in these indices when a three-class solution was compared to a two-class solution. As the number of classes increased, the improvement in fit became substantially smaller.

Across all the models for both male and female subjects, the addition of classes seemed to tap quantitative differences in severity of symptoms rather than qualitatively different

symptoms of ADHD based on an examination of plots of the estimated probabilities of symptom scores. In nearly all the classes produced, people were more likely to report experiencing hyperactive-impulsive symptoms than they were to report experiencing inattentive symptoms. The four-class solution for female subjects showed a highly affected class in which 8.8% of participants were categorized. The other three classes only appeared to differ based on two symptoms (fidgeted and talked too much). The six-class solution for male subjects showed a highly affected class in which 4.8% of participants were categorized. Three classes showed very low probabilities of any symptoms. Two classes differed from the less affected classes by showing a higher probability of experiencing three symptoms (easily distracted, fidgeted, and talked too much).

Simulation studies have shown that fit statistics tend to overestimate the number of classes in LCA models; consequently, researchers must typically rely on a combination of statistical information criteria and judgment in determining whether a solution is in agreement with substantive theory (Nylund et al., 2007). As the number of classes increased, the highly affected class became a smaller minority for both male and female subjects. Further, classes other than the highly affected class within each solution became difficult to distinguish from one another and called into question the utility of such complex solutions. Fit statistics supported a solution with at least four classes. As the number of classes increased, resulting classes differed on only a small number of symptoms. Substantive theory about ADHD would not suggest that an experience of one or two symptoms to a high degree should be categorized as a type of ADHD. Therefore, a model that allows for symptom severity differences within classes, such as FMM, may better characterize the data than either FA or LCA.

Factor Mixture Models

FMMs that allowed for two factors and varied the number of classes specified were fit to the data. In FMM, the factor structure is specified and the latent classes are estimated given that particular latent factor structure. In this way, models allow for both quantitative symptom severity differences and qualitatively distinct classes. As shown in Tables 2 and 3, the three statistical information criteria generally indicated better fit as the models became more complex. This was true through the two-factor five-class model, the most complex model tested, and was the case for both male and female subjects. We did not include models more complex than the two-factor five-class model because results of such models were not interpretable. As with the LCA analyses, we also examined the L-M-R tests for each model and considered substantive theory and an examination of estimated symptom probability plots to determine the best fitting FMM. Figure 1 shows plots for the two-factor two-class and the two-factor three-class models for male and female subjects. Plotted scores represent the probability of a person within each class experiencing each symptom “very often” (score of 3).

For male subjects, the L-M-R tests were not significant for any of the tested models, which did not support any model more complex than the two-factor one-class solution. The AIC, BIC, and adjusted BIC, however, showed a large decrease between the two-factor one-class and the two-factor two-class solutions. These three fit indices continued to show relatively smaller decreases with additional classes. The two-factor two-class solution produced a more affected class in which 27.8% of participants were categorized. Participants in this class were more likely to have experienced each of the symptoms at the highest level than were participants in the less affected class (see Figure 1). In the two-factor three-class solution, only 15.5% of participants were categorized into the most affected class. The two less affected classes were distinguished by an elevated probability of people in the class with 49.1% experiencing four symptoms (easily distracted, fidgeted, felt on the go, and talked too much).

For female subjects, the two-factor two-class model and the two-factor three-class model exhibited significant L-M-R tests indicating that each fit the data better than a model with one fewer class. The L-M-R test for the two-factor four-class model was not significant. As with the models for male subjects, the AIC, BIC, and adjusted BIC showed a large decrease between the two-factor one-class and the two-factor two-class solutions. Additional classes resulted in relatively smaller decreases in these three fit indices. The two-factor two-class solution produced a highly affected class in which 35.1% of participants were categorized. The two-factor three-class solution produced a highly affected class in which 17.6% of participants were categorized. Within this class, the estimated probability of a person experiencing each symptom at the most severe level (score of 3) was only 10 to 20% for most of the symptoms. The probability of experiencing three symptoms (easily distracted, fidgeted, and talked too much) was higher. The main difference between the two less affected classes was the likelihood of experiencing one symptom (talked too much). People in the class with 43.6% of participants had a 20% chance of experiencing one symptom (talked too much) at a severe level, whereas people in the class with 38.8% of participants had a 5% chance. For female subjects, the two-factor two-class model would be preferred given that the additional class gained in the two-factor three-class model only reflected a class that experienced the “talked too much” symptom. Considering this symptom by itself as a type of ADHD is not consistent with substantive theory.

Fit Statistic Comparisons Across Three Model Subtypes

An overall comparison of the models indicated a preference for the FMM models over both the FA and LCA models. This was true for models with male and female subjects. The statistical information criteria in Tables 2 and 3 indicate that the FMM models produced the overall best fit (lowest fit statistics). The two-factor two-class FMM model was most supported by fit statistics and substantive theory given that results of the two-factor three-class solutions showed quantitatively intermediate classes rather than a qualitatively distinct subtypes. For female subjects, 35.1% of the sample was classified into the affected class, whereas for male subjects, 27.8% of the sample was classified into the affected class. The probability of reporting a score in the highest response category for each of the 17 ADHD symptom items is shown in Figure 1 for each class. The highly affected class was more likely to experience both the inattentive symptoms and the hyperactive-impulsive symptoms, which indicates that the classes differed in terms of symptom severity rather than symptom type. The fact that the FA models provided a better fit than the LCA models also emphasizes the importance of severity differences rather than subtypes in the characterization of ADHD.

The correlation between the two factors was allowed to differ between the classes, allowing an examination of the extent to which IN and HI symptoms correlated within each class. For female and male subjects, the relation between the IN factor and HI factor was strong in both the affected and unaffected class (male affected class: correlation = .80, $SE = .02$; male unaffected class: correlation = .87, $SE = .02$; female affected class: correlation = .84, $SE = .03$; female unaffected class: correlation = .90, $SE = .02$). The correlations between the factors were higher in the FMM models as compared to the FA models because the correlations in the FMM models were based on homogenous persons within each class, whereas the correlations in the FA model were estimated with heterogeneous samples and thus were attenuated. These strong correlations between IN and HI symptoms support the conclusion that ADHD symptoms differ primarily by symptom severity rather than symptom type.

Relation of Class Membership to Reported History of ADHD Diagnosis

The two-factor two-class models for both male and female subjects were rerun with the addition of the ADHD diagnosis variable. The models tested the relation between the class

variable and ADHD diagnosis. Specifically, whether or not the person reported having been diagnosed with ADHD was regressed on the categorical latent class variable. Individuals who were categorized in the affected class were more likely to report a diagnosis of ADHD from a health care provider (*OR* male subjects = 4.03, 95% CI [2.65, 6.13]; *OR* female subjects = 5.65, 95% CI [3.15, 10.10]).

DISCUSSION

Better understanding of the phenotypic structure of ADHD symptoms is important for advancing etiologic and predictive research as well as improving clinical practice. Several authors have addressed this issue by using a variety of methodological approaches. These approaches have inherent assumptions, and few studies have compared and contrasted results acquired from the various methods. A major aim of this research was to test and compare different factor and class structures of ADHD symptoms in a U.S. population-based sample. The data from the current study support the use of factor mixture modeling to characterize ADHD symptoms. The best characterization of ADHD for both male and female subjects was a model with two factors (IN and HI) and two classes (affected group and unaffected group). These findings are in line with the *DSM-IV-TR* characterization of ADHD differentiating two categories of inattention and hyperactive-impulsive symptoms. Although two classes emerged, an affected and unaffected group, the results support a dimensional quality to symptom severity within these classes as opposed to viewing the disorder as purely categorical. This study is unique in that the current analyses use data from a large, nationally representative U.S. population-based sample and compare the performance of a relatively new analytic method, FMM, with the more commonly used LCA and FA approaches. Previous similar studies have used smaller European samples (Lubke et al., 2009; Lubke et al., 2007). Because the data contain self-reports of past ADHD diagnosis we were able to estimate the probability of likely ADHD given class membership, which provides some support for the validity of the solution. It should be noted that the proportions of participants assigned to the affected classes within FMM models exceeded typical ADHD prevalence estimates for both male and female subjects. This may be due to a number of methodological factors including the self-report and retrospective nature of the data. Further, models were based on symptom data only, not clinician assessment; therefore, assignment to the affected class in this analysis should not be considered equivalent to experiencing the ADHD disorder. A main goal of the current study was to compare empirically driven modeling approaches of ADHD symptoms. Conventional modeling approaches such as FA and LCA have inherent limitations. FA does not allow for model-based classification and as such is not well aligned with a preference among clinical providers to identify thresholds for the purpose of determining diagnostic status. LCA can be useful for classification purposes, but because it ignores within-class heterogeneity in symptom severity, potentially resulting in numerous classes with diminishing samples sizes within classes, it is not well aligned with a preference among researchers who are concerned with Type II error.

FMM provides a marriage of the FA and LCA approaches by providing model-based classification while retaining within-class heterogeneity in symptom severity (B. Muthén, 2006). By comparing different approaches it is possible to determine which phenotypic structure of a disorder best represents the data under analysis. In the case of ADHD, our conclusions largely validate those of Lubke and colleagues (Lubke et al., 2009; Lubke et al., 2007) who examined ADHD in two population-based (nonclinical) European samples. That is, within this U.S. population-based sample there is not strong evidence to identify multiple subgroups of qualitatively different symptoms. Rather, dimensional factors of inattentive and hyperactive-impulsive symptoms arranged along two quantitatively (high vs. low or affected vs. unaffected) different classes appear to best represent these data.

Despite the similarities between our findings and those of Lubke and colleagues (Lubke et al., 2009; Lubke et al., 2007) there was one notable difference. Whereas Lubke et al. (2007) found in their sample that people were more likely to experience IN symptoms than HI symptoms, our results showed the opposite. There are several possible reasons for this finding. First, it is possible that symptom experiences differ between U.S. and Northern European samples or that the interpretation of the symptom questions differed. Alternatively, these findings may be the result of different methodologies. Participants in the current analyses were asked to recall symptoms that occurred at a much earlier developmental age (age 5–12) as compared to the Finnish sample (age 16–18; Lubke et al. 2007). Findings may represent additional evidence that hyperactive-impulsive types of behavior decrease with age (Lahey & Willcutt, 2010). Further, the variation in measurement methods (retrospective self-report vs. caregiver report) may have contributed to this difference, as it may be easier to recall hyperactive-impulsive behaviors due to their potentially disruptive elements which may elicit disciplinary actions from caregivers or teachers. These are only some possible reasons for the higher reported incidence of HI symptoms and further research exploring this finding is needed before any firm conclusions can be made.

The IN and HI factors were strongly correlated in both classes for male and female subjects. It is possible that some people experience one type of ADHD symptom (i.e., inattentive) to a greater degree than the other, but in general, strong correlations between the factors suggest that in both the affected and unaffected classes people were experiencing both types of symptoms to a similar degree. The *DSM-IV-TR* assumes item equivalence of all ADHD symptom measures. Recently employed item response theory models provide evidence that this is an incorrect assumption; rather items may differ in their predictive power and may differ by informant (Gomez, 2007; Gomez, Vance, & Gomez, 2011). Items loadings across the two factors, within the current models, were allowed to differ, consistent with this recent suggestion that items should not be considered equivalent.

Subtypes of ADHD may be useful in clinical practice, and it is possible that subtypes would be supported by FMM models that utilize data from a clinical sample. Within this population-based sample, however, the best characterization of ADHD symptom experiences is two classes that differ primarily on symptom severity. Understanding ADHD symptoms prevalence and co-occurrence in a nonclinical sample is important given that ADHD symptoms, even at subclinical thresholds may represent risk factors for health behaviors including smoking and obesity (Kollins, McClernon, & Fuemmeler, 2005; Fuemmeler, Ostbye, Yang, McClernon, & Kollins, 2011; Fuemmeler, Kollins, & McClernon, 2007).

Limitations

A limitation of the current study was that the Add Health data only included retrospective self-report of ADHD symptoms. In addition, the ADHD diagnosis measure was self-reported, not obtained from a clinician using standardized diagnostic ratings, which would have provided a more accurate measure. On one hand, some have argued that retrospective report of ADHD symptoms and diagnosis lacks validity (Mannuzza, Klein, Klein, Bessler, & ShROUT, 2002). Mannuzza et al. (2002) found that 85% of persons with childhood ADHD were correctly identified using adult retrospective data and that 11% of those without childhood ADHD were misdiagnosed using adult retrospective data. After adjusting for ADHD population prevalence, they estimated that of adults retrospectively given a diagnosis of childhood ADHD, only 27% would be correctly identified. On the other hand, the use of retrospective report is common in clinical practice when working with adults with ADHD, and extant data do support the reliability and validity of this method (Epstein & Kollins, 2005; Murphy & Schachar, 2000; Stein et al., 1995; Ward, Wender, & Reimherr, 1993;

Zucker, Morris, Ingram, Morris, & Bakeman, 2002). Moreover, our prior research with the Add Health database found that individuals reporting six or more symptoms at Wave 3 were significantly more likely to have parents indicate the presence of learning or other behavioral problems at Wave 1 (Kollins et al., 2005), further supporting the validity of retrospective ADHD symptom data in general and of these data in particular.

Although the retrospective nature of the data limits some of the conclusions that can be drawn, the significant relation between ADHD diagnosis and the FMM results offers validation of the findings, especially given that the symptom ratings and ADHD diagnosis information were collected several years apart. The fact that one of the *DSM-IV-TR* symptoms was omitted from the data collection is another limitation. Having only self-report data did not allow for triangulation with caregiver and teacher reports, which would be another useful extension of this work. The use of multiple reporters, however, presents certain challenges. For instance, symptom patterns may be overrepresented or underrepresented in certain contexts for which the reporter is likely to be interacting with the child. Self-reports, on the other hand, should provide a general assessment of experiences across situations. Nevertheless, future studies using these methods with multiple reporters would add to extant findings. Finally, although much of the data in this cohort were prospectively obtained, questions pertaining to ADHD symptoms were not. As such, we are not able to determine the degree to which these phenotypic structures are stable over a range of developmental ages. Analysis of prospective longitudinal ADHD data examining the stability of various ADHD factor structures over time is needed.

Implications for Research, Policy and Practice

This research strengthens our understanding of the underlying phenotypic structure of ADHD symptoms. The current analyses, conducted using self-reported retrospective ratings of ADHD symptoms within a U.S. population-based sample, suggest that a small affected class of persons exists as well as a large unaffected class. Within both classes, the two factors were strongly correlated, supporting the conclusion that differences in ADHD symptoms have more to do with symptom severity than symptom type. Although some people may experience one type of symptom to a greater degree than the other, in this population sample, those people were not a large-enough group to produce classes that represent different types of ADHD as described in the *DSM-IV-TR*. It is yet to be determined whether the best fitting structure of ADHD phenotypes would be different if tested within a clinical sample. Results with a clinic-referred sample could show several affected classes or mirror the current results and show only one affected class. Further, a clinic sample may experience more severe ADHD symptoms and also additional psychopathology (Goodman et al., 1997). Estimating these models within a clinic sample is an important next step in understanding how ADHD symptoms co-occur. Future research would benefit from comparing ADHD phenotypes within clinical samples with general population samples or unaffected control groups.

The current research highlights important differences between different analytic strategies commonly used to characterize the phenotypic structure of ADHD. Our results indicate that for both male and female subjects, a two-factor two-class solution best represented ADHD symptoms. From a clinical perspective, it will be important for future work to assess the extent to which FMM latent class membership is predictive of relevant adverse outcomes such as substance use/smoking, occupational underachievement, interpersonal difficulties, and other life impairments. In addition, the relative contribution of genetic variation to membership in various FMM latent classes will be important to evaluate in the context of a more thorough understanding of the etiology of ADHD. A FMM model of ADHD symptomatology should better represent ADHD phenotypes than summed index measures and would therefore allow a better test of the relationships among ADHD and potential

correlates. Further, in the development of the *DSM-V*, the best way to address heterogeneity of symptom presentation is being explored (Nigg, Tannock, & Rohde, 2010). Research that examines ADHD symptoms through empirically driven methods, such as the current work, may prove invaluable in this endeavor.

Acknowledgments

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (<http://www.cpc.unc.edu/addhealth>). No direct support was received from grant P01HD31921 for this analysis. Authors received support for analyses and manuscript preparation from the National Cancer Institute K07CA124905, and the National Institute on Drug Abuse R01DA030487, K24DA023464 and P30DA023026.

REFERENCES

- Akaike, H. Information theory and an extension of the maximum likelihood principle.. In: Petrov, BN.; Csaki, F., editors. Second international symposium on information theory. Springer; Verlag: 1973. p. 267-281.
- Amador-Campos JA, Fornis-Santacana M, Martorell-Balanzo B, Guardia-Olmos J, Pero-Cebollero M. Confirmatory factor analysis of parents' and teachers' ratings of *DSM-IV* symptoms of attention deficit hyperactivity disorder in a Spanish sample. *Psychological Reports*. 2005; 97:847–860. doi: 10.2466/pr0.97.3.847-860. [PubMed: 16512303]
- Amador-Campos JA, Fornis-Santacana M, Martorell-Balanzo B, Guardia-Olmos J, Pero-Cebollero M. *DSM-IV* Attention Deficit Hyperactivity Disorder symptoms: Agreement between informants in prevalence and factor structure at different ages. *Journal of Psychopathology and Behavioral Assessment*. 2006; 28:23–32. doi:10.1007/s10862-006-4538-x.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed., text rev.. Author; Washington, DC: 2000.
- Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin*. 1997; 121(1):65–94. doi: 10.1037//0033-2909.121.1.65. [PubMed: 9000892]
- Barkley RA, Cox D. A review of driving risks and impairments associated with attention-deficit/hyperactivity disorder and the effects of stimulant medication on driving performance. *Journal of Safety Research*. 2007; 38(1):113–128. doi:10.1016/j.jsr.2006.09.004. [PubMed: 17303170]
- Bauermeister JJ, Canino G, Polanczyk G, Rohde LA. ADHD across cultures: Is there evidence for a bidimensional organization of symptoms? *Journal of Clinical Child and Adolescent Psychology*. 2010; 39:362–372. doi:10.1080/15374411003691743. [PubMed: 20419577]
- Castellanos FX, Tannock R. Neuroscience of attention-deficit/hyperactivity disorder: the search for endopheno-types. *Nature Reviews Neuroscience*. 2002; 3:617–628. doi:10.1038/nrn896.
- Eiraldi RB, Power TJ, Nezu CM. Patterns of comorbidity associated with subtypes of attention-deficit/hyperactivity disorder among 6- to 12-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36:503–514. doi:10.1097/00004583-199704000-00013. [PubMed: 9100425]
- Epstein JN, Kollins SH. Psychometric properties of an adult ADHD diagnostic interview. *Journal of Attention Disorders*. 2005; 9:504–514. doi:10.1177/1087054705283575. [PubMed: 16481667]
- Fergusson DM, Horwood LJ, Lynskey MT. Structure of DSM-III-R criteria for disruptive childhood behaviors: confirmatory factor models. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1994; 33:1145–1157. doi:10.1097/00004583-199410000-00010. [PubMed: 7982865]
- Froehlich TE, Lanphear BP, Epstein JN, Barbaresi WJ, Katusic SK, Kahn RS. Prevalence, recognition, and treatment of attention-deficit/hyperactivity disorder in a national sample of US children. *Archives of Pediatrics & Adolescent Medicine*. 2007; 161:857–864. doi:10.1001/archpedi.161.9.857. [PubMed: 17768285]

- Fuemmeler BF, Kollins SH, McClernon FJ. Attention deficit hyperactivity disorder symptoms predict nicotine dependence and progression to regular smoking from adolescence to young adulthood. *Journal of Pediatric Psychology*. 2007; 32:1203–1213. [PubMed: 17602186]
- Fuemmeler BF, Ostbye T, Yang C, McClernon FJ, Kollins SH. Association between attention-deficit/hyperactivity disorder symptoms and obesity and hypertension in early adulthood: A population-based study. *International Journal of Obesity*. 2011; 35:852–862. doi:10.1038/ijo.2010.214. [PubMed: 20975727]
- Gomez R. Australian parent and teacher ratings of the DSM–IV ADHD Symptoms: Differential symptom functioning, and parent–teacher agreement and differences. *Journal of Attention Disorders*. 2007; 11:17–27. doi:10.1177/1087054706295665. [PubMed: 17606769]
- Gomez R, Burns GL, Walsh JA, Hafetz N. A multitrait–multisource confirmatory factor analytic approach to the construct validity of ADHD and ODD rating scales with Malaysian children. *Journal of Abnormal Child Psychology*. 2005; 33:241–254. doi:10.1007/s10802-005-1831-1. [PubMed: 15839501]
- Gomez R, Vance A, Gomez A. Item response theory analyses of parent and teacher ratings of the ADHD symptoms for recoded dichotomous scores. *Journal of Attention Disorders*. 2011; 15:269–285. doi:10.1177/1087054709356404. [PubMed: 20019380]
- Goodman S, Lahey B, Fielding B, Dulcan M, Narrow W, Regier D. Representativeness of clinical samples of youths with mental disorders: A preliminary population-based study. *Journal of Abnormal Psychology*. 1997; 106:3–14. doi:10.1037//0021-843X.106.1.3. [PubMed: 9103713]
- Hancock, GR.; Samuelsen, KM. *Advances in latent variable mixture models*. Information Age; Charlotte, NC: 2008. p. 12
- Hardy KK, Kollins SH, Murray DW, Riddle MA, Greenhill L, Cunningham C, Chuang SZ. Factor structure of parent- and teacher-rated attention-deficit/hyperactivity disorder symptoms in the Preschoolers with Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescent Psychopharmacology*. 2007; 17:621–634. doi:10.1089/cap.2007.0073. [PubMed: 17979582]
- Hudziak JJ, Heath AC, Madden PF, Reich W, Bucholz KK, Slutske W, Todd RD. Latent class and factor analysis of *DSM–IV* ADHD: A twin study of female adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1998; 37:848–857. doi: 10.1097/00004583-199808000-00015. [PubMed: 9695447]
- Kessler RC, Adler L, Barkley R, Biederman J, Conners CK, Demler O, Zaslavsky AM. The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *The American Journal of Psychiatry*. 2006; 163:716–723. doi:10.1176/appi.ajp.163.4.716. [PubMed: 16585449]
- Kollins SH, McClernon FJ, Fuemmeler BF. Association between smoking and Attention-Deficit/Hyperactivity Disorder symptoms in a population-based sample of young adults. *Archives of General Psychiatry*. 2005; 62:1142–1147. doi:10.1001/archpsyc.62.10.1142. [PubMed: 16203959]
- Lahey BB, Willcutt EG. Predictive validity of a continuous alternative to nominal subtypes of attention-deficit/hyperactivity disorder for *DSM–V*. *Journal of Clinical Child & Adolescent Psychology*. 2010; 39:761–775. doi:10.1080/15374416.2010.517173. [PubMed: 21058124]
- Lo Y, Mendell N, Rubin DB. Testing the number of components in a normal mixture. *Biometrika*. 2001; 88:767–778. doi:10.1093/biomet/88.3.767.
- Lubke GH, Hudziak JJ, Derks EM, van Bijsterveldt TC, Boomsma DI. Maternal ratings of attention problems in ADHD: Evidence for the existence of a continuum. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2009; 48:1085–1093. doi:10.1097/CHI.0b013e3181ba3dbb. [PubMed: 19797980]
- Lubke GH, Muthén B, Moilanen IK, McGough JJ, Loo SK, Swanson JM, Smalley SL. Subtypes versus severity differences in attention-deficit/hyperactivity disorder in the Northern Finnish Birth Cohort. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2007; 46:1584–1593. doi:10.1097/chi.0b013e31815750dd. [PubMed: 18030080]
- Lubke GH, Neale MC. Distinguishing between latent classes and continuous factors: resolution by maximum likelihood? *Multivariate Behavioral Research*. 2006; 41:499–532. doi: 10.1080/00273170802490673.

- Lubke GH, Neale M. Distinguishing between latent classes and continuous factors with categorical outcomes: Class invariance of parameters of factor mixture models. *Multivariate Behavioral Research*. 2008; 43:592–620. doi:10.1080/00273170802490673. [PubMed: 20165736]
- Mannuzza S, Klein RG, Bessler A, Malloy P, Hynes ME. Educational and occupational outcome of hyperactive boys grown up. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36:1222–1227. doi:10.1097/00004583-199709000-00014. [PubMed: 9291723]
- Mannuzza S, Klein RG, Bessler A, Malloy P, LaPadula M. Adult outcome of hyperactive boys. Educational achievement, occupational rank, and psychiatric status. *Archives of General Psychiatry*. 1993; 50:565–576. [PubMed: 8317950]
- Mannuzza S, Klein RG, Bessler A, Malloy P, LaPadula M. Adult psychiatric status of hyperactive boys grown up. *The American Journal of Psychiatry*. 1998; 155:493–498. [PubMed: 9545994]
- Mannuzza S, Klein RG, Klein DF, Bessler A, ShROUT P. Accuracy of adult recall of childhood attention deficit hyperactivity disorder. *The American Journal of Psychiatry*. 2002; 159:1882–1888. doi:10.1176/appi.ajp.159.11.1882. [PubMed: 12411223]
- Mannuzza S, Klein RG, Moulton JL III. Lifetime criminality among boys with attention deficit hyperactivity disorder: a prospective follow-up study into adulthood using official arrest records. *Psychiatry Research*. 2008; 160:237–246. doi:10.1016/j.psychres.2007.11.003. [PubMed: 18707766]
- Martel MM, von Eye A, Nigg JT. Revisiting the latent structure of ADHD: Is there a ‘g’ factor? *The Journal of Child Psychology and Psychiatry*. 2010; 51:905–914. doi:10.1111/j.1469-7610.2010.02232.x.
- Murphy P, Schachar R. Use of self-ratings in the assessment of symptoms of attention deficit hyperactivity disorder in adults. *American Journal of Psychiatry*. 2000; 157:1156–1159. doi:10.1176/appi.ajp.157.7.1156. [PubMed: 10873926]
- Muthén B. Should substance use disorders be considered as categorical or dimensional? *Addiction*. 2006; 101:6–16. doi:10.1111/j.1360-0443.2006.01583.x.
- Muthén, B. Latent variable hybrids: Overview of old and new models.. In: Hancock, GR.; Samuelsen, KM., editors. *Latent variable mixture models*. Information Age; Charlotte, NC: 2008. p. 1-24.
- Muthén, L.; Muthén, B. *Mplus 5.21 User's Guide*. Muthén & Muthén; Los Angeles, CA: 1998–2009.
- Neuman RJ, Sitdhiraksa N, Reich W, Ji TH, Joyner CA, Sun LW, Todd RD. Estimation of prevalence of DSM-IV and latent class-defined ADHD subtypes in a population-based sample of child and adolescent twins. *Twin Research and Human Genetics*. 2005; 8(4):392–401. doi:10.1375/1832427054936646. [PubMed: 16176725]
- Nigg JT, Tannock R, Rohde LA. What is to be the fate of ADHD subtypes? An introduction to the special section on research on the ADHD subtypes and implications for the *DSM-V*. *Journal of Clinical Child & Adolescent Psychology*. 2010; 39:723–725. doi:10.1080/15374416.2010.517171. [PubMed: 21058120]
- Nylund KL, Asparouhov T, Muthén B. Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. *Structural Equation Modeling*. 2007; 14:535–569.
- Pelham WE, Foster EM, Robb JA. The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. *Journal of Pediatric Psychology*. 2007; 32:711–727. doi:10.1093/jpepsy/jsm022. [PubMed: 17556402]
- Rasmussen ER, Neuman RJ, Heath AC, Levy F, Hay DA, Todd RD. Replication of the latent class structure of Attention-Deficit=Hyperactivity Disorder (ADHD) subtypes in a sample of Australian twins. *The Journal of Child Psychology and Psychiatry*. 2002; 43:1018–1028. doi:10.1111/1469-7610.00229.
- Rasmussen ER, Neuman RJ, Heath AC, Levy F, Hay DA, Todd RD. Familial clustering of latent class and *DSM-IV* defined attention-deficit/hyperactivity disorder (ADHD) subtypes. *The Journal of Child Psychology and Psychiatry*. 2004; 45:589–598. doi:10.1111/j.1469-7610.2004.00248.x.
- Resnick MD, Bearman PS, Blum RW, Bauman KE, Harris KM, Jones J, Udry JR. Protecting adolescents from harm: findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association*. 1997; 278:823–832. doi:10.1001/jama.278.10.823. [PubMed: 9293990]

- Schwarz G. Estimating the dimension of a model. *Annals of Statistics*. 1978; 6:461–464. doi:10.1214/aos/1176344136.
- Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*. 1987; 52:333–343. doi:10.1007/BF02294360.
- Singh I. A disorder of anger and aggression: Children's perspectives on attention deficit/hyperactivity disorder in the UK. *Social Science & Medicine*. 2011 Advance online publication. doi:10.1016/j.socscimed.2011.03.049.
- Solanto MV, Gilbert SN, Raj A, Zhu J, Pope-Boyd S, Stepak B, Newcorn JH. Neurocognitive functioning in AD/HD, predominantly inattentive and combined subtypes. *Journal of Abnormal Child Psychology*. 2007; 35:729–744. doi:10.1007/s10802-007-9123-6. [PubMed: 17629724]
- Sonuga-Barke E, Bitsakou P, Thompson M. Beyond the dual pathway model: evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2010; 49:345–355. doi:00004583-201004000-00009 [pii]. [PubMed: 20410727]
- Stein MA, Sandoval R, Szumowski EK, Roizen N, Reinecke MA, Blondis T. Psychometric characteristics of the Wender Utah Rating Scale (WURS): Reliability and factor structure for men and women. *Psychopharmacology Bulletin*. 1995; 31:425–433. [PubMed: 7491401]
- Tabachnick, BG.; Fidell, LS. *Using multivariate statistics*. 5th ed.. Allyn & Bacon; New York, NY: 2007.
- Todd RD, Rasmussen ER, Neuman RJ, Reich W, Hudziak JJ, Bucholz KK, Heath A. Familiality and heritability of subtypes of attention deficit hyperactivity disorder in a population sample of adolescent female twins. *The American Journal of Psychiatry*. 2001; 158:1891–1898. doi:10.1176/appi.ajp.158.11.1891. [PubMed: 11691697]
- Volk HE, Neuman RJ, Todd RD. A systematic evaluation of ADHD and comorbid psychopathology in a population-based twin sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2005; 44:768–775. doi:10.1097/01.chi.0000166173.72815.83. [PubMed: 16034278]
- Ward MF, Wender PH, Reimherr FW. The Wender Utah Rating Scale: An aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *American Journal of Psychiatry*. 1993; 150:885–890. [PubMed: 8494063]
- Willcutt EG, Pennington BF, Chhabildas NA, Friedman MC, Alexander J. Psychiatric comorbidity associated with *DSM-IV* ADHD in a nonreferred sample of twins. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1999; 38:1355–1362. doi:10.1097/00004583-199911000-00009. [PubMed: 10560221]
- Wolraich ML, Lambert EW, Baumgaertel A, Garcia-Tornel S. Teachers' screening for attention deficit/hyperactivity disorder: Comparing multinational samples on teacher ratings of ADHD. *Journal of Abnormal Child Psychology*. 2003; 31:445–555. [PubMed: 12831232]
- Zucker M, Morris MK, Ingram SM, Morris RD, Bakeman R. Concordance of self- and informant ratings of adults' current and childhood attention-deficit/hyperactivity disorder symptoms. *Psychology Assessment*. 2002; 14:379–389. doi:10.1037//1040-3590.14.4.379.

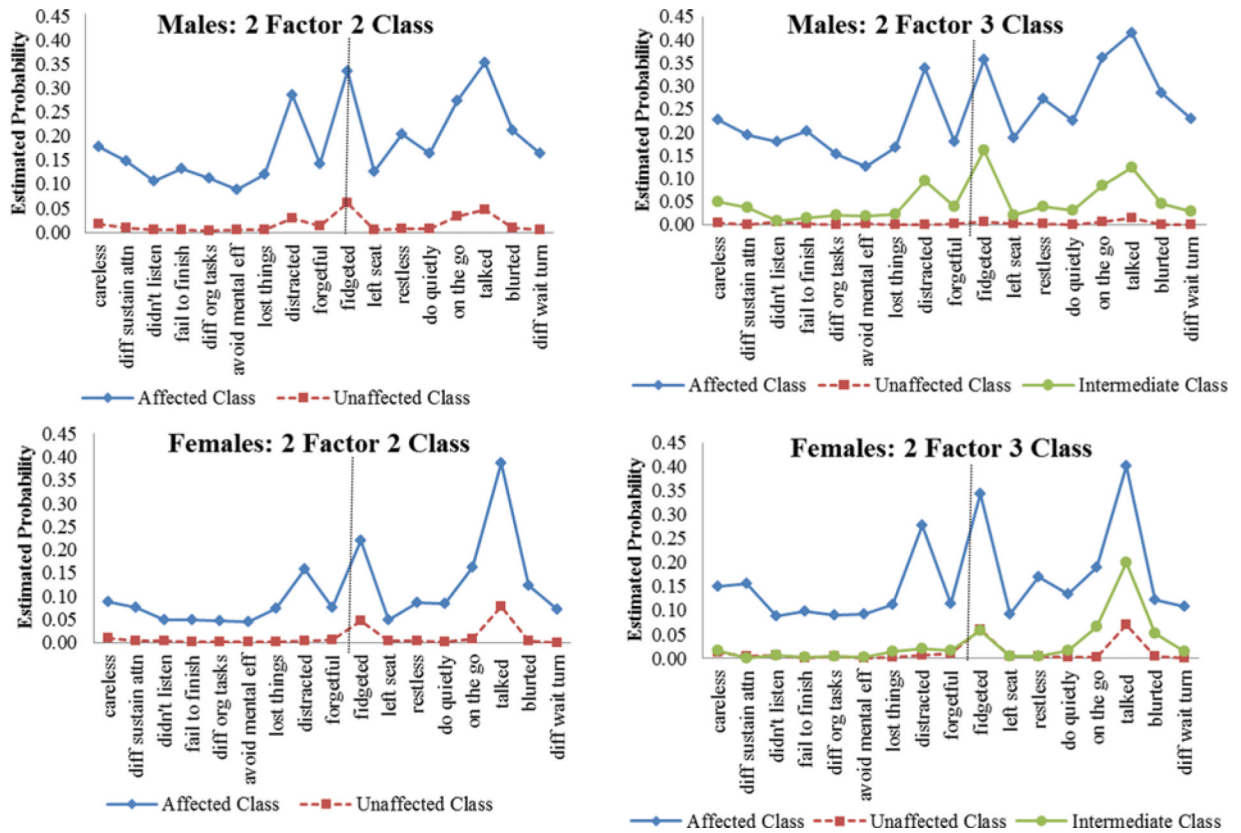


FIGURE 1. Estimated probabilities of experiencing each symptom “very often” within each class for male and female subjects. *Note:* Items to the left of the dotted lines are inattentive items; items to the right are hyperactive-impulsive items. (Figure appears in color online.)

TABLE 1

ADHD Items and Prevalence Among Male and Female Subjects

Question #	Question	Response Set	Males	Females
<i>Inattentive Items</i>				
1	You failed to pay close attention to details or made careless mistakes in your work.	Never or Rarely Sometimes Often Very Often	.203 .543 .193 .061	.295 .539 .127 .039
3	You had difficulty sustaining your attention in tasks or fun activities.	Never or Rarely Sometimes Often Very Often	.477 .360 .114 .050	.583 .319 .069 .030
5	You didn't listen when spoken to directly.	Never or Rarely Sometimes Often Very Often	.498 .379 .088 .035	.619 .311 .049 .021
7	You didn't follow through on instructions and failed to finish work .	Never or Rarely Sometimes Often Very Often	.426 .444 .089 .041	.604 .326 .051 .019
9	You had difficulty organizing tasks and civities.	Never or Rarely Sometimes Often Very Often	.455 .407 .104 .034	.567 .354 .060 .019
11	You avoided , disliked, or were reluctant to engage in work requiring sustained mental effort .	Never or Rarely Sometimes Often Very Often	.522 .359 .090 .030	.658 .275 .050 .018
13	You lost things that were necessary for tasks or activities.	Never or Rarely Sometimes Often Very Often	.412 .450 .101 .037	.505 .395 .072 .028
15	You were easily distracted .	Never or Rarely Sometimes Often Very Often	.254 .437 .209 .1	.370 .443 .128 .059
17	You were forgetful .	Never or Rarely Sometimes Often Very Often	.356 .473 .123 .048	.448 .434 .086 .032
<i>Hyperactive-Impulsive Items</i>				
2	You fidged with your hands or feet or squirmed in your seat.	Never or Rarely Sometimes Often Very Often	.226 .404 .232 .138	.287 .398 .207 .109
4	You left your seat in the classroom or in other situations when being seated was expected.	Never or Rarely Sometimes Often Very Often	.604 .275 .080 .041	.766 .170 .044 .020
6	You felt restless .	Never or Rarely Sometimes Often Very Often	.348 .448 .141 .063	.462 .420 .085 .032
8	You had difficulty doing fun things quietly .	Never or Rarely Sometimes Often Very Often	.464 .359 .126 .051	.605 .283 .081 .031
10	You felt "on the go" or "driven by a motor."	Never or Rarely Sometimes Often	.324 .379 .196 .101	.432 .362 .143 .063

Question #	Question	Response Set	Males	Females
		Very Often		
12	You talked too much.	Never or Rarely	.393	.329
		Sometimes	.305	.305
		Often	.171	.180
		Very Often	.131	.186
14	You blurted out answers before the questions had been completed.	Never or Rarely	.330	.453
		Sometimes	.435	.387
		Often	.168	.114
		Very Often	.066	.046
16	You had difficulty awaiting your turn.	Never or Rarely	.439	.557
		Sometimes	.386	.342
		Often	.126	.076
		Very Often	.050	.025

TABLE 2

Model Comparisons, Fit Indices, and Class Proportions for Male Subjects

	No. of Factors (f) or Classes (c)	Log Likelihood	No. of Estimated Parameters	AIC	BIC	Adjusted BIC	L-M-R p	Class Proportions (%)
EFA	1f	-115205.8	68	230547.5	231011.1	230795.0		
	2f	-113919.5	84	228006.9	228579.6	228312.7		
	3f	-113434.1	99	227066.2	227741.1	227426.5		
	4f	-113094.9	113	226415.8	227186.1	226827.0		
LCA	2c	-119805.5	103	239817.1	240519.3	240192.0	.0	56.7, 43.3
	3c	-115824.9	155	231959.7	233016.4	232523.9	.0	42.9, 37.9, 19.1
	4c	-114443.9	207	229301.8	230713.0	230055.2	.0	41.8, 30.2, 18.8, 9.1
	5c	-113521.3	259	227560.6	229326.3	228503.3	.0	33.7, 29.1, 17.4, 13.5, 6.2
	6c	-112786.6	311	226195.1	228315.3	227327.0	.0	30.8, 20.2, 17.7, 14.4, 12.0, 4.9
	7c	-112319.3	363	225364.6	227839.3	226685.8	.76	28.7, 17.1, 15.5, 12.5, 11.5, 9.6, 5.1
	8c	-112013.3	415	224856.6	227685.8	226367.0	.76	26.3, 18.4, 17.4, 9.3, 8.7, 8.5, 7.4, 4.0
FMM	2f1c	-114487.0	69	229112.0	229582.4	229363.2		
	2f2c	-112294.9	124	224837.9	225683.2	225289.2	.75	72.2, 27.8
	2f3c	-111769.2	179	223896.5	225116.8	224548.0	.71	49.1, 35.5, 15.5
	2f4c	-111268.9	234	223005.2	224600.5	223856.9	.76	35.5, 28.4, 26.0, 10.0
	2f5c	-110989.9	289	222577.9	224528.1	223609.7	.77	31.5, 23.1, 18.9, 15.2, 11.2

Note. $N = 6,751$. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; L-M-R = Lo-Mendell-Rubin; EFA = exploratory factor analysis; LCA = latent class analysis; FMM = factor mixture modeling. Boldface represents the best fitting model within each subtype.

TABLE 3

Model Comparisons, Fit Indices, and Class Proportions for Female Subjects

	No. of Factors (f) or Classes (c)	Log Likelihood	No. of Estimated Parameters	AIC	BIC	Adjusted BIC	L-M-R p	Class Proportions
EFA	1 f	-113667.8	68	227471.6	227942.9	227726.8		
	2 f	-112062.7	84	224293.4	224875.6	224608.6		
	3 f	-111602.7	99	223403.3	224089.4	223774.8		
	4 f	-111274.3	113	222774.7	223557.8	223198.7		
LCA	2 c	-118509.0	103	237223.9	237937.7	237610.4	.0	59.3, 40.7
	3c	-114792.4	155	229894.7	230968.9	230476.3	.0	47.1, 42.2, 10.7
	4 c	-113347.4	207	227108.6	228543.1	227885.3	.0	40.9, 30.0, 20.3, 8.8
	5c	-112415.9	259	225349.9	227144.7	226321.7	.79	30.0, 29.8, 17.9, 14.4, 7.9
	6c	-111805.9	311	224233.8	226389.0	225400.8	.0	28.9, 28.4, 15.1, 14.8, 9.6, 3.0
	7c	-111276.3	363	223278.5	225794.2	224640.6	.0	22.2, 21.4, 18.8, 15.6, 11.0, 7.9, 3.0
	8 c	-110926.5	415	222682.9	225558.9	224240.1	.78	25.1, 19.4, 16.5, 14.6, 11.5, 5.9, 4.5, 2.5
FMM	2 f 1 c	-113040.9	69	226219.9	226698.0	226478.8		
	2f 2c	-111404.9	124	223057.9	223917.2	223523.1	.0	64.9, 35.1
	2f 3c	-110591.9	179	221541.7	222782.2	222213.4	.0	43.6, 38.8, 17.6
	2f 4c	-110261.1	234	220990.2	222611.8	221868.2	.76	36.5, 29.7, 21.9, 11.9
	2f 5c	-110070.4	289	220718.9	222721.7	221803.3	.76	38.6, 28.6, 17.9, 7.9, 6.9

Note. $N = 7,556$. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; L-M-R = Lo-Mendell-Rubin; EFA = exploratory factor analysis; LCA = latent class analysis; FMM = factor mixture modeling. Boldface represents the best fitting model within each subtype.