Are There Distinct Cognitive and Motivational Sub-groups of Children with ADHD?

Rikke Lambek[1*], Edmund Sonuga-Barke[2], Rosemary Tannock[3], Anne Virring Sørensen[4], Dorte Damm[4], Per Hove Thomsen[1,4]

[1]Aarhus University, Denmark

[2]Department of Child & Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, Kings College, United Kingdom & Ghent University, Belgium

[3]University of Toronto & Hospital for Sick Children, Canada

[4]Aarhus University Hospital, Denmark

Wordcount: 4820

*Corresponding author: Rikke Lambek, Department of Psychology and Behavioral Sciences, Aarhus University, Bartholins Allé 9, 8000 Aarhus C, Denmark. Email: rikkel@psy.au.dk

Abstract

Background. Attention-deficit/hyperactivity disorder (ADHD) is proposed to be a neuropsychologically heterogeneous disorder that encompasses two distinct sub-groups, one with executive function (EF) deficits and one with delay aversion (DA). However, such claims have often been based on studies that have operationalized neuropsychological deficits using a categorical approach - using intuitive but rather arbitrary, clinical cut-offs. The current study applied an alternative empirical approach to sub-grouping in ADHD, latent profile analysis (LPA), and attempted to validate emerging subgroups through clinically relevant correlates.

Methods. One-hundred medication-naïve children with ADHD and 96 typically developing children (6-14 years) completed nine EF and three DA tasks as well as an odor identification test. Parents and teachers provided reports of the children's behavior (ADHD and EF). Models of the latent structure of scores on EF and DA tests were contrasted using confirmatory factor analysis (CFA). LPA was carried out based on factor scores from the CFA and sub-groups were compared in terms of odor identification and behavior.

Results. A model with one DA and two EF factors best fit the data. LPA resulted in four sub-groups that differed in terms of general level of neuropsychological performance (ranging from high to very low), odor identification, and behavior. The subgroups did not differ in terms of the relative EF and DA performance. Results in the ADHD group were replicated in the control group.

Conclusions. While EF and DA appear to be dissociable constructs; they do not yield distinct sub-groups when sub-grouping is based on a statistical approach like LPA.

Keywords: ADHD, executive function, delay aversion, latent profile analysis (LPA).

Introduction

Over the past decade there has been a shift towards accepting that attention-deficit/hyperactivity disorder (ADHD) is a neuropsychologically heterogeneous condition (Faraone *et al.* 2015). This is based on evidence demonstrating that individuals with ADHD display strikingly different profiles of impairment across neuropsychological domains (Coghill *et al.* 2014). Scientifically, this has been important as it has led to the emergence of multiple pathway models of ADHD pathophysiology. The dual pathway model (Sonuga-Barke, 2002), the first such account published, for instance, hypothesized two distinct neuro-developmental pathways to ADHD, one implicating alterations in dorsal front-striatal circuits and marked by deficits in executive function (EF) and a second implicating disrupted ventral rewards circuits and marked by delay aversion (DA) - a motivational orientation that leads individuals to escape or avoid delay (Sonuga-Barke *et al.* 1992; Marco *et al.* 2009). Since its inception this model has been extended to encompass three (Sonuga-Barke *et al.* 2010) and then six (Coghill *et al.* 2014) pathways and inspired other multiple pathway accounts. The shift in perspective is also important clinically because it suggests that tailored treatment may be needed that allow for the more effective targeting of the particular pattern of underlying impairment seen in specific patients. For instance, a patient with EF deficits may benefit from computerized training in executive control skills while those with DA should respond to DA training.

Initially evidence for heterogeneity came from the finding that different domains of neuropsychological impairment within ADHD were dissociable – for instance, Solanto *et al.* (2001) found DA and EF deficits to be uncorrelated but independently associated with ADHD. Researchers have since taken this way of thinking to the next step – by proposing that ADHD populations can be divided up into sub-groups with distinct neuropsychological profiles, that is, one with EF deficits, one with DA, one with EF deficits *plus* DA and one without EF deficits or DA. However, studies

vary considerably with respect to the proportion of individuals found in a given sub-group and whether a ”pure” DA subgroup is identified (Sonuga-Barke *et al.* 2003; Nigg *et al*., 2005; Sonuga-Barke *et al.* 2010; Karalunas & Huang-Pollock, 2011; Sjöwall *et al.* 2013).

Such studies have typically operationalized neuropsychological deficits in categorical terms with some individuals meeting certain thresholds defined as having a deficit and those not meeting that threshold as impairment free. Furthermore, the thresholds used have been arbitrary - cut-offs equivalent to the top 10 percent of impaired individuals within the normal population. While such an approach has the benefit of being clinically intuitive it lacks a solid empirical foundation which potentially biases analyses towards finding pure neuropsychologically impaired sub-groups. In this paper we adopt a different approach to the task of neuropsychological sub-grouping - latent profile analysis (LPA). This is a statistical method for identifying homogeneous sub-groups of participants (latent classes) who share similar patterns of observed scores such as those from neuropsychological tasks. This method has been used recently to successfully identify neuropsychological sub-groups in ADHD. Gomez *et al.* (2014) found three working memory (WM) sub-groups differentiated primarily by level of performance. Rajendran *et al.* (2015) found a weak attention/EF sub-group and a pervasive neuropsychological dysfunction sub-group. Rommelse *et al*. (2016) identified four sub-groups that mainly differed in their speed–accuracy tradeoffs. Finally, van Hulst *et al.* (2015) found a sub-group with poor cognitive control and one with poor timing. Reward mechanisms were only examined in the latter study, where no evidence for a sub-group with reward processing problems was found.

In the current study we extend the use of LPA to see if distinct classes with DA and EF deficits can be identified in the ADHD population. The current study had a number of strengths compared to previous studies of neuropsychological heterogeneity. First, we employed multiple tasks within the EF and the DA domains and applied confirmatory factor analysis (CFA) to obtain

"purer" latent variables. Second, the children were all medication-naïve. Third, we attempted to validate any sub-groups found by comparing them in terms of a range of clinically relevant correlates. Odor identification was examined because of the proposed association between this ability and orbitofrontal cortex integrity (Killgore *et al*. 2008). As DA is hypothesized to implicate ventral regions, distinct DA and EF sub-groups may differ with respect to odor identification. A questionnaire-based EF assessment was included because it is hypothesized to capture other facets of EF than the more fine-grained controlled task-based assessment, specifically the child's ability to problem-solve in everyday situations (Huizinga & Smidts, 2011; Toplak *et al.* 2013) and possibly task-based sub-groups will differ with respect to this EF facet. ADHD behavior as well as comorbidity were also examined. Finally, the study included a large sample of unaffected individuals so that the results could be examined as a function of ADHD.

The current study therefore had three major aims. The first aim was to determine whether EF and DA are dissociable constructs, by testing alternative models using CFA based on EF and DA tasks. The second aim was to test for the existence of distinct homogeneous neuropsychological classes using LPA. The final aim was to investigate the external validity of the neuropsychological sub-groups identified, by testing for differences in terms of ADHD symptomatology, questionnaire-based EF, and odor identification.

On the basis of prior work we hypothesized that (i) EF and DA would be dissociable and that (ii) LPA would result in at least three ADHD sub-groups - one with "pure" EF deficits, one with "pure" DA, and one with EF deficits plus DA. No formal hypotheses were formulated regarding potential sub-group differences due to inconclusive previous research.

Methods

Participants

Over a period of two years, all children (n = 308) between the ages of 6 and 14 years, consecutively referred to an ADHD specialist clinic at a Danish public regional child and adolescent psychiatric hospital for a diagnostic assessment of ADHD were invited to participate in the study together with their parents. Almost half of the invited families (n = 137) agreed to participate. To be included in the present study the children had to meet DSM-IV criteria for an ADHD diagnosis (n = 129) based on the Development and Well-Being Assessment (DAWBA; Goodman *et al.* 2000). The DAWBA (a structured diagnostic parent interview and a teacher questionnaire) was administered by a psychologist or a doctor and integrated and reviewed by a child and adolescent psychiatrist as part of the clinical examination at the hospital. Children with pervasive developmental disorders (n = 12), an IQ below 70 (n = 6), or a history of treatment with psychotropic medication[1] (n = 10) were excluded, as was one member of one sib-pair, leaving a final sample of 100 children. The medication status of the children at the time of the study was due to the fact that pharmacological treatment is not initiated until after the diagnostic assessment at the hospital. Demographic characteristics are presented in Table 1.

A total of 120 same-age controls attending schools located within the same region as the hospital agreed to participate. According to parent reports, none of these children had previous or current psychiatric diagnoses or use of psychotropic medication. Children with an IQ below 70 (n = 2), impaired vision (n = 1), or missing parent questionnaires (n = 2) as well as 19 siblings were excluded, leaving 96 children in the control sample. One child met thresholds on the Strengths and Weaknesses of ADHD-Symptoms and Normal-Behavior (SWAN; Swanson *et al.* 2001), but was not excluded in order to obtain a representative control sample.

[1]One child who received methylphenidate for one week, one and a half years prior to participating in the study, was not excluded.

Families were contacted via mail through the hospital or schools. Children with ADHD were tested at the hospital in two sessions each lasting approximately 1½-hours, whereas typically developing children were tested at the local university/school in one session lasting approximately 2.5-hours. Questionnaires were distributed to parents at the child's test session and sent to teachers via mail. The children earned small prizes (e.g., stationery and coins) on some tasks and a small present at the end of the test session. The study was approved by the Danish regional ethics committee and informed consent was obtained from the participants.

>Table 1<

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Measures

Tasks

The neuropsychological test battery consisted of nine EF (three WM, three shifting [SH], and three response inhibition [IN]) tasks and three DA tasks. One outcome per task was selected a priori and based on previous research involving the specific task (see below) and/or meta-analyses (see e.g., Willcutt *et al.* 2005) as well as previous research on the organization of EF and DA (e.g., Coghill *et al.*, 2014; Huizinga *et al.* 2006; Miyake *et al.* 2000). All outcome measures were transformed such that lower scores were indicative of poorer performance (task details are provided in Supplementary Table S1 available online).

Working Memory

Tic Tac Toe (Huizinga *et al.* 2006). The respondent indicated when a visual target pattern was reproduced. Outcome: Mean accuracy.

Mental Counters (Huizinga *et al.* 2006). The respondent kept track of the numerical value of two independent counters. Outcome: Mean accuracy.

Finger Windows Backwards (Bedard & Tannock, 2008). The respondent indicated the reverse sequential placement of a series of holes in a card. Outcome: Total accuracy.

Shifting

Local–Global (LGT; Huizinga *et al.* 2006). The respondent switched between rule sets (local vs. global). Outcome: Median reaction time (RT) in milliseconds (ms).

Dots-Triangles (DOT; Huizinga *et al.* 2006). The respondent switched between rule sets (up-down vs. left–right). Outcome: Median RT in ms.

Trail Making Test (Reitan, 1971). The respondent switched between consecutive numbers and letters. Outcome: Total RT in ms.

Response inhibition

Flanker Task (Huyser *et al.* 2011). The respondent focused on a central arrow while ignoring either congruent or incongruent peripheral arrows. Outcome: Total accuracy on incongruent trials.

Stop-signal task (SST; Williams *et al.* 1999). The respondent responded to a visual go signal but refrained from responding when an auditory stop signal was presented. Outcome: Mean stop-signal reaction time in ms.

Walk Don't Walk from the Test of Everyday Attention of Children (Manly *et al.* 1999). The respondent responded to one auditory (go) signal but refrained from responding to a different auditory (no-go) signal. Outcome: Total accuracy.

Delay aversion

Choice-Delay Task (C-DT; Sonuga-Barke et al. 1992; Solanto et al. 2001). The respondent chose between a small immediate reward and a larger delayed one. Outcome: Percent short small rewards.

Maudsley's Index of Childhood Delay Aversion (MIDA; Kuntsi *et al.* 2001). The respondent chose between shooting one spaceship (yielding a small immediate reward) or two spaceships (yielding a larger delayed reward). Outcome: Percent short small rewards.

Delay Frustration Task (DeFT; Bitsakou *et al.* 2006). The respondent solved a series of math questions with incorporated delays. Outcome: Mean total duration in ms.

General cognitive function

Vocabulary and Matrix Reasoning from the Wechsler Intelligence Scale for Children, 4th edition (Wechsler, 2003) were applied to estimate general cognitive function.

Questionnaires and odor identification test

This battery included a background questionnaire about the parents' educational level (years) applied as an indicator of socioeconomic status (SES), the SWAN (Inattention and Hyperactivity/impulsivity scales; Swanson et al. 2001), and the Behavior Rating Inventory of Executive Function (BRIEF; WM, Shift, and Inhibit scales; Gioia *et al.* 2000) which measures DSM congruent ADHD behavior and EF behaviors in children respectively. The Brief Smell Identification Test (BSIT; Doty, 2001) was applied as a test of olfactory function in the child. Means were applied in the questionnaire analyses, whereas total number of errors was examined in the odor identification test. In both cases higher scores were indicative of more perceived difficulties (questionnaire details and a description of the odor identification test are provided in Supplementary Table S1 online). Information on child comorbidity was based on the DAWBA. The correlations between the BRIEF scales and the executive function tasks are provided in online Supplementary Table S2.

Data preparation

Missing task-data at case level (< 1%) resulted from apparatus malfunction etc. If accuracy of performance on the LGT, the DOT, and the SST go task was lower than expected by chance, these values were coded as missing, as were blocks on the SST if inhibition occurred on fewer than 20% or more than 80% of stop trials. Following Huizinga *et al.* (2006) trials on tasks with an RT outcome were excluded if they were incorrect, preceded by an incorrect response, shorter than 100 ms, or more than 2.5 SD from the mean for the individual child and task. The missing data and the data removed as a result of the subsequent trimming procedures ranged between 0-9% depending on the task. There were no missing parent questionnaires, but questionnaires from 59 teachers (9 in the ADHD group; 50 in the control group) and the BSIT from two children with ADHD were missing. There were less than 1% missing data at item level.

Analysis

First, the latent structure of EF and DA was tested using CFA. Six alternative models were specified and tested (see Supplementary Figure S3 available online). The latent variables were specified to be correlated and all error variances were uncorrelated. Model parameters were estimated using robust maximum likelihood (MLR) using all available data. Model fit was assessed using the chi-square statistic (χ^2), the comparative fit index (CFI), the Tucker-Lewis Index (TLI), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). A non-significant χ^2, CFI and TLI values above .95, and RMSEA and SRMR values below .08 indicate reasonable fit. AIC and BIC are used to compare competing models with lower values indicating better fit. A difference greater than 10 between two BIC values was considered indicative of a "significant" difference (Raftery, 1995). Finally, parsimony was considered when selecting the best model.

Second, LPA was conducted on the factor scores. Age, SES, and IQ were entered as covariates as they have each been associated with neuropsychological performance (Brydges *et al.* 2012; Hackman *et al.* 2015). LPA models with two to six classes were estimated. Model fit was assessed using the likelihood ratio chi-square (LRχ^2), the AIC, the BIC, entropy values, and the Lo-Mendell-Rubins adjusted likelihood ratio test (LMR-A). AIC and BIC values were interpreted as before. The AIC and BIC values were also taken into consideration when evaluating the improvement made by adding an additional class as a small difference is considered indicative of little model improvement (DiStefano & Kamphaus, 2006). Entropy values greater than .80 indicate acceptable classification accuracy (Burnham & Anderson, 2004). The LMR-A is used to compare models with increasing numbers of latent classes. A non-significant value ($p > .05$) suggests that the model with one less class should be accepted (Lo *et al.* 2001). Again parsimony was taken into consideration when selecting the best solution.

Finally, χ^2 and one-way Analysis of Variance tests were applied to examine potential sub-group differences. Post hoc tests were (Bonferroni) adjusted for multiple testing.

CFA and LPA were carried out using Mplus 7.4 (Muthén & Muthén, 1998-2016). IBM SPSS Statistics V24.0 (IBM Corp, 2016) was applied when comparing sub-groups.

Consistent with previous research within the field (e.g., van Hulst *et al.*, 2015) and to be able to address neuropsychological function in ADHD specifically, children with ADHD and typically developing children were examined separately.

Results

Confirmatory factor analysis

One indicator of DA (the DeFT) was unrelated to the other two (C-DT and MIDA) with non-significant correlations of .117 and .093 in the ADHD group (and .090 and .148 in the control group) and was excluded from further analysis. The fit statistics from the CFA based on the ADHD

sample are reported in Table 2. Models 4 and 6 clearly fitted the data better than the other models as both had a non-significant χ^2, a CFI value above .95, a TLI value of .93, and RMSEA and SRMR values below .08. The values for the AIC and the BIC were lower for Model 4 than Model 6 and the difference between the BIC values was greater than 10. Furthermore, Model 4 was more parsimonious (having fewer parameters) than Model 6. On this basis, Model 4 with three factors (WMIN, SH, and DA) was considered to be the "best" model. Factor loadings and factor correlations for Model 4 are presented in Table 3. The factor loadings were statistically significant and ranged from .346 to .839. The factor correlations were statistically significant with absolute values ranging from .427 to .646. Model 4 was also considered the "best" model when tested with typically developing children ($\chi2$ = 43.445, df = 41, p = .368; CFI = .989; TLI = .985; RMSEA = .025, 90% CI = [.000-.076]; SRMR = .056; AIC = 8090.636; BIC = 8182.953; see online Supplementary Tables S4 and S5)[2]. Finally, standardized factor scores for each latent variable were calculated based on Model 4.

>Table 2<

>Table 3<

Latent profile analysis

When tested with children with ADHD (see Table 4) the four class solution had a statistically significant LMR-A value and a high entropy value (> .80). This solution also had the lowest BIC value and a negligible lowering of the AIC value from the four to the six class solution. Based on these results and taking parsimony into account, the four class solution was considered the "best" solution. The three class solution was considered "best" for typically developing children as it had a significant LMR-A value, a high entropy value, and the lowest BIC value (LRχ^2 = -221.465; AIC =

[2]The factor model prevailed when tested with the overall sample (children with ADHD plus typically developing children).

482.931; BIC = 534.218; Entropy = .933; LMR-A = 67.773, p = .010; see online Supplementary Table S6).

>Table 4<

The performance profiles for the four classes in the ADHD group are illustrated in Fig. 1a. Inspection of the classes suggested that Class 1 (n = 32) represented children with high WMIN, SH, and DA factor scores. Class 2 (n = 35) included children with average scores. Class 3 (n = 26) included children with low scores and class 4 (n = 7) children with very low scores. The four classes differed significantly (all ps < .00001 and large effect sizes) with respect to WMIN, SH, and DA performance, suggesting that they constituted different severity levels along a skewed underlying continuum rather than qualitatively distinct subgroups. This pattern was reproduced when the control group was examined (Fig. 1b)[3].

>Fig. 1<

Sub-group comparisons

The four ADHD classes differed significantly (all ps < .05 and medium to large effect sizes) with respect to inattention (parent-report), hyperactivity-impulsivity (parent- and teacher-report), WM behavior (teacher-report), SH and IN behavior (parent-report), and odor identification. In all instances post hoc analyses indicated that one or more classes with lower WMIN, SH, and DA factor scores (indicative of poorer performance) obtained higher ratings (indicative of more perceived difficulties) on the SWAN and the BRIEF scales and identified fewer odors correctly than one or more classes with higher factor scores. There were no significant associations between class and any type of comorbidity (externalising or internalising disorders; Supplementary Table S7 available online). The three classes in the typically developing group differed significantly (all ps <

[3]A number of classes that differed only quantatively were also found when the LPA was conducted on the overall sample (children with ADHD plus typically developing children) or without IQ and SES as covariates.

.05 and medium effect sizes) with respect to WM and SH behavior (parent-report) as well as odor identification (again classes with lower factor scores obtained higher parent-ratings and identified fewer odors; online Supplementary Table S7). In supplementary analyses similar latent classes from each sample were compared[4] (i.e., ADHD class 1 vs. typically developing class 1 etc.) with respect to the correlates. This showed that all ADHD classes had significantly higher levels of ADHD and EF behavior than the corresponding typically developing classes (all ps < .0001 and η^2 ranging from .38-.78). Individuals in the ADHD classes also identified fewer odors correctly than individuals in the corresponding typically developing classes, although only the class 1 comparison was statistically significant ($p < .05$, $\eta^2 = .07$).

Discussion

This study examined the presence of distinct cognitive and motivational sub-groups in children with ADHD using LPA. The key findings were as follows.

First, EF and DA were found to be dissociable constructs. This is consistent with our initial hypothesis and the basic tenets of the dual pathway model (Sonuga-Barke, 2002). However, contrary to some previous research (e.g., Solanto *et al.* 2001), they were not found to be unrelated as the DA factor correlated moderately with each of the two EF factors (ranging from .427 to .552 in the ADHD group). While the nature of this relationship is less clear, it may be one where DA influences learning over time thereby hindering cognitive development (Sonuga-Barke, 2002). This idea has received some support as DA has been found to predict WM scores for school-age children with and without ADHD (Karalunas & Huang-Pollock, 2011). To determine whether DA and EF are less associated in younger children and become increasingly interrelated over the course of development, future studies should examine preschool samples and follow them longitudinally.

[4]As there was no class 4 equivalent in the typically developing group only classes 1-3 were compared.

Second, the LPA supported the presence of four ADHD sub-groups distinguished by level of neuropsychological performance (ranging from high to very low), but not in terms of the relative levels of EF and DA performance. Consequently, there was no evidence for a group affected by either "pure" EF deficits or by "pure" DA. This strongly supports neuropsychological heterogeneity in ADHD but at the same time is more consistent with a quantitative rather than a qualitative model of heterogeneity. As distinct EF and DA sub-groups have previously been identified using a categorical approach (e.g., Sonuga-Barke *et al*. 2003; Nigg *et al*. 2005; Sonuga-Barke *et al.* 2010), this was a surprising finding. It is unlikely that the heterogeneity in findings are explained by some unique characteristic of the current ADHD sample. While the children were medication-naïve, their medication status reflected the fact that any medication initiation occurred after the examination at the hospital as is practice in Denmark and not their degree of ADHD. The sample had an ADHD subtype distribution, a gender distribution, a mean group IQ, and a comorbid profile comparable to what is traditionally reported in the ADHD literature for this age group suggesting that the children included in the present study constituted a representative ADHD sample. Further, IQ and SES were included as covariates in our LPA models, but this does not appear to be the reason why no distinct EF and DA sub-groups were found, as running the analyses without the covariates also yielded a quantitative model. Possibly, traditional categorical sub-grouping based on a thresholding method identifies the most impaired individuals on a given measure (relative to controls), but not necessarily those individuals who have the most in common "neuropsychologically-speaking". By comparison, this may be the strength of LPA, as this approach identifies sub-groups of children who are similar based on a number of characteristics (such as neuropsychological performance) and without using a control group as reference point or *a priori* cut-offs. Rather, LPA makes no a priori assumptions about the number of classes as they are decided upon based on objective statistical criteria. With respect to the results from our LPA, it is important to note that although EF and DA

appear to go hand in hand this implies neither unity of domains nor primacy of one over the other. It is possible that EF and DA are supported by different neural pathways and make distinctive contributions to child development and ADHD, but do so concomitantly in school-age.

Recently, Gomez *et al.* (2014) also found a continuum when applying LPA to examine WM performance in ADHD. However, other studies applying empirical approaches such as LPA have identified qualitative differences in neuropsychological performance in ADHD when examining domains such as response variability (Fair *et al.* 2012), processing speed (Bergwerff *et al.*, 2017), timing (van Hulst *et al.* 2015), or speed-accuracy tradeoffs (Rommelse *et al*. 2016) and it is possible that inclusion of these domains would have resulted in qualitatively distinct sub-groups in the present study. Van Hulst *et al.* examined reward sensitivity using an anticipation paradigm, but did not find a sub-group with poor reward sensitivity. By comparison Mostert *et al.* (2015) found a sub-group with impaired delay discounting when examining an adult sample. The different results found in sub-grouping studies may be mainly attributable to the use of different task paradigms, dissimilar outcomes (e.g., accuracy vs. reaction times), varying levels at which the variables are measured (e.g., observed vs. latent), different statistical approaches, and age-varying samples (e.g., children vs. adults). Future studies should attempt to cross-validate previously identified sub-groups using several competing and rigorous methods, and examine their clinical utility over time, for instance as moderators of treatment effects.

Third, we found some evidence for the levels of neuropsychological performance distinguishing the sub-groups based on experimental tasks to be mirrored in the children's levels of ADHD symptoms, (questionnaire-based) EF behavior, and odor identification ability. Olfactory function has been proposed as a biological marker for prefrontal, hippocampal, and dopaminergic function (Killgore *et al*. 2008; Schecklmann *et al*. 2011; Hagemeier *et al.* 2016). Whereas odor identification logically did not differentiate between DA and EF in the present study as qualitatively

distinct sub-groups were not found, children with lower levels of neuropsychological performance did identify fewer odors correctly than did children with higher levels. This could be interpreted as suggesting that individuals with lower levels of neurological performance have more compromised brain function. However, olfactory function has only been examined sporadically in pediatric ADHD (e.g., Karsz *et al.* 2008; Romanos *et al.* 2008; Ghanizadeh *et al.* 2012) and more research is clearly needed before any conclusions can be drawn as to the association between olfactory and neuropsychological function in ADHD. Overall, the findings from the sub-group comparisons support the external validity of the identified sub-groups and suggest the presence of a continuum with the most impaired individuals at one end.

Finally, the neuropsychological profiles identified in the ADHD group were replicated in the control group, suggesting that there are also different levels of neuropsychological performance in typically developing children. The idea of neuropsychological heterogeneity in the general population has also received support in previous studies (e.g., Bergwerff *et al.* 2017; Fair *et al.* 2012; Rommelse *et al.*, 2016; van Hulst *et al.* 2015). Together these results suggest that not only are there similar patterns of heterogeneity across samples, but that some children with ADHD probably have more in common with non-ADHD counterparts than with other children with ADHD with respect to neuropsychological performance - such similarities may potentially inform on risk and protective factors associated with ADHD.

Strengths and Limitations

The study has several strengths, including being (to our knowledge) the first time LPA has been applied to examine EF *and* DA using a choice-delay paradigm in school-age children with ADHD. CFA was applied to obtain "purer" latent variables, just as CFA models were theory-driven and examined based on outcomes from eleven different tasks and one outcome per task. In addition, a large sample of medication-naïve children with ADHD was included in the study. There are also

limitations. Due to the amount of missing teacher questionnaires in the control group, results from these analyses should be interpreted with caution. As several neuropsychological sub-groups were identified and some groups included relatively few children, the study may not have had sufficient statistical power to detect additional intergroup differences. However, the intergroup differences that were detected were all in the same direction which supports the findings.

Conclusions

Sub-grouping based on LPA results in quantitatively rather than qualitatively distinct sub-groups characterized by different levels of neuropsychological performance but similar levels of EF and DA performance at each level. The neuropsychological performance levels distinguishing the sub-groups based on experimental tasks are to some degree mirrored in the children's levels of ADHD symptoms and questionnaire-based EF as well as their ability to identify odors suggesting a continuum with the most impaired individuals at one end. It is possible that a focus on neuropsychological severity may be informative for implementing interventions; that is, the efficacy of interventions that target neuropsychological problems may depend on level of impairment in the intervention group. Furthermore, examination of neuropsychological heterogeneity in the general population may contribute to the understanding of factors associated with the development of ADHD. As this was the first time EF and DA as measured using a choice-delay paradigm have been examined together using LPA, the results are in need of replication.

Supplementary material

For supplementary material accompanying this paper visit >insert link<.

Acknowledgements

The study was supported by research grants from the TRYG Foundation (grant number 7-10-0098, 101183). We thank the children, parents, and teachers, who participated in the study. We also thank Anne Katrine Hansen, Mette Daubjerg, Rikke Hansen, and Jytte Breining for their contribution

during the data collection process and the researchers who kindly gave us permission to use their tasks, especially Mariëtte Huizinga.

Declaration of interests

Edmund Sonuga-Barke has received speaker fees, consultancy, research funding, and conference support from Shire, speaker fees from American University of Beirut and Janssen Cilag, consultancy from Neurotech solutions, Copenhagen University and Berhanderling, Skolerne, and KU Leuven, and book royalties from OUP and Jessica Kingsley. He is Editor-in-Chief JCPP – supported by a buy-out of time to University of Southampton and a personal Honorarium.

Rosemary Tannock has received honoraria from Shire, Eli Lilly, Medice, and OpoPharma, research support from Pearson-Cogmed, and was a member of the DSM-5 Workgroup for ADHD and Externalizing Disorders.

Dorte Damm has received honoraria from Shire.

Per Hove Thomsen has received honoraria from Shire and Medice.

References

Bedard AC, Tannock R (2008). Anxiety, methylphenidate response, and working memory in children with ADHD. *Journal of Attention Disorders* **11**, 546-557.

Bergwerff CE, Luman M, Weeda WD, Oosterlaan J (2017). Neurocognitive profiles in children with ADHD and their predictive value for functional outcomes. *Journal of Attention Disorders*. Published online 30 January 2017. doi: 10.1177/1087054716688533.

Bitsakou P, Antrop I, Wiersema JR, Sonuga-Barke EJ (2006). Probing the limits of delay intolerance: Preliminary young adult data from the delay frustration task (DeFT). *Journal of Neuroscience Methods* **151**, 38-44.

Brydges CR, Reid CL, Fox AM, Anderson M (2012). A unitary executive function predicts intelligence in children. *Intelligence* **40**, 458-469.

Burnham KP, Anderson DR (2004). Multimodal inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research* **33**, 261–304.

Coghill DR, Seth S, Matthews K (2014). A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: Advancing beyond the three-pathway models. *Psychological Medicine* **44**, 1989-2001.

DiStefano C & Kamphaus RW (2006). Investigating subtypes of child development: A comparison of cluster analysis and latent class cluster analysis in typology creation. *Educational and Psychological Measurement* **66**, 778-794.

Doty RL (2001). The Brief Smell Identification Test Administration Manual, Sensonics Inc: Haddon Heights.

Fair DA, Bathula D, Nikolas MA, Nigg JT (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *PNAS Proceedings of the National Academy of Sciences of the United States of America* **109**, 6769-6774.

Faraone SV, Asherson P, Banaschewski T, Biederman J, Buitelaar JK, Ramos-Quiroga JA, Rohde LA, Sonuga-Barke EJS, Tannock R, Franke B (2015). Attention-deficit/hyperactivity disorder. *Nature Reviews: Disease Primers* **1**, 15020.

Ghanizadeh A, Bahrani M, Miri R, Sahraian A (2012). Smell identification in children with attention deficit hyperactivity disorder. *Psychiatry Investigation* **9**, 150-153.

Gioia GA, Isquith PK, Guy SC, Kenworthy L (2000). *Behavior rating of executive function*. Psychological Assessment Resources: Lutz, FL.

Gomez R, Gomez RM, Winther J, Vance A (2014). Latent profile analysis of working memory performance in a sample of children with ADHD. *Journal of Abnormal Child Psychology* **42**, 1367-1379.

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**, 645-656.

Hackman DA, Gallop R, Evans GW, Farah MJ (2015). Socioeconomic status and executive function: Developmental trajectories and mediation. *Developmental Science* **18**, 686-702.

Hagemeier J, Woodward MR, Rafique UA, Amrutkar CV, Bergsland N, Dwyer MG, Benedict R, Zivadinov R, Szigeti K (2016). Odor identification deficit in mild cognitive impairment and Alzheimer's disease is associated with hippocampal and deep gray matter atrophy. *Psychiatry Research: Neuroimaging* **255**, 87-93.

Huizinga M, Dolan CV, van der Molen MW (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia* **44**, 2017-2036.

Huizinga M, Smidts DP (2011). Age-related changes in executive function: A normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology* **17**, 51-66.

Huyser C, Veltman DJ, Wolters LH, de Haan E, Boer F (2011). Developmental aspects of error and high-conflict-related brain activity in pediatric obsessive-compulsive disorder: A fMRI study with a flanker task before and after CBT. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **52**, 1251-1260.

IBM Corp (2016). *IBM SPSS statistics for windows, version 24.0*. IBM Corp: Armonk, N.Y.

Karalunas SL, Huang-Pollock CL (2011). Examining relationships between executive functioning and delay aversion in attention deficit hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology* **40**, 837-847.

Karsz FR, Vance A, Anderson VA, Brann PG, Wood SJ, Pantelis C, Brewer WJ (2008). Olfactory impairments in child attention-deficit/hyperactivity disorder. *The Journal of Clinical Psychiatry* **69**, 1462-1468.

Killgore WDS, Mcbride SA, Killgore DB, Balkin TJ, Kamimori GH (2008). Baseline odor identification ability predicts degradation of psychomotor vigilance during 77 hours of sleep deprivation. *International Journal of Neuroscience* **118**, 1207-1225.

Kuntsi J, Oosterlaan J, Stevenson J (2001). Psychological mechanisms in hyperactivity: I. response inhibition deficit, working memory impairment, delay aversion, or something else? *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **42**, 199-210.

Lo Y, Mendell NR, Rubin DB (2001). Testing the number of components in a normal mixture. *Biometrika* **88**, 767–778.

Manly T, Robertson IH, Anderson V, Nimmo-Smith I (1999). *Test of Everyday Attention For Children: TEA-Ch.* Thames Valley Test Company: Bury St. Edmunds, UK.

Marco R, Miranda A, Schlotz W, Melia A, Mulligan A, Muller U, Andreou P, Butler L, Christiansen H, Gabriels I, Medad S, Albrecht B, Uebel H, Asherson P, Banaschewski T, Gill M, Kuntsi J, Mulas F, Oades R, Roeyers H, Steinhausen HC, Rothenberger A, Faraone SV, Sonuga-Barke EJ (2009). Delay and reward choice in ADHD: An experimental test of the role of delay aversion. *Neuropsychology* **23**, 367-380.

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology* **41**, 49-100.

Mostert JC, Hoogman M, Onnink AMH, van Rooij D, von Rhein D, van Hulzen KJE, Dammers J, Kan CC, Buitelaar JK, Norris DG, Franke B (2015). Similar subgroups based on cognitive performance parse heterogeneity in adults with ADHD and healthy controls.

Journal of Attention Disorders. Published online 14 September 2015. doi: 10.1177/1087054715602332.

Muthén LK, Muthén BO (1998-2016). *Mplus user's guide* (Sixth edn.). Muthén & Muthén: Los Angeles, CA.

Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJ (2005). Causal heterogeneity in attention-deficit/ hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry* **57**, 1224-1230.

Raftery AE (1995). Bayesian model selection in social research. *Sociological Methodology* **25**, 111-163.

Rajendran K, O'Neill S, Marks DJ, Halperin JM (2015). Latent profile analysis of neuropsychological measures to determine preschoolers' risk for ADHD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **56**, 958-965.

Reitan RM (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills* **33**, 575-581.

Romanos M, Renner TJ, Schecklmann M, Hummel B, Roos M, von Mering C, Pauli P, Reichmann H, Warnke A, Gerlach M (2008). Improved odor sensitivity in attention-deficit/hyperactivity disorder. *Biological Psychiatry* **64**, 938-940.

Rommelse NNJ, van der Meer JMJ, Hartmann CA, Buitelaar JK (2016). Cognitive profiling useful for unraveling cross-disorder mechanisms: Support for a step-function endophenotype model. *Clinical Psychological Science* **4**, 957-970.

Schecklmann M, Schaldecker, M, Aucktor, S, Brast J, Kirchgäßner K, Mühlberger A, Warnke A, Gerlach M, Fallgatter AJ, Romanos M (2011). Effects of methylphenidate on olfaction and frontal and temporal brain oxygenation in children with ADHD. *Journal of Psychiatric Research* **45**, 1463-1470.

Sjöwall D, Roth L, Lindqvist S, Thorell LB (2013). Multiple deficits in ADHD: Executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *Journal of Child Psychology and Psychiatry* **54**, 619-627.

Solanto MV, Abikoff H, Sonuga-Barke E, Schachar R, Logan GD, Wigal T, Hechtman L, Hinshaw S, Turkel E (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: A supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology* **29**, 215-228.

Sonuga-Barke EJ (2002). Psychological heterogeneity in AD/HD-a dual pathway model of behaviour and cognition. *Behavioural Brain Research* **130**, 29-36.

Sonuga-Barke E, Bitsakou P, Thompson M (2010). 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* **49**, 345-355.

Sonuga-Barke EJS, Dalen L, Remington B (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *American Academy of Child and Adolescent Psychiatry* **42**, 1335-1342.

Sonuga-Barke EJ, Taylor E, Sembi S, Smith J (1992). Hyperactivity and delay aversion: I. the effect of delay on choice. *Child Psychology & Psychiatry & Allied Disciplines* **33**, 387-398.

Swanson J, Schuck S, Mann M, Carlson C, Hartman C, Sergeant J *et al.* (2001). Over-identification of extreme behaviour in evaluation and diagnosis of ADHD/HKD. (http://www.adhd.net). Accessed 1 May 2006.

Toplak ME, West RF, Stanovich KE (2013). Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry* **54**, 131-143.

van Hulst BM, de Zeeuw P, Durston S (2015). Distinct neuropsychological profiles within ADHD: A latent class analysis of cognitive control, reward sensitivity and timing. *Psychological Medicine* **45**, 735-745.

Wechsler D (2003). *Wechsler intelligence scale for children-4th edition (WISC-IV®).* Harcourt Assessment: San Antonio, TX.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF (2005).Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry* **57**, 1336–1346.

Williams BR, Ponesse JS, Schachar RJ, Logan GD, Tannock R (1999). Development of inhibitory control across the life span. *Developmental Psychology* **35**, 205-213.

Table 1. *Demographic Characteristics*

Variables	ADHD (n = 100) Mean (SD)/n (%)	TD (n = 96) Mean (SD)/n (%)
Age	9.38 (1.84)	9.81 (2.14)
Gender (males)	77 (77.00)	53 (55.21)
WISC-IV		
Estimated IQ	91.60 (14.30)	97.48 (14.43)
Parental education	14.10 (2.34)	15.91 (1.86)
SWAN		
Parent-report ADHD	1.33 (.64)	-.51 (.78)
Teacher-report ADHD[a]	1.55 (.65)	-.88 (1.19)
DAWBA		
ADHD combined type	73 (73.00)	-
ADHD predominantly inattentive type	24 (24.00)	-
ADHD predominantly hyperactive-impulsive type	3 (3.00)	-
Externalising disorders	32 (32.00)	-
Internalising disorders	23 (23.00)	-

DAWBA = Development and Well-Being Assessment; Externalising disorders = conduct disorder, oppositional defiant disorder; Internalising disorders = specific phobia, social phobia, obsessive-compulsive disorder, generalized anxiety disorder, separation anxiety disorder, anxiety disorder NOS, depression, depressive disorder NOS; SWAN = Strengths and Weaknesses of ADHD-Symptoms and Normal-Behavior; TD = typically developing; WISC-IV = Wechsler Intelligence Scale for Children, 4th edition. [a]Missing teacher questionnaires (9 in ADHD group; 50 in TD group).

Table 2. *Fit statistics for the CFA for children with ADHD*

Model[a]	χ^2 (df) *p*	CFI	TLI	RMSEA (90% CI)	SRMR	AIC	BIC
1	93.988 (44) .000	.834	.793	.107 (.077-.136)	.092	8249.785	8335.756
2	75.132 (43) .002	.894	.864	.086 (.053-.118)	.083	8232.146	8320.722
3	72.013 (41) .002	.897	.862	.087 (.052-.120)	.081	8233.237	8327.023
4	54.759 (41) .074	.954	.939	.058 (.000-.095)	.060	8216.814	8310.600
5	72.725 (41) .002	.895	.859	.088 (.054-.120)	.075	8231.101	8324.887
6	51.766 (38) .067	.954	.934	.060 (.000-.098)	.057	8219.848	8321.449

AIC = Akaike information criterion; BIC = Bayesian information criterion; χ^2 = Chi-square; CFI = comparative fit index; CI = confidence interval; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; TLI = Tucker Lewis index.

[a]Model 1 = One factor - common neuropsychological function; Model 2 = Two factors - executive function and delay aversion; Model 3 = Three factors - working memory/shifting, inhibition, and delay aversion; Model 4 = Three factors - working memory/inhibition, shifting, and delay aversion; Model 5 = Three factors - working memory, shifting/inhibition, and delay aversion; Model 6 = Four factors - working memory, shifting, inhibition, and delay aversion.

Table 3. *Factor loadings (standard error) and factor correlations (standard error) for Model 4 for children with ADHD*

	Working Memory/ Inhibition	Shifting	Delay Aversion
Tic Tac Toe (Mean accuracy)	.839 (.041)		
Mental Counters (Mean accuracy)	.721 (.050)		
Finger Windows Backwards (Total accuracy)	.750 (.053)		
Local–Global (Median RT; ms)		.628 (.127)	
Dots–Triangles (Median RT; ms)		.687 (.115)	
Trail Making Test (Total RT; ms)		.714 (.096)	
Flanker Task (IT total accuracy)	.640 (.059)		
Stop-signal task (Mean SSRT; ms)	.346 (.097)		
TEA-Ch (Walk don't Walk total accuracy)	.598 (.072)		
Choice-Delay Task (% short small rewards)			.697 (.094)
Maudsley's index of Childhood DAv (% short small rewards)	.		.751 (.104)
Working Memory/Inhibition	1.00		
Shifting	.646 (.081)	1.00	
Delay Aversion	.552 (.105)	.427 (.125)	1.00

DAv = delay aversion; IT = Incongruent trials; ms = millisecond; RT = reaction time; SSRT = Stop-signal reaction time; TEA-Ch = Test of Everyday Attention for Children.

Table 4. *Fit statistics for the LPA with covariates age, SES, and IQ for children with ADHD*

Model	LRχ^2	AIC	BIC	Entropy	LMR-A (*p*)
2	-308.772	643.543	677.410	.915	154.360 (.063)
3	-272.547	585.094	637.198	.936	70.269 (.043)
4	-251.312	556.624	626.964	.914	41.192 (.016)
5	-248.758	565.517	654.092	.931	7.851 (.077)
6	-235.825	553.651	660.463	.961	45.886 (.240)

LRχ^2 = the likelihood ratio chi-square; AIC = Akaike information criterion; BIC = Bayesian information criterion; LMR-A = Lo-Mendell-Rubins adjusted likelihood ratio test.

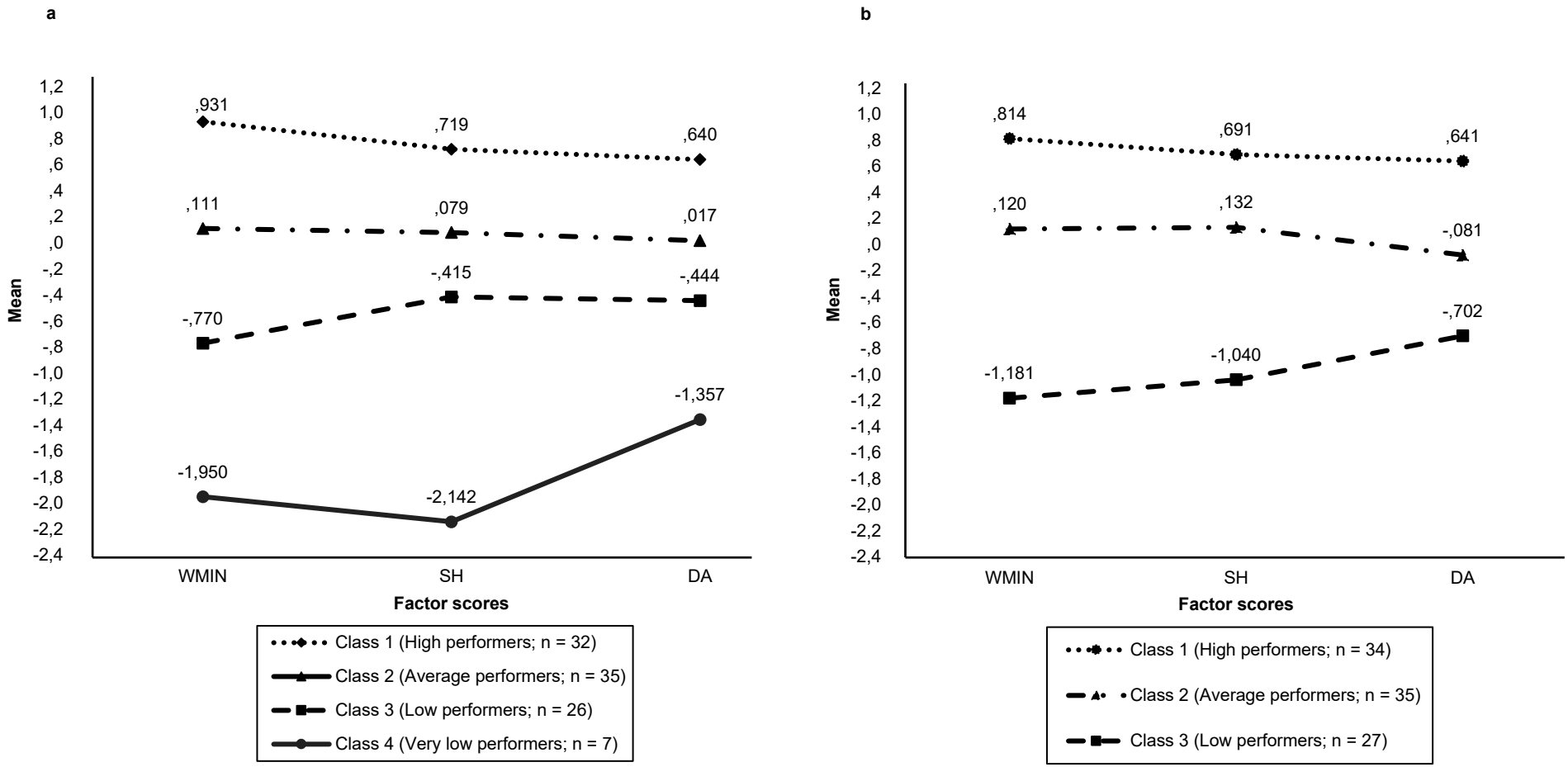

Fig. 1. *LPA classes and factor score means for children with ADHD (a) and for typically developing children (b).*