The Accuracy of Retrospective Recall of Childhood ADHD: Results from a Longitudinal Study

Elena von Wirth¹ · Janet Mandler¹ · Dieter Breuer¹ · Manfred Döpfner^{1,2}

Accepted: 2 November 2020 / Published online: 7 December 2020

© The Author(s) 2020

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a childhood-onset condition that may continue into adulthood. When assessing adult patients, clinicians usually rely on retrospective reports of childhood symptoms to evaluate the age-of-onset criterion. Since inaccurate symptom recall may impede the diagnosis and treatment of ADHD, knowledge about the factors influencing retrospective reports is needed. This longitudinal study investigated (a) the accuracy of retrospective symptom ratings by adult participants with a childhood diagnosis of ADHD (self-ratings) and parents or significant others (proxy ratings), and (b) the influence of current ADHD symptom severity and ADHD-associated impairments on retrospective symptom ratings. Participants ($N = 55$) were members of the Cologne Adaptive Multimodal Treatment (CAMT) study who had been referred and treated for ADHD in childhood and were reassessed in adulthood (average age 27 years). Participants' retrospective self-ratings were substantially lower than, and did not correlate with, parents' ADHD symptom ratings provided at study entry, while retrospective symptom ratings provided by proxy respondents correlated moderately with parents' childhood ratings. In addition, participants were more likely to underreport childhood symptoms (79%) and more frequently denied the presence of three or more childhood symptoms (17%) compared to proxy respondents (65% underreporting, 10% false-negative recall). Proxy respondents' symptom recall was best predicted by childhood ADHD, while participants' symptom recall was best predicted by current ADHD symptom severity. ADHD-associated impairments were not correlated with symptom recall after controlling for childhood ADHD. Together, these findings suggest a recall bias in adult patients and question the validity of retrospective reports, even in clinical samples.

Keywords Attention-deficit/hyperactivity disorder · Retrospective symptom recall · Age-of-onset criterion · Longitudinal study · Clinical sample · Cross-informant agreement

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common condition in children and adolescents (worldwide prevalence: 2.6–4.5%, Polanczyk, Salum, Sugaya, Caye and Rohde 2015). Over the last two decades, longitudinal studies have repeatedly shown that a substantial proportion of children with ADHD (41% - 77%) continue to struggle

with symptoms and functional impairments in adulthood (Faraone, Biederman and Mick 2006; Sibley, Swanson, Arnold, Hechtman, Owens, Stehli, et al. 2017; Uchida, Spencer, Faraone and Biederman 2018). Yet, despite increased awareness of ADHD as a life-long condition, adult ADHD is often underdiagnosed and undertreated (Kooij, Bijlenga, Salerno, Jaeschke, Bitter, Balázs, et al. 2019). To facilitate classification in adolescents and adults, the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association (APA) 2013) introduced some important revisions of the diagnostic criteria for ADHD (e.g., Epstein and Loren, 2013). For instance, the DSM-5 provides examples describing how ADHD symptoms may manifest across the lifespan. In addition, the symptom threshold was reduced to five symptoms (instead of the previous six) for individuals older than 17 years, and the age-of-onset criterion now requires several inattentive or hyperactive-impulsive symptoms to be present prior to age 12 years.

✉ Elena von Wirth
elena.von-wirth@uk-koeln.de

¹ Faculty of Medicine and University Hospital Cologne, School for Child and Adolescent Cognitive Behavior Therapy (AKiP), University of Cologne, Pohligstr. 9, 50969 Cologne, Germany

² Faculty of Medicine and University Hospital Cologne, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Cologne, Robert-Koch-Str. 10, 50931 Cologne, Germany

In the 4th edition of the DSM (DSM-IV-TR; APA 2000), the age-of-onset criterion specified that some symptoms and impairments needed to be present before age 7, but a large number of studies questioned the utility of this criterion (see Kielsing, Kielsing, Rohde, Frick, Moffitt, Nigg, et al. 2010). The DSM-IV field trials and other studies (e.g., Applegate, Lahey, Hart, Biederman and Hynd 1997; Barkley and Biederman, 1997; Todd, Huang and Henderson 2008) showed that many adolescents and adults do not accurately recall ADHD-related impairments that occurred before age 7. Similarly, Kessler, Berglund, Demler, Jin, Merikangas and Walters (2005) showed that only 50% of adults with clinical features of ADHD retrospectively recalled the presence of ADHD symptoms before age 7, while 95% recalled a symptom onset before age 12, suggesting that the decision to extend the age-of-onset criterion in the DSM-5 will improve the identification of adult ADHD. The recently published 11th edition of the International Classification of Diseases (ICD-11; World Health Organization [WHO] 2018) even went a step further by requiring a symptom onset “during the developmental period, typically early to mid-childhood”, instead of specifying an upper age limit.

The European Network Adult ADHD (ENAA) proposes that the diagnosis of adult ADHD should include the assessment of childhood-onset and current symptoms of ADHD using self-reports and collateral information from significant others (Kooij et al. 2019). However, problems with retrospective recall of ADHD symptoms remain. It has consistently been shown that informants disagree not only regarding the rating of current ADHD symptoms but also when asked to recall the severity and onset of ADHD symptoms in childhood (Breda, Rovaris, Vitola, Mota, Blaya-Rocha, Salgado, et al. 2016; Dias, Mattos, Coutinho, Segenreich, Saboya and Ayrão 2008; Henry, Moffitt, Caspi, Langley and Silva 1994; Kooij, Boonstra, Swinkels, Bekker, De Noord and Buitelaar 2008; Moffitt, Houts, Asherson, Belsky, Corcoran, Hammerle, et al. 2015). Moreover, longitudinal studies have revealed that retrospective reports of ADHD are often inaccurate. For instance, Barkley, Fischer, Smallish and Fletcher (2002) found that adults' retrospective self-reports of ADHD symptoms and parent ratings collected in childhood were only moderately correlated ($r = .39$ to $r = .44$). In a study by Loney, Ledolter, Kramer and Volpe (2007), participants rated their ADHD symptoms in adolescence (average age 14 years) and again in young adulthood (average age 21 years). The results revealed that participants' self-ratings in adolescence did not correlate significantly with their retrospective age-14 self-ratings made during the follow-up in young adulthood. In a longitudinal study by Mannuzza, Klein, Donald, Bessler and ShROUT (2002), adults participated in semi-structured psychiatric interviews designed to generate a lifetime diagnosis of ADHD. About one fifth (22%) of the participants did not receive a retrospective diagnosis of childhood ADHD according to DSM-III-R criteria

(APA 1987), although all participants had been diagnosed with ADHD in childhood. Similarly, a longitudinal study by Miller, Newcorn and Halperin (2010) examined the recall of childhood ADHD in a sample of adolescents and young adults with a childhood diagnosis of ADHD. The authors found that only 63% of participants and 78% of parents reported sufficient childhood symptoms to substantiate a past diagnosis of ADHD according to the DSM-IV (i.e., at least six childhood symptoms of inattention and/or hyperactivity-impulsivity prior to age 7). Together, these longitudinal studies demonstrate that retrospective reports of ADHD symptoms are, at most, moderately correlated with childhood ratings, and that even in samples with a confirmed childhood diagnosis of ADHD, a substantial proportion of adult participants (22–37%) falsely denies the presence of childhood ADHD symptoms.

Two large birth cohort studies confirmed that inaccurate reports of childhood ADHD are common, even if the age-of-onset criterion is extended to symptom onset before the age of 12 (as suggested by the DSM-5). Findings from the Dunedin Multidisciplinary Health and Development Study (Moffitt et al. 2015) revealed that only 23% of adults with a confirmed childhood diagnosis of ADHD had parents who correctly recalled that their offspring had core ADHD symptoms or had been diagnosed with ADHD before the age of 12 (true-positive recall). The parents of the remaining participants with confirmed childhood ADHD (77%) did not recall that their offspring had core ADHD symptoms before the age of 12 (false-negative recall). Only 4% of non-ADHD comparison subjects had parents who incorrectly recalled ADHD core symptoms during childhood (false-positive recall). In the 1993 Pelotas Birth Cohort study (Breda, Rohde, Menezes, Ansemi, Caye, Rovaris, et al. 2019), individuals with a current ADHD syndrome at the age of 22 (i.e. at least five ADHD symptoms in the last 6 months; symptoms in two or more settings; ‘a lot of’ or ‘very much’ impairment) were asked about the presence of symptoms before the age of 12. Only 33% of the individuals with confirmed childhood symptoms positively endorsed the presence of at least several ADHD symptoms in childhood (true-positive recall), and 67% falsely denied the presence of childhood symptoms (false-negative recall). Of the individuals without childhood ADHD, 30% positively endorsed the presence of childhood symptoms (false-positive recall), and 70% denied the presence of childhood symptoms (true-negative recall). Thus, the results of these birth cohort studies suggest that inaccurate recall of childhood ADHD is a common phenomenon and that false-negative recall was more frequently observed (67–77% of individuals with confirmed childhood ADHD) than false-positive recall (4–30% of individuals without childhood ADHD; Breda et al. 2019, Moffitt et al. 2015).

Since inaccurate symptom recall may lead to misdiagnosis and impede access to appropriate treatment options, knowledge about the factors influencing the accuracy of

retrospective ADHD symptom recall is needed in order to inform clinical decision makers about potential biases. To date, only a small number of studies have focused on this issue. Breda et al. (2019) investigated factors associated with true- and false-recall of ADHD symptoms in the above-mentioned population-based birth cohort sample. False-negative endorsement of childhood ADHD was more likely in male participants, non-white participants, and participants with less years of schooling. False-positive endorsement of childhood ADHD was more likely in individuals with social phobia and current ADHD symptoms. Similarly, the longitudinal study by Miller et al. (2010) found that the severity of current ADHD symptoms, but not the severity of childhood symptoms, was positively associated with the accuracy of retrospective symptom recall in a sample of adolescents and young adults with confirmed childhood ADHD (Miller et al. 2010).

The present study is one of the few to have followed children with a confirmed diagnosis of ADHD (age 6 to 10 years, $N = 55$) longitudinally into adulthood (average age 27 years). We collected parent ratings of ADHD symptoms in childhood as well as self- and proxy ratings of retrospective (childhood) and current (adult) ADHD symptoms in adulthood. The goal of the present analyses was to add to the limited body of knowledge on factors influencing the accuracy of ADHD symptom recall. We therefore aimed to determine how accurately adult participants and their parents or significant others recall childhood ADHD symptoms. In addition, we sought to determine whether retrospective symptom ratings are influenced by the severity of adult ADHD or ADHD-associated impairments.

We first analyzed changes in ADHD symptom severity over time. Based on the existing literature (e.g., Faraone et al. 2006), we expected that ratings of ADHD symptom severity would show a significant reduction from childhood to young adulthood. Next, we examined cross-informant agreement. Based on previous evidence, we expected to find low to moderate correlations between participants' and collateral informants' retrospective ratings of childhood symptoms (Breda et al. 2016; Dias et al. 2008; Henry et al. 1994; Murphy and Barkley 1996; Zucker, Morris, Ingram, Morris and Bakeman 2002). Similarly, we expected to find low to moderate correlations between self- and proxy ratings of current ADHD symptoms in adulthood (Kooij et al. 2008; Murphy and Barkley 1996; Mörsstedt, Corbisiero, Bitto and Stieglitz 2015).

We then examined the accuracy of ADHD symptom recall. Based on previous studies, we expected low to moderate correlations between ratings of ADHD symptom severity collected in childhood and retrospective ratings of childhood symptoms (Barkley et al. 2002; Loney et al. 2007). Given the substantial proportions of false-negative recall documented in previous studies, we further expected that both adult

participants and proxy respondents recall significantly fewer ADHD childhood symptoms than were originally reported by parents when participants were aged 6 to 10 years. However, since the present sample was clinically referred and treated for ADHD, we expected the proportion of participants and proxy respondents who false deny the presence of at least several ADHD childhood symptoms to be lower than in the Pelotas Birth Cohort study (67% false-recall; Breda et al. 2019) and the Dunedin Multidisciplinary Health and Development Study (77% false-recall; Moffitt et al. 2015). Subsequently, we examined whether the severity of ADHD symptoms and ADHD-associated impairments in adulthood influence the accuracy of retrospective symptom ratings. There is evidence that recall of past symptoms is fostered by emotional and physical distress (Barsky 2002). We therefore expected that participants with more severe ADHD symptoms and higher ADHD-associated impairments in adulthood would provide higher ratings of childhood ADHD.

Method

Participants

All participants were members of the Cologne Adaptive Multimodal Treatment (CAMT) study and had received individualized ADHD treatment consisting of behavior therapy and / or medication management in childhood. Inclusion criteria for the CAMT study were: (a) age 6 to 10 years; (b) attending the first, second, third or fourth school grade; (c) a nonverbal IQ of 80 or higher; and (d) fulfillment of DSM-III-R (APA 1987) or ICD-10 criteria (WHO 2004) for ADHD. ADHD symptoms were assessed through semi-structured interviews with parents and teachers (DISYPS-KJ; Döpfner and Lehmkuhl 2000). All children met diagnostic criteria for a DSM-III-R diagnosis of ADHD (including the presence of at least six ADHD symptoms) according to either parent or teacher interview. A detailed description of the initial study is provided by Döpfner, Breuer, Schürmann, Wolff Metternich, Rademacher and Lehmkuhl (2004). Follow-up [FU] assessments were conducted 18 months (Döpfner, Ise, Wolff Metternich-Kaizman, Schürmann, Rademacher and Breuer 2015), 8 years (Döpfner, Ise, Breuer, Rademacher, Wolff Metternich-Kaizman and Schürmann 2020) and 18 years after treatment (Döpfner, Mandler, Breuer, Dose, Walter and von Wirth 2020). Informed consent was obtained from parents or legal guardians prior to study inclusion in childhood and all participants and respondents gave their informed consent prior to data collection at the 18-year FU.

The present analyses are based on a subsample of $N = 55$ participants with available childhood ADHD ratings at study entry (pre-intervention) and at least one retrospective rating of childhood ADHD symptoms (self-rating or proxy rating)

assessed at the 18-year FU. At the 18-year FU, participants ($N = 55$) were 22 to 32 years of age ($M = 26.9$, $SD = 2.2$). The large age range at the 18-year FU resulted from the fact that initial study recruitment and the 18-year follow-up study each spread over several years. The majority were male ($n = 51$ male, $n = 4$ female).

Childhood Assessment (Pre-Intervention)

ADHD Rating Scale (FBB-ADHS) The FBB-ADHS is a proxy rating scale for ADHD symptoms (German: *Fremdbeurteilungsbogen ADHS*), and is part of the German ICD- and DSM-based Diagnostic System for the Assessment of Mental Disorders in Children and Adolescents (DISYPS, Döpfner, Görtz-Dorten, Lehmkuhl, Breuer and Goletz 2008). The first version of the FBB-ADHS, which was used at pre-intervention, contained 23 items assessing the occurrence of ADHD symptoms according to the diagnostic criteria of DSM-III-R and the preliminary research diagnostic criteria of ICD-10 rated on 4-point Likert scales ranging from 0 (not at all) to 3 (very much) (Döpfner and Lehmkuhl, 2000). There are two subscales: Inattention (11 items) and Hyperactivity-Impulsivity (12 items). Item scores are averaged to yield scale scores ranging from 0 to 3. Table S1 (Supplementary Material) provides an overview of the items. Parent ratings ($n = 55$) were provided by $n = 47$ mothers and $n = 6$ fathers (two missing values). Research has shown that the FBB-ADHS is a reliable (internally consistent; Cronbach's alpha .84 to .88) and valid instrument (Erhart, Döpfner, Ravens-Sieberger and the Bella Study Group 2008). Estimates of internal consistency (Cronbach's alpha) in the present sample were .89 for Inattention and .85 for Hyperactivity-Impulsivity.

18-Year Follow-Up (FU) Assessment

Rating Scale for Current ADHD Symptoms (FEA-AFB / FEA-ASB) The proxy rating scale for current ADHD symptoms in adults FEA-AFB (German: *Fragebogen zur Erfassung von ADHS im Erwachsenenalter – aktuelle Probleme Fremdbeurteilung*) and the self-rating scale for current ADHD symptoms in adults FEA-ASB (German: *Fragebogen zur Erfassung von ADHS im Erwachsenenalter – aktuelle Probleme Selbstbeurteilung*) (Döpfner, Lehmkuhl and Steinhausen, 2006) are adapted from the FBB-ADHS and the self-rating scale SBB-ADHS (German: *Selbstbeurteilungsbogen ADHS*), which are part of the DISYPS (Döpfner et al. 2008). Both rating scales were sent to participants by mail.

The FEA-AFB can be completed by parents or significant others (such as friends, partners or siblings), while the FEA-ASB assesses the patient's self-rating. Both scales comprise 20 items assessing the occurrence of ADHD symptoms according to the DSM-5 and ICD-10. There are three subscales:

Inattention (9 items), Hyperactivity-Impulsivity (11 items), and Impairment (10 items). The items of the Impairment scale assess functioning and psychological strain associated with ADHD in ten domains: 1. Performance in educational and occupational settings, 2. Relationships with teachers or supervisors, 3. Relationships with partners, parents, peers or colleagues, 4. During leisure time / recreational activities, 5. Community activities, 6. During learning / acquiring new contents, 7. Dating / Marriage, 8. Finances, 9. Driving, 10. Daily responsibilities. All items are rated on 4-point Likert scales ranging from 0 (not at all) to 3 (very much). Item scores are averaged to yield scale scores from 0 to 3. The total score is computed by averaging the item scores of the subscales Inattention and Hyperactivity-Impulsivity. In addition, the number of ADHD symptoms was determined (symptom count). Item scores ≥ 2 (2 = quite a bit, 3 = very much) are considered to reflect fulfillment of the specific symptom criterion (Inattention / Hyperactivity-Impulsivity scale) or substantial impairment (Impairment scale). In this study, proxy ratings (FEA-AFB, $n = 48$) were completed by $n = 30$ mothers, $n = 15$ partners, $n = 1$ grandparent, $n = 1$ friend, and $n = 1$ sibling. The FEA-AFB and FEA-ASB subscales were internally consistent in the present sample (Cronbach's alpha was .90 / .86 for Inattention, .86 / .85 for Hyperactivity-Impulsivity and .93 / .87 for Impairment, respectively).

Rating Scale for Retrospective Recall of ADHD Symptoms

(FEA-FFB / FEA-FSB) The proxy rating scale for retrospective recall of childhood ADHD symptoms FEA-FFB (German: *Fragebogen zur Erfassung von ADHS im Erwachsenenalter – frühere Probleme Fremdbeurteilung*) can be completed by parents and other informants who knew the patient during the developmental period. The self-rating scale for retrospective recall of childhood ADHD symptoms FEA-FSB (German: *Fragebogen zur Erfassung von ADHS im Erwachsenenalter – frühere Probleme Selbstbeurteilung*) assesses the patient's self-rating. Both rating scales were sent to participants by mail. Items of the FEA-FFB and FEA-FSB are analogous to those of the Inattention and Hyperactivity-Impulsivity scales of the FEA-AFB and FEA-ASB. Informants are asked to evaluate the behavior of the participant when he/she was aged 6 to 12 years. As in the FEA-AFB and FEA-ASB, there are nine items assessing inattentive behaviors and 11 items assessing hyperactive-impulsive behaviors. All items are rated on 4-point Likert scales ranging from 0 (not at all) to 3 (very much). Scale scores (0–3) are averaged across items. Additionally, the number of ADHD symptoms was determined (symptom count). In this study, proxy ratings (FEA-FFB, $n = 49$) were provided by $n = 45$ mothers and $n = 4$ fathers. FEA-FFB and FEA-FSB subscales were internally consistent in the present sample (Cronbach's alpha was .91 / .81 for Inattention, and .94 / .92 for Hyperactivity-Impulsivity, respectively).

Clinician's Rating of ADHD (DCL-ADHS-E) A semi-structured face-to-face interview assessing DSM-5 criteria for ADHD was conducted by a senior psychologist / psychotherapist (JM) at the university hospital Cologne to determine the prevalence rates of current ADHD diagnoses according to DSM-5 criteria in the 18-year FU sample. The *Diagnostic Checklist ADHD for Adults* (DCL-ADHS-E; German: *Diagnosecheckliste ADHS für Erwachsene*) was constructed for use in the present study. This is an adapted version of the *Diagnostic Checklist ADHD* (DCL-ADHS; German: *Diagnosecheckliste ADHS*), which is part of the DISYPS (Döpfner et al. 2008) and assesses diagnostic criteria for ADHD in children and adolescents. The DCL-ADHS is a reliable (internally consistent) instrument (Cronbach's alpha was .89 for symptoms of inattention, and .93 for symptoms of hyperactivity-impulsivity; Döpfner et al. 2008).

ADHD Inattentive subtype was diagnosed if at least five symptoms of inattention were present, but less than five symptoms of hyperactivity-impulsivity. ADHD Hyperactive-impulsive subtype was diagnosed if at least five symptoms of hyperactivity-impulsivity were present, but less than five symptoms of inattention. ADHD Combined subtype was diagnosed if at least five symptoms of inattention and at least five symptoms of hyperactivity-impulsivity were reported. ADHD in partial remission was diagnosed if less than five symptoms of inattention and / or less than five symptoms of hyperactivity-impulsivity were present. No ADHD diagnosis was given if no ADHD symptoms were reported. Consistent with the DSM-5, an adult diagnosis of ADHD also required substantial functional impairments. The interview did not assess ADHD symptoms in childhood, because (a) all participants had a clinical ADHD diagnosis at study entry and therefore all met the age of onset criterion of DSM-5, and (b) the accuracy of the retrospective recall of ADHD symptoms is an outcome of the present study.

Data Analysis

The statistical analyses were conducted using the Statistical Package for the Social Sciences, SPSS version 26 (IBM Corporation, Armonk, NY). If two or less items of the FBB-ADHS, FEA-FSB, or FEA-FFB were missing, the scale scores were computed by averaging the available item scores. For repeated measures analysis of variance (ANOVA), partial eta squared (η_p^2) was used as a measure of effect size. According to Cohen (1988), η_p^2 values below .06 are considered as small, values between .06 and .14 as moderate, and values above .14 as large. For paired samples *t*-tests, Cohen's *d* was used as a measure of effect size ($d = [M_2 - M_1] / SD_{pooled}$). According to Cohen (1988), *d* values of .20 are considered a small effect, .50 are considered a medium effect, and .80 are considered a large effect. A power analysis revealed that the present sample size of $n = 55$ would be

sufficient to detect effect sizes of $d = 0.4$ or larger with 80% power and a significance level of 0.05 (two-sided). For hierarchical regression analyses, f^2 was used as a measure of effect size ($f^2 = \sqrt{(\text{adjusted } R^2 / (1 - \text{adjusted } R^2))}$). According to Cohen (1992), f^2 values of .02 are considered a small effect, .15 are considered a medium effect, and .35 are considered a large effect. The Benjamin-Hochberg procedure was applied to *p* values retrieved from ANOVAs, *t*-tests, and regression analyses to control for the false discovery rate (FDR). In case of non-significant results for paired samples *t*-tests (i.e. $p > .05$), we tested for equivalence using the "two one-sided tests" (TOST) procedure (Lakens 2017) with equivalence bounds of ± 0.3 scale points.

To investigate whether ADHD symptom severity declines over time, we first compared parent ratings of childhood ADHD symptom severity provided at study entry (FBB-ADHS) to proxy ratings of current ADHD symptom severity provided at the 18-year FU (FEA-AFB) using a repeated measures ANOVA with the dependent variable ADHD symptom severity (proxy rating) and the two within-subjects variables Time (pre-intervention / 18-year FU) and Scale (Inattention / Hyperactivity-Impulsivity). We then analyzed whether changes in ADHD symptom severity over time are evident when participants' self-ratings of current ADHD symptoms in adulthood (FEA-ASB) are compared to parent ratings of childhood ADHD symptoms at study entry (FBB-ADHS) using a repeated measures ANOVA with the dependent variable ADHD symptom severity and the two within-subjects variables Time (pre-intervention / 18-year FU) and Scale (Inattention / Hyperactivity-Impulsivity).

We next sought to examine cross-informant agreement for ADHD symptom ratings. Due to the children's young age at study entry (6 to 10 years), we did not collect self-ratings at pre-intervention. We therefore used data collected at the 18-year FU to estimate the degree of cross-informant agreement. Paired samples *t*-tests were conducted to compare (i) proxy respondents' retrospective ratings of childhood ADHD symptoms (FEA-FFB) to participants' retrospective self-reports of childhood ADHD symptoms (FEA-FSB), and (ii) proxy respondents' ratings of current ADHD symptoms (FEA-AFB) to participants' self-reports of current ADHD symptoms (FEA-ASB). In addition, we present Pearson correlations (two-tailed) between ADHD symptom ratings of different raters and different time points.

To investigate the accuracy of retrospective recall of ADHD childhood symptoms, we first conducted paired samples *t*-tests that compared retrospective self-ratings of childhood symptom severity (FEA-FSB) and parent ratings collected in childhood (FBB-ADHS), as well as retrospective proxy ratings of childhood symptom severity (FEA-FFB) and parent ratings collected in childhood (FBB-ADHS). We also considered Pearson correlations (two-tailed) between parent ratings collected in childhood (FBB-ADHS) and retrospective ratings

of childhood symptom severity. To determine the rate of cases who fail to confirm the age-of-onset criterion of the DSM-5 by denying the presence of at least several childhood symptoms (defined as the presence of at least three ADHD symptoms between the ages of 6 and 12 years), we determined the number of childhood ADHD symptoms recalled by adult participants and their parents or significant others (proxy respondents). In addition, we conducted paired samples *t*-tests with number of symptoms as the dependent variable to compare the number of ADHD childhood symptoms reported at the 18-year FU (FEA-FSB / FEA-FFB) and at pre-intervention (childhood measure, FBB-ADHS). To determine the number of participants who over- or underreported ADHD childhood symptoms, we calculated difference scores between the number of ADHD childhood symptoms reported by parents at study entry and the number of ADHD childhood symptoms recalled by participants at the 18-year FU (DIFF-SELF) or the number of ADHD childhood symptoms recalled by proxy informants at the 18-year FU (DIFF-PROXY). Negative values indicate underreporting (i.e. less childhood symptoms recalled than originally reported by parents at study entry). Positive values indicate overreporting (i.e. more childhood symptoms recalled than originally reported by parents at study entry).

To investigate whether the recall accuracy of childhood ADHD symptoms is influenced by the severity of current ADHD symptoms or impairments associated with adult ADHD, two hierarchical regression analyses were performed. In the first regression analysis, participants' retrospective ratings of childhood ADHD symptoms (FEA-FSB, self-report) served as the outcome variable. The control variable (childhood ADHD assessed at pre-intervention [FBB-ADHS, parent rating]) was entered into the model as a potential predictor in the first step. Self-ratings of adult ADHD symptoms (FEA-ASB total score) and ADHD-associated impairments (FEA-ASB impairment scale) were added as potential predictors in the second step using the forced entry (ENTER) method. The second regression analysis was conducted with retrospective ratings of childhood ADHD provided by parents and significant others (FEA-FFB, proxy report) as the outcome variable. Again, predictors were entered hierarchically (Step 1: FBB-ADHS; Step 2: FEA-AFB total score, FEA-AFB impairment scale) using the forced entry (ENTER) method.

Results

Prevalence of ADHD Diagnoses in Adulthood

Forty-four participants were willing to participate in the clinical interview DCL-ADHS-E at the 18-year FU. Eight participants (18.2%) qualified for a diagnosis of ADHD Inattentive subtype, two participants (4.5%) qualified for a diagnosis of

ADHD Hyperactive-impulsive subtype, and another two participants (4.5%) qualified for a diagnosis of ADHD Combined subtype. Approximately half of the participants ($n = 24$; 54.5%) qualified for a diagnosis of ADHD in partial remission. Eight participants (18.2%) did not qualify for any ADHD diagnosis in adulthood.

ADHD Symptom Reduction from Childhood to Adulthood

Table 1 displays the severity of ADHD symptoms (M and SD of ADHD scale scores) reported in childhood (pre-intervention) and adulthood (18-year FU). The repeated measures ANOVA with the dependent variable ADHD symptom severity (proxy rating) revealed a large and significant effect of Time, $F_{(1,47)} = 124.03$, $p = .003$, $\eta_p^2 = .73$. Parent ratings provided in childhood were significantly higher than proxy ratings of adult ADHD symptom severity completed at the 18-year FU, suggesting a substantial decline of ADHD symptoms from childhood to young adulthood. There was no significant effect of Scale, $F_{(1,47)} = 2.60$, $p = .141$, $\eta_p^2 = .05$, but the Time*Scale interaction reached statistical significance, $F_{(1,47)} = 10.40$, $p = .005$, $\eta_p^2 = .18$. Ratings of hyperactivity-impulsivity showed a greater decline over time compared to ratings of inattention. Since ratings of adult ADHD symptom severity provided by parents tended to be lower than those provided by other informants (Inattention: $d = 0.41$; Hyperactivity-Impulsivity: $d = 0.44$), we infer that the observed decline in symptom severity cannot be attributed to the fact that a substantial proportion (37.5%) of proxy ratings of adult ADHD was provided by informants other than parents.

We next analyzed whether the decline of ADHD symptoms over time is also evident when participants' self-ratings of current ADHD symptoms in adulthood (FEA-ASB) are compared to parent ratings of childhood ADHD symptoms at study entry (FBB-ADHS). Again, repeated measures ANOVA revealed a large and significant decline of ADHD symptom severity over time, $F_{(1,53)} = 168.55$, $p = .003$, $\eta_p^2 = .76$. There was no significant effect of Scale, $F_{(1,53)} = 0.10$, $p = .749$, $\eta_p^2 = .002$, and no significant Time*Scale interaction, $F_{(1,53)} = 3.44$, $p = .094$, $\eta_p^2 = .06$. Given the medium effect size of the interaction, the lack of a statistically significant interaction effect might be due to the restricted sample size and associated limited statistical power.

Cross-Informant Agreement

Proxy respondents' retrospective ratings of childhood ADHD symptoms (FEA-FFB) were significantly higher than participants' retrospective self-reports of childhood ADHD symptoms (FEA-FSB) (Inattention: $t_{(46)} = 3.51$, $p = .003$, $d = 0.60$; Hyperactivity-Impulsivity: $t_{(46)} = 3.10$, $p = .007$, $d = 0.47$). In

Table 1 Descriptive statistics (*M*, *SD*) for ADHD rating scales

	ADHD total	Inattention	Hyperactivity-Impulsivity
self-report			
Retrospective rating of childhood ADHD symptoms (FEA-FSB, <i>n</i> = 53) ^a	1.35 (0.63)	1.45 (0.59)	1.27 (0.76)
Rating of adult ADHD symptoms (FEA-ASB, <i>n</i> = 54) ^a	0.62 (0.46)	0.67 (0.50)	0.58 (0.50)
Informant report			
Rating of childhood ADHD symptoms (parent rating, FBB-ADHS, <i>n</i> = 55) ^b	1.86 (0.55)	1.84 (0.61)	1.89 (0.58)
Retrospective rating of childhood ADHD symptoms (parent rating, FEA-FFB, <i>n</i> = 49) ^{a, c}	1.76 (0.71)	1.88 (0.67)	1.67 (0.81)
Rating of adult ADHD symptoms (proxy rating, FEA-AFB, <i>n</i> = 48) ^{a, d}	0.76 (0.58)	0.87 (0.68)	0.67 (0.58)

Note. ^a Symptom rating assessed in adulthood (18-year FU). ^b Symptom rating assessed in childhood (pre-intervention). ^c Informants were *n* = 45 mothers and *n* = 4 fathers. ^d Informants were *n* = 30 mothers, *n* = 15 partners, *n* = 1 grandparent, *n* = 1 friend, and *n* = 1 sibling

addition, external informants provided higher ratings of current adult symptoms of Inattention (FEA-AFB) than did participants themselves (FEA-ASB), $t_{(46)} = 2.36$, $p = .041$, $d = 0.35$. Proxy ratings (FEA-AFB) and self-ratings (FEA-ASB) did not differ significantly regarding the severity of current adult symptoms of Hyperactivity-Impulsivity, $t_{(47)} = 1.26$, $p = .254$, $d = 0.15$. The TOST procedure indicated that the observed effect size was significantly within the equivalence bounds of -0.3 and 0.3 scale points, $t_{(47)} = 3.47$, $p = .001$. We therefore conclude that the observed effect is close enough to zero to be practically equivalent. According to Cohen (1988), the correlation between self- and proxy ratings of ADHD symptoms was moderate for retrospective ratings of childhood ADHD (see Table 2; FEA-FFB / FEA-FSB: $r = .44$, $p = .002$) and somewhat higher for ratings of current ADHD in adulthood (FEA-AFB / FEA-ASB: $r = .63$, $p < .001$), but the correlation coefficients did not differ significantly ($z = -1.33$, $p = .18$; Dunn and Clark, 1969).

Accuracy of Retrospective Recall of ADHD Childhood Symptoms

Retrospective self-ratings of childhood symptom severity (FEA-FSB) were significantly lower than parent ratings collected in childhood (FBB-ADHS) for both the Inattention Scale, $t_{(52)} = 3.54$, $p = .003$, $d = 0.66$, and the Hyperactivity-Impulsivity Scale, $t_{(52)} = 5.66$, $p = .003$, $d = 0.91$. In contrast, retrospective proxy ratings of childhood symptom severity (FEA-FFB) did not differ significantly from parent ratings collected in childhood (FBB-ADHS) (Inattention: $t_{(48)} = 0.42$, $p = .705$, $d = 0.07$; Hyperactivity-Impulsivity: $t_{(48)} = 1.95$, $p = .089$, $d = 0.31$). The TOST procedure indicated that the observed effect size for Inattention was significantly within the equivalence bounds of -0.3 and 0.3 scale points, $t_{(48)} = 2.44$, $p = .009$. The observed effect size for Hyperactivity-Impulsivity was not significantly within these bounds, $t_{(48)} = 0.72$, $p = .239$, suggesting that the present study may not have sufficient statistical power to detect an effect large enough to be considered meaningful.

As shown in Table 2, parent ratings of ADHD symptom severity collected in childhood (FBB-ADHS) and participants' retrospective self-ratings of childhood symptom severity (FEA-FSB) did not correlate significantly ($r = .24$). The correlation between parent ratings in childhood (FBB-ADHS) and retrospective proxy ratings (FEA-FFB) was moderate ($r = .39$, $p = .006$).

All participants fulfilled diagnostic criteria for a DSM-III-R diagnosis of ADHD at study entry, which required the presence of at least six ADHD symptoms in childhood. Table 3 shows the number and percentage of respondents that (i) falsely deny the presence of at least several ADHD childhood symptoms (0–2 childhood symptoms reported), (ii) report the presence of 3 to 5 ADHD childhood symptoms, or (iii) report the presence of six or more ADHD childhood symptoms, as well as the mean number of ADHD childhood symptoms reported. When the participants were aged 6 to 10 years (pre-intervention), the vast majority of parents (95.6%) indicated the presence of six or more ADHD symptoms on the FBB-ADHS. None of the parents reported two or less symptoms on the FBB-ADHS. When the participants were aged 22 to 32 years (18-year FU), 75.5% of the proxy respondents indicated the presence of six or more childhood symptoms on the FEA-FFB. Five proxy respondents (10.2%) recalled two or less childhood symptoms, and hence falsely denied the presence of at least several childhood symptoms. Participants' retrospective self-ratings of childhood symptoms at the 18-year FU (FEA-FSB) revealed that 17.0% of the sample falsely denied the presence of several childhood symptoms. Both participants and proxy respondents reported significantly fewer ADHD childhood symptoms at the 18-year FU (FEA-FSB / FEA-FFB) than did the parents at pre-intervention (childhood measure, FBB-ADHS), $t_{(52)} = 6.73$, $p = .003$, $d = 1.16$, and $t_{(48)} = 3.28$, $p = .005$, $d = 0.51$, respectively.

The mean difference score DIFF-SELF was $M = -6.28$ ($SD = 6.80$), providing further support for the finding that

Table 2 Correlations between childhood ADHD symptom severity, retrospective recall of childhood ADHD symptom severity, and adult ADHD symptom severity ($n = 47\text{--}53$)

	Childhood ADHD ^a (parent rating)	Childhood ADHD ^b (retrospective self-rating)	Adult ADHD ^c (self-rating)	Childhood ADHD ^d (retrospective parent rating)	Adult ADHD ^e (proxy rating)
Childhood ADHD ^a (parent rating)		$r = 0.24$	$r = 0.07$	$r = 0.39^{**}$	$r = 0.29^*$
Childhood ADHD ^b (retrospective self-rating)			$r = 0.52^{**}$	$r = 0.44^{**}$	$r = 0.40^{**}$
Adult ADHD ^c (self-rating)				$r = 0.18$	$r = 0.63^{**}$
Childhood ADHD ^d (retrospective parent rating)					$r = 0.32^*$

Note. ^a ADHD parent rating scale (FBB-ADHS) at pre-intervention. ^b Self-rating scale for retrospective recall of childhood ADHD symptoms (FEA-FSB) at 18-year FU. ^c Self-rating scale for current ADHD symptoms (FEA-ASB) at 18-year FU. ^d Parent rating scale for retrospective recall of childhood ADHD symptoms (FEA-FFB) at 18-year FU. ^e Proxy rating scale for current ADHD symptoms (FEA-AFB) at 18-year FU. * $p < .05$. ** $p < .01$

participants recalled, on average, less childhood symptoms than their parents had reported at study-entry. The mean difference score DIFF-PROXY was also negative ($M = -2.98$, $SD = 6.37$) confirming that proxy informants recalled, on average, less childhood symptoms than the parents had reported at study-entry. Retrospective underreporting of childhood symptoms was found in $n = 42$ (79.2%) of participants and $n = 32$ (65.3%) of proxy informants. Retrospective overreporting of childhood symptoms was found in $n = 10$ (18.9%) of participants and $n = 15$ (30.6%) of proxy informants. Figure 1 depicts the distribution of DIFF-SELF and DIFF-PROXY scores in the sample. Descriptive item-level analyses consisting of mean, stand deviation, and percentage of sample with item score ≥ 2 (reflecting fulfillment of the specific symptom criterion) of the FBB-ADHD, the FEA-FFB, and the FEA-FSB are provided in Table S1 – S3 (Supplementary Material).

Does Adult ADHD or Impairment Associated with Adult ADHD Influence the Retrospective Recall of Childhood Symptoms?

Table 4 shows the results of the regression analyses. Parent ratings of ADHD in childhood were not a significant predictor of young adults' retrospective recall of their childhood ADHD in Step 1 ($R^2 = .05$, $\beta = .23$, $p = .139$). Adding ratings of adult ADHD and impairments as additional predictors significantly improved the model fit ($\Delta R^2 = .27$, $p = .003$). Participants' self-ratings of adult ADHD were a significant predictor of their retrospective symptom recall ($\beta = .41$, $p = .047$), but self-ratings of ADHD-associated impairments ($\beta = .13$, $p = .549$) were not. The model explained a significant proportion of the variance (32%, adjusted $R^2 = .27$, $f^2 = 0.61$) in participants' retrospective recall of childhood ADHD, $F_{(3,45)} = 7.05$, $p = .003$, which can be considered a large effect (Cohen 1992).

Table 3 Number of ADHD childhood symptoms reported by participants (self-rating) and their parents

	Parent rating in childhood (FBB-ADHS) ^a $n = 55$	Retrospective parent rating (FEA-FFB) ^b $n = 49$	Retrospective self-rating (FEA-FSB) ^c $n = 53$
0–2 ADHD symptoms	$n = 0$ 0.0%	$n = 5$ 10.2%	$n = 9$ 17.0%
3–5 ADHD symptoms	$n = 3$ 5.5%	$n = 7$ 14.3%	$n = 8$ 15.1%
≥ 6 ADHD symptoms	$n = 52$ 94.6%	$n = 37$ 75.5%	$n = 36$ 67.9%
Number of ADHD symptoms	$M = 15.00$ $SD = 5.06$	$M = 12.10$ $SD = 6.78$	$M = 8.70$ $SD = 5.69$

Note. ^a Parent rating scale FBB-ADHS completed at study entry (pre-intervention). ^b Parent rating scale for retrospective recall of childhood ADHD symptoms (FEA-FFB) completed at 18-year FU. ^c Self-rating scale for retrospective recall of childhood ADHD symptoms (FEA-FSB) completed at 18-year FU

Note. DIFF-SELF = (nr. of ADHD childhood symptoms recalled at the 18-year FU [FEA-FSB, self-report] – nr. of ADHD childhood symptoms reported at study-entry [FBB-ADHS, parent rating]); DIFF-PROXY = (nr. of ADHD childhood symptoms recalled at the 18-year FU [FEA-FFB, proxy report] – nr. of ADHD childhood symptoms reported at study-entry [FBB-ADHS, parent rating]). Negative values indicate underreporting of childhood symptoms at 18-year FU; positive values indicate overreporting of childhood symptoms at 18-year FU.

Fig. 1 Distribution of DIFF-SELF and DIFF-PROXY scores

In contrast, parent ratings of ADHD in childhood significantly predicted proxy respondents retrospective ratings of childhood ADHD symptoms ($R^2 = .19$, $\beta = .44$, $p = .008$). Entering ratings of adult ADHD and impairments as additional predictors did not significantly improve the model fit ($\Delta R^2 = .11$, $p = .094$). Neither proxy ratings of adult ADHD ($\beta = -.10$, $p = .693$), nor proxy ratings of ADHD-associated impairments in adulthood ($\beta = .40$, $p = .089$) were significant predictors after controlling for childhood ADHD. The overall fit of the second regression model was statistically significant, $F_{(3,37)} = 5.35$, $p = .008$. The predictors explained 30% of the variance in retrospective ratings of childhood ADHD symptoms by proxy respondents (adjusted $R^2 = .25$, $f^2 = 0.58$), which can be considered a large effect (Cohen 1992).

Discussion

The diagnostic criteria for ADHD according to the DSM-5 (APA 2013) require the presence of at least several symptoms prior to age 12. To evaluate this criterion in adult patients, clinicians often have to rely on retrospective recall of ADHD symptoms by patients and significant others who knew the patients during the developmental period, usually one or both parents. However, it has repeatedly been shown that a substantial proportion of adolescents and adults do not accurately recall their childhood functioning (e.g., Applegate et al. 1997; Kessler, Berglund et al. 2005). The aim of the present study was therefore to investigate (a) the accuracy of retrospective ADHD childhood ratings provided by adult participants and their parents or significant others, and b) the influence of ADHD symptom severity and ADHD-associated impairments in adulthood on retrospective

ADHD symptom ratings in order to inform clinical decision makers about potential biases.

Participants ($N = 55$) were members of the CAMT study who had been diagnosed and treated for ADHD during childhood and were reassessed in adulthood when they were 22 to 32 years of age (18-year FU). We first analyzed changes in ADHD symptom severity over time. As we expected based on previous evidence (e.g. Faraone et al. 2006), both self- and proxy ratings of adult ADHD were substantially and significantly lower than parent ratings of ADHD collected in childhood ($\eta_p^2 = .73$ and $\eta_p^2 = .76$), suggesting that ADHD symptoms lessened from childhood to adulthood. Also consistent with previous evidence (e.g. Döpfner, Hautmann, et al. 2015), proxy ratings of hyperactivity-impulsivity showed a substantially greater decline over time compared to ratings of inattention ($\eta_p^2 = .18$). Our analyses further suggest that this observed decline cannot be attributed to the fact that a substantial proportion (37.5%) of proxy ratings of adult ADHD was provided by informants other than parents. We therefore conclude that our data confirm previous findings of a developmental reduction in ADHD symptomatology (e.g., Langberg, Epstein, Altaye, Molina, Arnold and Vitiello 2008), although we cannot completely rule out that other factors (e.g., changes in item formulations, altered perception of ADHD-related behaviors) influenced our results.

We then examined cross-informant agreement for ADHD symptom ratings. Previous studies found that adults report more ADHD childhood symptoms than do their parents or other informants (Magnússon, Smári, Baldursson, Kristjánsson, Sigurbjörnsdóttir and Guomundsson 2006; Murphy and Schachar 2000), but that they tend to underreport current ADHD symptoms (Kooij et al. 2008; Zucker et al. 2002). The results of the present study differ from these findings in an important aspect: Adult participants' retrospective

Table 4 Results of hierarchical regression analyses predicting retrospective ADHD symptom ratings from childhood ADHD symptom ratings and adult ADHD symptom ratings

	Retrospective rating of childhood ADHD symptoms (self-rating, FEA-FSB) ^a				Retrospective rating of childhood ADHD symptoms (parent rating, FEA-FFB) ^a			
	<i>B</i>	<i>SE B</i>	β	ΔR^2	<i>B</i>	<i>SE B</i>	β	ΔR^2
Step 1				.05				.19**
Constant	0.88	0.31			0.68	0.38		
Rating of childhood ADHD symptoms (parent rating, FBB-ADHS) ^b	0.26	0.16	.23		0.58	0.19	.44**	
Step 2				.27**				.11
Constant	0.50	0.29			0.63	0–36		
Rating of childhood ADHD symptoms (parent rating, FBB-ADHS) ^b	0.22	0.14	.20		0.52	0.19	.39*	
Rating of adult ADHD symptoms (self-report, FEA-ASB) ^a	0.57	0.25	.41*					
ADHD associated impairment (self-report, FEA-ASB)	0.18	0.25	.13					
Rating of adult ADHD symptoms (proxy report, FEA-AFB) ^a					- 0.12	0.26	- .10	
ADHD associated impairment (proxy report, FEA-AFB)					0.41	0.21	.40	

Note. ΔR^2 = Change in R^2 associated with the variable(s) entered in each step. ^a Symptom rating assessed in adulthood (18-year FU). ^b Symptom rating assessed in childhood (pre-intervention)

* $p < .05$. ** $p < .01$

ratings of their childhood symptoms were substantially and significantly *lower* than proxy respondents' retrospective ratings ($d = 0.60$), and their ratings of current symptoms of Hyperactivity-Impulsivity differed only slightly and not statistically significant from those provided by other informants ($d = 0.15$). Consistent with previous findings, adult participants underreported current symptoms of Inattention compared to proxy respondents, although the size of this effect can be considered small ($d = 0.35$). Also consistent with previous evidence (Kooij et al. 2008; Mörtstedt et al. 2015; Murphy and Barkley, 1996; Zucker et al. 2002), there was a moderate correlation ($r = .44$) between participants' and proxy respondents' retrospective ratings of childhood symptoms and a somewhat (but not significantly) higher correlation ($r = .63$) between participants' and other informants' ratings of current symptoms in adulthood. Together, these analyses suggest that, in our sample, cross-informant agreement was moderate for retrospective ratings of childhood ADHD and high for ratings of current adult ADHD.

As an index of recall accuracy, we first compared retrospective symptom ratings with ratings of ADHD symptom severity collected in childhood (parent rating). Participants retrospectively described their childhood symptoms of Inattention and Hyperactivity-Impulsivity as substantially less severe than their parents had done when the participants were children ($d = 0.66$ and $d = 0.91$). In contrast, retrospective ratings of Inattention provided by parents and other proxy respondents differed only slightly and not statistically significant from parents' childhood ratings ($d = 0.07$). Their retrospective ratings of Hyperactivity-Impulsivity also differed only slightly and not statistically from parents' childhood ratings

($d = 0.31$), but equivalence testing suggested that our design may lack sufficient statistical power to detect a meaningful effect. In line with previous findings (Barkley et al. 2002; Miller et al. 2010; Loney et al. 2007), participants' retrospective symptom ratings did not correlate statistically significantly with parent ratings provided in childhood ($r = .24$, ns), while proxy respondents' retrospective symptom ratings correlated moderately with parents' previous ADHD ratings ($r = .39$). Interestingly, as noted above, adult participants' retrospective symptom ratings correlated moderately with proxy respondents' retrospective ratings ($p = .44$), suggesting that adult patients' symptom recall is more closely associated with how their parents or significant others recall and communicate about childhood symptoms than with how parents judged the severity of ADHD symptoms in childhood.

As another index of recall accuracy, we determined how many informants falsely denied the presence of at least several ADHD symptoms in childhood. Consistent with previous studies (e.g., Mannuzza et al. 2002; Miller et al. 2010), adult participants and proxy respondents recalled significantly fewer childhood symptoms than the parents had reported at study entry. This effect can be considered large for participants ($d = 1.15$) and moderate for proxy respondents ($d = 0.51$). However, the majority of the adult participants and proxy respondents (68% and 76%, respectively) reported sufficient childhood symptoms to substantiate a past diagnosis of ADHD (≥ 6 ADHD childhood symptoms). Even more participants and proxy respondents (83% and 90%, respectively) accurately recalled the presence of at least several ADHD childhood symptoms (defined as the presence of at least three ADHD symptoms between the ages of 6 and 12 years). Yet,

17% of participants and 10% of proxy respondents falsely denied the presence of at least three childhood symptoms, even though all participants had received multimodal treatment for their childhood ADHD including six sessions of psychoeducation in all treatment arms (Döpfner et al. 2004). As we expected, recall accuracy in nonclinical samples has been found to be even lower. For example, results from the Dunedin Multidisciplinary Health and Development Study revealed that only 23% of parents of adults with a confirmed childhood diagnosis of ADHD recalled that their offspring had core symptoms or had been diagnosed with ADHD before the age of 12 (Moffitt et al. 2015). Similarly, results from the 1993 Pelotas Birth Cohort study showed that only 33% of adult participants with childhood ADHD recalled the presence of childhood symptoms (Breda et al. 2019). We therefore conclude that having received clinical diagnosis and treatment during childhood improves the accuracy of retrospective symptom recall.

In addition, we determined the number of participants who over- or underreported ADHD childhood symptoms. Underreporting of childhood symptoms was a common phenomenon. The majority of participants (79%) and proxy respondents (65%) recalled less childhood symptoms than parents had reported at study entry. Our data do not confirm previous reports of symptom exaggeration in 25–48% of college students self-referred for ADHD evaluation (Sullivan, May and Galbally 2007), possible because the present sample was not faced with secondary gain potentials such as academic accommodation or other forms of assistance.

To sum up, we found that proxy respondents recalled significantly fewer ADHD childhood symptoms than parents had reported at study entry, but when severity ratings were considered, proxy respondents' retrospective ratings correlated moderately with, and did not differ significantly from, parents' previous ratings. In contrast, adult participants' retrospective self-ratings were substantially lower than, and did not correlate with, their parents' ratings at study entry. Adult participants were also more likely to underreport ADHD childhood symptoms (79%) and to falsely deny the presence of at least several ADHD childhood symptoms (17%) compared to proxy respondents (65% underreporting; 10% false-negative recall). These results might suggest that participants have limited retrospective recall of childhood symptoms. However, we did not collect self-ratings in childhood, and hence compared participants' retrospective ratings to parent ratings collected in childhood. Since cross-informant agreement on retrospective ADHD ratings has repeatedly been shown to be at most modest (see above), a comparably lower level of accuracy for retrospective self-ratings in our data was to be expected.

Finally, we were interested in finding out whether the severity of ADHD symptoms and ADHD-associated impairments in adulthood influences retrospective symptom ratings. Our results only partly support our expectation that participants with

more severe ADHD symptoms and higher ADHD-associated impairments in adulthood would provide higher ratings of childhood ADHD. Hierarchical regression analyses revealed that adult participants' retrospective symptom ratings were significantly predicted by their ratings of current ADHD symptoms, but not by childhood ADHD ratings or ratings of ADHD-associated impairments in adulthood. Proxy respondents' retrospective ratings were best predicted by childhood ADHD ratings. Neither ratings of current ADHD nor ratings of ADHD-associated impairments in adulthood were significant predictors of their retrospective symptom ratings after controlling for childhood ratings of ADHD. To summarize, we found that adult participants' symptom recall was influenced by the severity of current ADHD symptoms. More severe symptoms in adulthood predicted higher retrospective symptom ratings ($\beta = .41$). Participants' symptom recall was not influenced by the severity of childhood ADHD or ADHD-associated impairments in adulthood. Proxy respondents' retrospective ratings were associated with the severity of childhood ADHD, but they were not influenced by the severity of current ADHD or ADHD-associated impairments. These findings suggest a recall bias in adult patients, but not in external informants.

Several limitations of the present study need to be mentioned. First, our sample size was restricted, which limited the power to detect small effects and thus limited the interpretation of our results. Second, since our sample was referred and treated for ADHD in childhood, our data can provide information on the percentage of participants that falsely denied the presence of childhood symptoms, but it is not suitable for detecting false positive or true negative reports of childhood ADHD. In addition, our results cannot be generalized to samples not treated for ADHD. Symptom recall might be less accurate in adults who did not undergo diagnostic assessment and treatment during childhood, since the severity of childhood symptoms has been found to be associated with recall accuracy (Kessler, Adler, Barkley, Conners, Faraone, Greenhill, et al. 2005). Moreover, adults who were being told about their ADHD symptoms by a professional during childhood might be more likely to remember having had these symptoms. The generalizability of results is further limited by the fact that 51 of 55 participants (93%) were male. Future studies should recruit both clinically referred and population-based samples of children with and without ADHD that include both male and female participants to study the effects of gender and prior diagnosis and treatment on the accuracy of retrospective ADHD symptom recall. Another limitation worth mentioning is the difference between the FBB-ADHS and the FEA-FFB / FEA-FSB. The first version of the FBB-ADHS, which was used at pre-intervention, contained 23 items assessing the occurrence of ADHD symptoms according to the diagnostic criteria of DSM-III-R and the preliminary research diagnostic criteria of ICD-10. The FEA-FSB and the FEA-FFB each contain 20 items assessing the

occurrence of ADHD symptoms according to the DSM-5 and ICD-10. We were, therefore, not able to perform a direct comparison of the specific symptoms that tended to be under- or overreported at the 18-year FU.

Despite these limitations, our study adds valuable information about the accuracy of ADHD symptom recall. Our results question the validity of retrospective recall of childhood ADHD symptoms, since retrospective symptom ratings showed only low to moderate correlations ($p = .24 - p = .39$) with ratings of childhood symptoms assessed at ages 6 to 10 years. Moreover, 17% of adults with a childhood diagnosis of ADHD and 10% of proxy respondents falsely denied the presence of at least three childhood symptoms. We further demonstrated that participants' retrospective symptom recall was associated with the severity of current ADHD symptoms. This finding fits well with the concept of a state-dependent recall, which postulates that recall of past symptoms and episodes is affected by a patient's emotional and physical state at the time of recall (Barsky 2002). The relationship between current distress and recall of past emotions or symptoms has previously been supported by evidence of a negative recall bias in patients with depression (Ben-Zeev and Young, 2010), panic disorders (Margraf, Taylor, Ehlers, Roth and Agras 1987), borderline personality disorder (Ebner-Priemer, Kuo, Welch, Thielgen, Witte, Bohus and Linehan 2006), or acute pain (Eich, Reeves, Jaeger and Graff-Radford 1985).

A possible clinical implication of our findings is that adult patients who report low levels of current ADHD symptomatology might underestimate their childhood symptoms. Clinicians should be vigilant for underreporting of childhood symptoms among patients with ADHD symptoms who do not regard their behavior as ADHD-related, and may try to increase recall accuracy by asking the patient to describe specific situations and / or by providing examples of specific symptom behaviors. These examples may serve as retrieval cues that remind the patient about childhood behaviors he or she might otherwise have overlooked. In addition, clinicians may consider accepting a comparably low number of retrospectively recalled childhood symptoms as sufficient for fulfillment of the age-of-onset criterion of the DSM-5. Our data suggest that the diagnosis of ADHD should not be dismissed in adults who have a full ADHD syndrome but recall less than three ADHD childhood symptoms. Finally, the present findings demonstrate that cross-informant agreement for retrospective ratings of childhood ADHD is at most moderate, and therefore support the recommendation of the European Network Adult ADHD (ENAA) that information from different sources should be used when diagnosing ADHD in adults (Kooij et al. 2019).

Supplementary Information The online version contains supplementary material available at (<https://doi.org/10.1007/s10862-020-09852-1>)

Authors' Contributions Manfred Döpfner and Elena von Wirth developed the study conception and design. Data collection was performed by Janet Mandler and Manfred Döpfner. Data analysis was performed by Elena von Wirth and Dieter Breuer. The first draft of the manuscript was written by Elena von Wirth and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; grant number LE 575/3–1).

Compliance with Ethical Standards

Conflict of Interest Manfred Döpfner served in an advisory or consultancy role for Lilly, Medice, Novartis, Shire, Takeda and Vifor Pharma. He received conference attendance support, conference support, or speaker's fees from Lilly, Medice, Novartis, Takeda and Shire. He is or has been involved in clinical trials conducted by Lilly, Shire, and Vifor. He is an author of rating scales investigated in this study and receives royalties from the publisher Hogrefe, Göttingen. Elena von Wirth, Janet Mandler, and Dieter Breuer declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Consent to Publish Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., revised)*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders - text revision (4th ed.)*. Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: American Psychiatric Association.
- Applegate, B., Lahey, B. B., Hart, E. L., Biederman, J., & Hynd, G. W. (1997). Validity of the age-of-onset criterion for ADHD: A report from the DSM-IV field trials. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 1211–1221.

- Barkley, R. A., & Biederman, J. (1997). Toward a broader definition of the age-of-onset criterion for attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 1204–1210.
- Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology*, 111, 279–289.
- Barsky, A. J. (2002). Forgetting, fabricating and telescoping: The instability of the medical history. *Archives of Internal Medicine*, 162(9), 981–984.
- Ben-Zeev, D., & Young, M. A. (2010). Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports. *Journal of Nervous and Mental Disease*, 198, 280–285.
- Breda, V., Rohde, L. A., Menezes, A. M. B., Ansemi, L., Caye, A., Rovaris, D. L., Vitola, E. S., Bau, C. H. D., & Grevet, E. H. (2019). Revisiting ADHD age-of-onset in adults: To what extent should we rely on the recall of childhood symptoms? *Psychological Medicine*, 1–10.
- Breda, V., Rovaris, D. L., Vitola, E. S., Mota, N. R., Blaya-Rocha, P., Salgado, C. A. I., Victor, M. M., Picon, F. A., Karam, R. G., Silva, K. L., Rohde, L. A., Bau, C. H. D., & Grevet, E. H. (2016). Does collateral retrospective information about childhood attention-deficit/hyperactivity disorder symptoms assist in the diagnosis of attention-deficit/hyperactivity disorder in adults? Findings from a large clinical sample. *Australian & New Zealand Journal of Psychiatry*, 50, 557–565.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Dias, G., Mattos, P., Coutinho, G., Segenreich, D., Saboya, E., & Ayrão, V. (2008). Agreement rates between parent and self-report on past ADHD symptoms in an adult clinical sample. *Journal of Attention Disorders*, 12, 70–75.
- Döpfner, M., Breuer, D., Schürmann, S., Wolff Metternich, T., Rademacher, C., & Lehmkuhl, G. (2004). Effectiveness of an adaptive multimodal treatment in children with attention-deficit hyperactivity disorder - global outcome. *European Child and Adolescent Psychiatry*, 13(Suppl 1), I/117–I/129.
- Döpfner, M., Görtz-Dorten, A., Lehmkuhl, G., Breuer, D., & Goletz, H. (2008). *Diagnostik-system für psychische Störungen nach ICD-10 und DSM-IV für kinder und Jugendliche-II (DISYPS-II) [diagnostic system for mental disorders in child and adolescence according to ICD-10 and DSM-IV: DISYPS-II]*. Bern, Switzerland: Huber.
- Döpfner, M., Hautmann, C., Görtz-Dorten, A., Klasen, F., Ravens-Sieberger, U., & The BELLA study group. (2015). Long-term course of ADHD symptoms from childhood to early adulthood in a community sample. *European Child & Adolescent Psychiatry*, 24, 665–673.
- Döpfner, M., Ise, E., Wolff Metternich-Kaizman, T., Schürmann, S., Rademacher, C., & Breuer, D. (2015). Adaptive multimodal treatment for children with attention-deficit / hyperactivity disorder: An 18 month follow-up. *Child Psychiatry & Human Development*, 46, 44–56.
- Döpfner, M., Ise, E., Breuer, D., Rademacher, C., Wolff Metternich-Kaizman, T., & Schürmann, S. (2020). Long-term course after adaptive multimodal treatment for children with ADHD: An 8-year follow-up. *Journal of Attention Disorders*, 24(1), 365–375.
- Döpfner, M., & Lehmkuhl, G. (2000). *Diagnostik-system für Psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV (DISYPS-KJ) [diagnostic system for mental disorders in child and adolescence according to ICD-10 and DSM-IV: DISYPS-KJ]*. Bern, Switzerland: Huber.
- Döpfner, M., Lehmkuhl, G., & Steinhausen, H. C. (2006). *Kinder-Diagnostik-System (KIDS), Band 1: Aufmerksamkeitsdefizit- und Hyperaktivitätsstörungen (ADHS) [children-diagnostic-system (KIDS), Volume 1: attention-deficit/hyperactive disorders]*. Göttingen: Hogrefe.
- Döpfner, M., Mandler, J., Breuer, D., Dose, C., Walter, D., & von Wirth, E. (2020). Children with attention-deficit/hyperactivity disorder grown up: An 18-year follow-up after multimodal treatment. *Journal of Attention Disorders*. <https://doi.org/10.1177/1087054720948133>.
- Dunn, O. J., & Clark, V. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, 64(325), 366–377.
- Ebner-Priemer, U. W., Kuo, J., Welch, S. S., Thielgen, T., Witte, S., Bohus, M., & Linehan, M. M. (2006). A valence-dependent group-specific recall bias of retrospective self-reports. *Journal of Nervous and Mental Disease*, 194, 775–779.
- Eich, E., Reeves, J. L., Jaeger, B., & Graff-Radford, S. B. (1985). Memory for pain: Relation between past and present pain intensity. *Pain*, 23(4), 375–380.
- Epstein, J. N., & Loren, R. E. A. (2013). Changes in the definition of ADHD in DSM-5: Subtle but important. *Neuropsychiatry*, 3, 455–458.
- Erhart, M., Döpfner, M., Ravens-Sieberger, U., & Bella Study Group. (2008). Psychometric properties of two ADHD questionnaires: Comparing the Conners' scale and the FBB-HKS in the general population of German children and adolescents – Results of the BELLA study. *European Child and Adolescent Psychiatry*, 17(Suppl 1), 106–115.
- Faraone, S. V., Biederman, J., & Mick, E. (2006). The age-dependent decline of attention deficit hyperactivity disorder: A meta-analysis of follow-up studies. *Psychological Medicine*, 36, 159–165.
- Henry, N., Moffitt, T. E., Caspi, A., Langley, J., & Silva, J. M. (1994). On the “remembrance of things past”: A longitudinal evaluation of the retrospective method. *Psychological Assessment*, 6, 92–101.
- Kessler, R. C., Adler, L. A., Barkley, J., Conners, K. C., Faraone, S. V., Greenhill, L. L., et al. (2005). Patterns and predictors of attention deficit/hyperactivity disorder persistence into adulthood: Results from the national comorbidity survey replication. *Biological Psychiatry*, 57, 1442–1451.
- Kessler, R., Berglund, P., Demler, O., Jin, R., Merikangas, K., & Walters, E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 593–602.
- Kieling, C., Kieling, R. R., Rohde, L. A., Frick, P. J., Moffitt, T., Nigg, J. T., Tannock, R., & Castellanos, F. X. (2010). The age at onset of attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 167, 14–16.
- Kooij, J. J. S., Bijlenga, D., Salerno, L., Jaeschke, R., Bitter, I., Balázs, J., Thome, J., Dom, G., Kasper, S., Nunes Filipe, C., Stes, S., Mohr, P., Leppämäki, S., Casas, M., Bobes, J., McCarthy, J. M., Richarte, V., Kjemis Philipsen, A., Pehlivanidis, A., Niemela, A., Styr, B., Semerci, B., Bolea-Alamanac, B., Edvinsson, D., Baeyens, D., Wynchank, D., Sobanski, E., Philipsen, A., McNicholas, F., Caci, H., Mihailescu, I., Manor, I., Dobrescu, I., Saito, T., Krause, J., Fayyad, J., Ramos-Quiroga, J. A., Foeken, K., Rad, F., Adamou, M., Ohlmeier, M., Fitzgerald, M., Gill, M., Lensing, M., Motavalli Mukaddes, N., Brudkiewicz, P., Gustafsson, P., Tani, P., Oswald, P., Carpentier, P. J., de Rossi, P., Delorme, R., Markovska Simoska, S., Pallanti, S., Young, S., Bejerot, S., Lehtonen, T., Kustow, J., Müller-Sedgwick, U., Hirvikoski, T., Pironti, V., Ginsberg, Y., Félegyházy, Z., Garcia-Portilla, M. P., & Asherson, P. (2019). Updated European consensus statement on diagnosis and treatment of adult ADHD. *European Psychiatry*, 56, 14–34.
- Kooij, J. J. S., Boonstra, A. M., Swinkels, S. H. N., Bekker, E. M., De Noord, I., & Buitelaar, J. K. (2008). Reliability, validity, and utility of instruments for self-report and informant report concerning symptoms of ADHD in adult patients. *Journal of Attention Disorders*, 11(4), 445–458.

- Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362.
- Langberg, J. M., Epstein, J. N., Altaye, M., Molina, B. S. G., Arnold, L. E., & Vitiello, B. (2008). The transition to middle school is associated with changes in the developmental trajectory of ADHD symptomatology in young adolescents with ADHD. *Journal of Clinical Child & Adolescent Psychology*, 37, 651–663.
- Loney, J., Ledolter, J., Kramer, J. R., & Volpe, R. J. (2007). Retrospective ratings of ADHD symptoms made at young adulthood by clinic-referred boys with ADHD-related problems, their brothers without ADHD, and control participants. *Psychological Assessment*, 19, 269–280.
- Magnússon, P., Smári, J., Baldursson, D. S., Kristjánsson, J. S., Sigurbjörnsdóttir, S. H., & Guomundsson, O. O. (2006). Validity of self-report and informant rating scales of adult ADHD symptoms in comparison with a semistructured diagnostic interview. *Journal of Attention Disorders*, 9(3), 494–503.
- Mannuzza, S., Klein, R. G., Donald, D. F., Bessler, A., & Shrout, P. (2002). Accuracy of adult recall of child attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 159, 1882–1888.
- Margraf, J., Taylor, C. B., Ehlers, A., Roth, W. T., & Agras, W. S. (1987). Panic attacks in the natural environment. *Journal of Nervous and Mental Diseases*, 175, 558–565.
- Miller, C. J., Newcorn, J. H., & Halperin, J. M. (2010). Fading memories – Retrospective recall inaccuracies in ADHD. *Journal of Attention Disorders*, 14, 7–14.
- Moffitt, T. E., Houts, R., Asherson, P., Belsky, D. W., Corcoran, D. L., Hammerle, M., Harrington, H. L., Hogan, S., Meier, M. H., Polanczyk, G. V., Poulton, R., Ramrakha, S., Sugden, K., Williams, B., Rohde, L. A., & Caspi, A. (2015). Is adult ADHD a childhood-onset neurodevelopmental disorder? Evidence from a four-decade longitudinal cohort study. *American Journal of Psychiatry*, 172, 967–977.
- Mörstedt, B., Corbisiero, S., Bitto, H., & Stieglitz, R. D. (2015). Attention-deficit/hyperactivity disorder (ADHD) in adulthood: Concordance and differences between self and informant perspectives on symptoms and functional impairment. *PLoS One*, 10, e0141342.
- Murphy, K., & Barkley, R. A. (1996). Attention deficit hyperactivity disorder in adults: Comorbidities and adaptive impairments. *Comprehensive Psychiatry*, 37, 393–401.
- Murphy, P., & Schachar, R. (2000). Use of self-ratings in the assessment of symptoms of attention hyperactivity disorder in adults. *American Journal of Psychiatry*, 157, 1156–1159.
- Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, 56(3), 345–365.
- Sibley, M. H., Swanson, J. M., Arnold, E. E., Hechtman, L. T., Owens, E. B., Stehli, A., et al. (2017). Defining ADHD symptom persistence in adulthood: Optimizing sensitivity and specificity. *Journal of Child Psychology and Psychiatry*, 58, 655–662.
- Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, 14(3), 189–207.
- Todd, R., Huang, H., & Henderson, C. (2008). Poor utility of the onset criterion for DSM-IV attention deficit/hyperactivity disorder: Recommendations for DSM-V and ICD-11. *Journal of Child Psychology and Psychiatry*, 49, 942–949.
- Uchida, M., Spencer, T. J., Faraone, S. V., & Biederman, J. (2018). Adult outcome of ADHD: An overview of results from the MGH longitudinal family studies of Pediatrically and psychiatrically referred youth with and without ADHD of both sexes. *Journal of Attention Disorders*, 22, 523–534.
- World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems 10th revision*. Geneva: World Health Organization.
- World Health Organization (2018). *International Classification of Diseases 11th Revision*. <https://icd.who.int/en>.
- Zucker, M., Morris, M. K., Ingram, S. M., Morris, R. D., & Bakeman, R. (2002). Concordance of self- and informant ratings of adults' current and childhood attention-deficit/hyperactivity disorder symptoms. *Psychological Assessment*, 14, 379–389.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.