Published in final edited form as:

J Child Psychol Psychiatry. 2015 May ; 56(5): 577–585. doi:10.1111/jcpp.12325.

How interview questions are placed in time influences caregiver description of social communication symptoms on the ADI-R

Rebecca M. Jones¹, Susan Risi², Diana Wexler¹, Deborah Anderson¹, Christina Corsello³, Andrew Pickles⁴, and Catherine Lord¹

¹Weill Cornell Medical College, Center for Autism and the Developing Brain, New York Presbyterian Hospital, White Plains, NY, USA

²Cornell Rd, Ypsilanti, MI, USA

³Rady Children's Hospital, San Diego, CA, USA

⁴Department of Biostatistics, Institute of Psychiatry, King's College London, UK

Abstract

Background—Caregiver report is crucial for the diagnosis of childhood onset psychiatric disorders, particularly autism. Three experiments were conducted to determine whether caregiver reports of past and current behaviors are affected by question timing and ordering.

Methods—Using the Autism Diagnostic Interview – Revised (ADI-R), two studies systematically varied the order in which caregivers were asked about behaviors. In a third study, descriptions of children's current behaviors at age 5 were compared to retrospective descriptions of behaviors at age 5 collected at age 10.

Results—Caregivers who were first asked about a history of symptoms, described less severe past and present behavior than caregivers reporting current behaviors as well as caregivers reporting current and history of symptoms together. Caregivers retrospectively reported more severe behaviors for age 5 when their children were age 10 than they had when their children were age 5.

Conclusions—Caregivers describe past behaviors differently depending on whether they are asked about current symptoms first. Methods of caregiver reporting can influence interpretations of symptom severity with effects on diagnoses and research findings.

Keywords

Autism spectrum disorders; ADI-R; parent report; longitudinal; diagnosis

Correspondence Rebecca M. Jones, Weill Cornell Medical College, Center for Autism and the Developing Brain, New York Presbyterian Hospital, 21 Bloomingdale Road, White Plains, NY 10605, USA; rej2004@med.cornell.edu.

Conflicts of interest statement: See Acknowledgements for disclosure.

C.L. receives royalties for the ADI-R and ADOS, for which profits are donated to charity. The other resources have declared that they have no competing or potential conflicts of interest.

Introduction

A history of core features of autism, often acquired through caregiver report, is standard diagnostic practice for autism spectrum disorder (ASD: American Psychiatric Association, 1994; 2013). Yet research suggests caregivers' description of developmental milestones are affected by multiple factors, including the child's current or past level of impairment and the distance in time from the past to the present (Hus et al., 2011).

Severity of ASD features in parent reports can sometimes be accounted for by severity of nonspecific general problem behaviors and adaptive skills as much by autism symptoms (Hus and Lord, 2012, Snow et al., 2009). Current functioning also influences caregiver histories in disorders such as ADHD (Miller et al., 2010). Thus, such issues (Taylor et al., 2014) have significant clinical implications for diagnosis but even more for determining changes in conditions over time.

Our hypothesis is that caregivers anchor their perception of past symptoms on the current functioning of the child. We test this hypothesis in three experimental studies using the ADI-R. In study 1, the standard format of the ADI-R was manipulated so that a) caregivers were first administered all questions about current functioning and then asked the same items about past behavior or b) the order was reversed with all questions about the past questions preceding present. In study 2, data from study 1 was compared to data for comparable participants who received the ADI-R administered in standard format (both current and history are queried together). A third study examined longitudinal data collected at both age 5 and age 10 to determine whether caregivers retrospectively reported more severe symptoms recalling 5 year-old behaviors at age 10 compared to descriptions of the behaviors when children were actually 5. For study 1 and 2, we predicted that caregivers would describe more severe symptoms overall when asked to report about current behaviors before reporting about the past, due to a tendency to anchor perceptions of the past as generally worse than the present. Similarly, we hypothesized that, when queried at the 10 year-old assessment, caregivers would recall the child's behaviors at age 5 as being relatively more severe than the description obtained contemporaneously.

Methods

Participants

Study 1—Participants were the caregivers of children and adults who were 5 years 1 month to 38 years of age, recruited at the University of Michigan Autism and Communication Disorders Center (UMACC), who were seeking a clinical evaluation for their children with autism. Caregivers were randomized across consecutive appointments for initial evaluations at UMACC, with examiner and sex of the participant stratified, into the following groups: 1) Caregivers of 41 children (32 male) asked all of the ADI-R questions regarding their children's current behavior on the ADI-R and then all of the same questions again, this time pertaining to prior symptoms (*Current 1st*) or 2) Caregivers of 40 children (29 male) were asked to report prior history of symptoms on the ADI-R first followed by questions regarding current behaviors (*History 1st*). In the *Current 1st* and *History 1st* groups respectively, the ADI-R was completed by 14 biological mothers, 22 biological mothers and

fathers together and the remaining completed by the father or other relatives. Two caregivers in the *Current 1st* group and one caregiver in the *History 1st* group were excluded due to missing data. Sample demographics (39 participants in each group) are summarized in Table 1.

Study 2—Participants were caregivers of children and adults 5 years, 1 month of age and older (oldest 34 years of age) who completed the ADI-R at UMACC during the same years as those in study 1, age matched to study 1. Caregivers of 160 children (114 male) were asked about current behaviors and prior history of symptoms on the ADI-R according to the typical format (*Standard*). Caregivers included had either participated in other research projects that required a standard ADI-R or were not able to complete the ADI-R as formatted in study 1 due to time constraints. The ADI-R was completed by 71 biological mothers, 19 biological fathers, 57 biological mothers and fathers together and 7 legal guardians or other relatives; caregiver relationship was missing for 6 participants. Data was excluded from eight caregivers whose children were missing cognitive testing. See Table 1 for demographics (152 participants).

Study 3—102 caregivers of children with suspected ASD or nonspectrum developmental disorders were recruited from agencies in North Carolina for young children with developmental delays to participate in a longitudinal study. Children and their caregivers were seen at 4-5 years and at age 10. The families reported here represent a subset of children who had ADI-R data at ages 4-5 and 10; children were excluded if missing the ADI-R at either time point or an item on the algorithm. See Table 2 for demographics (94 participants). Details about the entire sample at various time points are published elsewhere (Lord et al., 2006, Richler et al., 2010).

For 70 of the 94 cases, the same caregiver or combination of caregivers at T1 completed the ADI-Rs at T2. In the remaining 24: 9 cases at T2 the ADI-R was completed by both caregivers when previously only completed by one caregiver; 4 cases at T2 one caregiver completed the ADI-R and was completed by both caregivers at T1, 3 cases the ADI-R was completed by the other caregiver (mother versus father) from T1 to T2; and 8 cases there was a switch from the mother and/or father to another relative. Control analyses for changes in informant are described below. Two individuals were missing informant data; three individuals were missing T2 ADOS scores and one individual was missing a best estimate diagnosis; these individuals were excluded from informant, ADOS CSS and diagnosis analyses respectively. Caregivers provided informed consent and children provided assent when possible, approved by the Institutional Review Boards at the University of Michigan, the University of Chicago and the University of North Carolina.

Measures

Autism Diagnostic Interview – Revised—The Autism Diagnostic Interview – Revised (ADI-R) (Rutter et al., 2003) is a semi-structured diagnostic interview administered to a caregiver by a clinician trained on the instrument. Items are coded for current behaviors and prior history of symptoms (either ‘Most Abnormal at 4.0-5.0 years of age’ or ‘Most Abnormal Ever’ which includes current behaviors). During the standard ADI-R, caregivers

report about current and prior history of symptoms in response to a single item, and then describe current and prior behaviors for the next item.

Following the ADI-R conventions, items are scored '0', '1', '2' or '3.' When scoring the algorithm, codes of 2s and 3s are collapsed. Items were summed to provide a total of past (diagnostic 'Most Abnormal/Ever') and current ('Current') algorithm items. Algorithm items are divided into three domains: social, communication and restricted, repetitive behaviors. The communication domain consists of verbal and nonverbal subdomains. In order to pool data from verbal and nonverbal participants, verbal totals were excluded for most analyses. Secondary analyses summed scores from social and nonverbal communication into a 'social communication domain' and a 'restricted, repetitive, stereotyped patterns of behavior (RRB) domain' reflecting DSM-5 and previous research (Snow et al., 2009).

Study 1—Questions on the ADI-R were altered from the original format to test whether caregiver report differs based upon the order in which current and 'ever' questions were asked. The *Current 1st* group received questions about all current behaviors followed by the same questions regarding those behaviors in the past. Conversely, caregivers in the *History 1st* group received questions about prior history of behaviors in their children, and then were asked about the same behaviors in the present. Unlike the typical ADI-R, participants in the *Current 1st* and *History 1st* received all questions regardless of age.

Study 2—Caregivers received the standard ADI-R format. Questions about current behaviors not typically asked for all ages (Direct Gaze, all A2 items, Failure to Develop Peer Relationships and B4 Nonverbal Communication items) were excluded.

Study 3—Caregivers were administered the standard ADI-R at T1 and T2. To determine whether caregivers had differential perceptions of behaviors at age 5 when retrospectively reporting, the '4-5 most abnormal' items were analyzed. A second analysis focused on 'ever' scores. A single total at T1 and a single total at T2 on the 'ever' time questions included two items from A4 ('Use of Other's Body to Communicate' and 'Inappropriate Facial Expressions'), and all C items (Restricted, Repetitive and Stereotyped Patterns of Behavior).

Cognitive Testing and Autism Diagnostic Observation Schedule—Cognitive ability was assessed to determine whether caregiver report was influenced by their children's cognitive skills. Nonverbal IQ (NVIQ) and verbal IQ (VIQ) scores were derived using age and language appropriate cognitive tests selected according to a standard hierarchy used by the clinic: Differential Scales of Ability (Elliott, 1990); Mullen Scales of Early Learning (Mullen, 1995), Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999); Wechsler Intelligence Scale for Children (Wechsler, 2003); Ravens Progressive Matrices (Raven et al., 2003) and Peabody Picture Vocabulary Test (Dunn and Dunn, 1997).

The Autism Diagnostic Observation Schedule (ADOS) is a standardized observational assessment organized into four modules based on the participant's age and expressive language level. ADOS-2 algorithms (Lord et al., 2012) yield a total score which can be used

to derive a calibrated severity score (CSS) (1 -10) that is relatively independent of developmental age and cognitive ability (Gotham et al., 2009, Hus and Lord, 2014).

Following an evaluation with the ADI-R and the ADOS, a diagnosis of ‘autism’, ‘PDD-NOS’, ‘Asperger's Disorder’ or ‘non-spectrum’ was determined by trained staff. The best estimate diagnosis was based upon psychometric and diagnostic algorithm scores as well as direct child observations (Risi et al., 2006). Diagnoses of Autism, PDDNOS or Asperger's were pooled within ASD. In Study 1, 13 of the 39 *Current 1st* group were classified as “non-spectrum” VIQ (M = 68.31, SD = 28.56), NVIQ (M = 68.31, SD = 28.56) and 26 individuals were on the spectrum VIQ (M = 81.77, SD = 37.21), NVIQ (M = 85.54, SD = 32.88). In the *History 1st* group, 17 of the 39 were classified as “non-spectrum” VIQ (M = 93.35, SD = 26.25), NVIQ (M = 88.94, SD = 28.55) and 22 on the spectrum VIQ (M = 94.18, SD = 32.74) and NVIQ (M = 94.59, SD = 29.28). In study 2, 61 of the 152 individuals in the *Standard* group were classified as non-spectrum VIQ (M = 82.92, SD = 27.25), NVIQ (M = 81.61, SD = 26.97) and 91 on the spectrum VIQ (M = 89.82, SD = 28.87), NVIQ (M = 93.46, SD = 25.04). In study 3, at T1, 29 children were classified as non-spectrum. Down Syndrome, Williams Syndrome, Mental Retardation, Fetal Alcohol Syndrome, Mood/Anxiety Disorders, Tourette's, and Oppositional Defiant Disorder/Attention Deficit Hyperactivity Disorder were common diagnoses in the non-spectrum groups.

Data Analysis

Study 1 & 2—ADI-R Diagnostic algorithm totals and Current totals were analyzed separately in two univariate ANOVA analyses with a fixed factor of group: study 1 (*Current 1st*; *History 1st*) and study 2 (*Current 1st*; *History 1st*; *Standard*), covarying age of the participant at the time of the ADI-R. In study 1, there was a difference in VIQ ($p = 0.03$) and a trend ($p = 0.07$) of a difference in NVIQ between *Current 1st* and *History 1st*, when covarying for age. Groups were randomly assigned; therefore the group IQ difference is by chance. Consequently, VIQ and NVIQ were included as separate covariates in all analyses and centered to the mean of each group (see table 1 for group averages). Paternal and maternal education were correlated in the three groups (p 's < 0.001), therefore secondary analyses included maternal education as a proxy for education, as well as diagnosis (ASD vs. non-spectrum) and ADOS CSS as covariates. We did not examine gender effects due to insufficient statistical power.

To determine group differences within ADI-R domains, social communication and restricted, repetitive behavior totals were entered separately as the dependent variable into a univariate ANOVA analysis with a fixed factor of group: study 1: (*Current 1st*; *History 1st*) or study 2: (*Current 1st*; *History 1st*; *Standard*) covarying age of the child at the time of the ADI-R. Verbal communication totals were assessed in separate analyses, which omitted 4 children in *Current 1st* and 11 children in *Standard* who were non-verbal. Post-hoc analyses on significant main effects of group used independent sample t -tests, equal variances not assumed. All analyses were conducted using SPSS 19.0 and considered significant at two-tailed $p < 0.05$.

Study 3—A single ADI-R social and nonverbal communication total from T1 (5 years of age) and T2 (10 years of age) on ‘age 4-5’ ADI-R items was entered as a dependent variable in a 2X1 Repeated Measures ANOVA. Mean centered NVIQ and VIQ (see Table 2 for mean) from T1 were entered as separate covariates to control for the cognitive functioning of the child. Analyses were repeated with mean centered NVIQ and VIQ separately from T2 to determine whether cognitive functioning of the child at age 10 influenced caregivers’ retrospective reporting. Covariates of 1) whether a change in informant on the ADI-R ($N = 24$ had a change) occurred between T1 and T2 (1 = different informant, 0 = no change) 2) whether the child received an ASD diagnosis versus non-spectrum 3) ADOS severity scores (T1 $N = 92$; T2 $N = 91$) and 4) maternal education as a proxy for education, were included in independent analyses.

To determine a global effect of time, i.e. ‘Ever’ versus ‘Current,’ on how caregivers report symptoms, a single total at T1 and a single total at T2 on ‘Ever’ ADI-R questions were entered in a 2X1 Repeated Measures ANOVA. Current behaviors from the 4-5 questions at T2 were compared to the same questions from T1. NVIQ and VIQ at T1 and T2 were entered as covariates as described above.

Results

Study 1

Caregivers who were asked questions about their child's current behavior first (*Current 1st*) reported more severe current symptoms, compared to caregivers who were asked to report prior history of the same behaviors first (*History 1st*) ($F(1,74) = 9.200, p < 0.003, \eta^2 = 0.11$) (Figure 1a). There was an effect of age at the time of the ADI-R ($F(1,74) = 5.275, p < 0.024, \eta^2 = 0.07$), caregivers reported more severe current behaviors in children who were younger ($r = -0.254, p < 0.025$). In addition, caregivers in the *Current 1st* group reported more severe behaviors in the past compared to caregivers in the *History 1st* group ($F(1,74) = 6.618, p < 0.012, \eta^2 = 0.08$). There was a significant interaction with VIQ ($F(1,74) = 9.052, p < 0.004, \eta^2 = 0.109$), as all caregivers regardless of their order of the questions that they received, described more severe symptoms in the past in children with lower VIQ ($r = -0.375, p < 0.001$). There were no effects of NVIQ or maternal education.

Similar to the findings for totals, caregivers in the *Current 1st* group reported more severe current social communication symptoms, as measured by the sum of the social and nonverbal communication domains, compared to caregivers in the *History 1st* group ($F(1,74) = 8.462, p < 0.005, \eta^2 = 0.1$) (Figure 1b). Caregivers in the *Current 1st* group also reported more severe social nonverbal communication behaviors in the past compared to caregivers in the *History 1st* group ($F(1,74) = 6.954, p < 0.010, \eta^2 = 0.09$). In contrast, there were no significant differences in how the groups described current and past restricted and repetitive behaviors (p 's > 0.105). Caregivers in the *Current 1st* group reported more severe current verbal communication behaviors compared to caregivers in the *History 1st* group ($F(1,70) = 5.438, p < 0.02, \eta^2 = 0.07$) but no difference in past verbal communication behaviors ($p > 0.098$).

There was a significant relationship between ADOS severity scores and caregivers' current reporting of symptoms ($F(1,74) = 26.124, p < 0.001, \eta^2 = 0.261$) and prior history reporting ($F(1,74) = 17.130, p < 0.001, \eta^2 = 0.188$), when including NVIQ as a covariate. There was a positive correlation between severity of current symptoms on the ADOS and both the current ($r = 0.512, p < 0.001$) and the past ($r = 0.438, p < 0.001$) reports of severity from the ADI-R. As expected, there was a significant relationship between clinical diagnosis (ASD versus non-spectrum) and caregivers' reporting of current symptoms ($F(1,73) = 14.243, p < 0.001, \eta^2 = 0.163$) and prior history of symptoms ($F(1,73) = 13.114, p < 0.001, \eta^2 = 0.152$), when including NVIQ and age of the child as covariates. These results are important because they support the validity of the parents' reporting in general. The main effect of "order of questions" remained after including diagnosis as a covariate for current symptoms ($p < 0.009$) and was $p = 0.064$ for prior history of symptoms. Post-hoc analyses, not surprisingly, demonstrated that caregivers described more severe current symptoms ($t(1,76) = -3.796, p < 0.001$) and prior history of symptoms ($t(1,76) = -3.442, p < 0.001$) in individuals who had ASD versus those who were nonspectrum.

Study 2

Consistent with our hypotheses, the perceived severity of current symptoms depended on the order of the questions caregivers received (*Current 1st vs. History 1st vs. Standard*) ($F(2,225) = 3.865, p < 0.022, \eta^2 = 0.033$) (Figure 2a). Caregivers who answered prior history questions first (*History 1st*) reported less severe current symptoms than those who were asked about current behaviors and prior history of symptoms together (*Standard*) ($t(1,189) = -2.733, p < 0.008$). As found in study 1, caregivers who first reported current behaviors (*Current 1st*) reported more severe current symptoms, compared to caregivers who first reported prior history of behaviors (*History 1st*) ($t(1,76) = 2.916, p < 0.005$). There was no difference in reporting current symptoms between *Current 1st* and *Standard* ($p = 0.303$). When including VIQ, as a covariate, the perceived severity of current symptoms interacted with VIQ ($F(1,225) = 3.989, p < 0.047, \eta^2 = 0.017$), all caregivers regardless of their order of the questions that they received, described more current severe symptoms in children with lower VIQ ($r = -0.119, p = 0.072$) and the main "order of questions" group effect became a trend $p = 0.065$.

The perceived severity of past symptoms also depended on how the ADI-R was administered to caregivers (*Current 1st vs. History 1st vs. Standard*) ($F(2,225) = 3.102, p < 0.047, \eta^2 = 0.027$) (Figure 2a). Caregivers who first reported past history of behaviors (*History 1st*) reported less severe past symptoms, compared to caregivers who first described current behaviors (*Current 1st*) ($t(1,76) = 2.582, p < 0.012$), and compared to caregivers who described current and past behaviors during the same question (*Standard*) ($t(1,189) = -2.284, p < 0.025$). The perceived severity of past symptoms interacted with VIQ ($F(1,225) = 14.563, p < 0.001, \eta^2 = 0.061$), all caregivers regardless of their order of the questions that they received, described more severe symptoms in the past in children with lower VIQ ($r = -0.258, p < 0.001$). The main "order of questions" group effect remained significant with VIQ as a covariate ($p < 0.035$). There was no difference in current symptoms between the *Current 1st* and *Standard* group ($p = 0.336$) or effects of NVIQ, age of the child at time of the ADI-R or maternal education.

Consistent with the domain totals, there was a difference in how caregivers reported social communication behaviors currently ($F(2,225) = 3.808, p < 0.024$), and in the past ($F(2,225) = 3.169, p < 0.044$) (Figure 2b) based upon the order of questions. Caregivers who first reported past social communication behaviors (*History 1st*) reported less severe behaviors for current and past social communication compared to caregivers who described current and past symptoms together (*Standard*) (current: $t(1,189) = -2.670, p < 0.009$; past: $t(1,189) = -2.147, p = 0.035$). As reported in Study 1, caregivers who first reported past social communication behaviors (*History 1st*) reported less severe behaviors for both current social communication ($t(1,76) = -2.777, p < 0.007$) and past social communication ($t(1,76) = -2.636, p < 0.010$), compared to caregivers who first described current behaviors (*Current 1st*). There was no difference of current or past social communication symptoms between the *Current 1st* and *Standard* groups (p 's > 0.268). There was no difference between current and past restricted, repetitive behaviors (p 's > 0.254) or current and past verbal communication scores (p 's > 0.074) across groups.

There was an effect of ADOS severity on reporting of current behaviors ($F(1,225) = 39.110, p < 0.001, \eta^2 = 0.148$) and past behaviors ($F(1,225) = 51.603, p < 0.001, \eta^2 = 0.187$) when including NVIQ as a covariate. Post hoc analyses revealed a positive correlation between caregivers' reports of current behaviors in the ADI-R and ADOS severity scores ($r = 0.388, p < 0.001$) and caregivers' reports of past behaviors ($r = 0.428, p < 0.001$). There was also a relationship between diagnosis (ASD versus non-spectrum) and caregivers' reporting of current symptoms ($F(1,224) = 40.836, p < 0.001, \eta^2 = 0.154$) and prior history of symptoms ($F(1,224) = 55.540, p < 0.001, \eta^2 = 0.199$), when including NVIQ and age of the child as covariates. As expected, caregivers' described more severe current symptoms ($t(1,228) = -6.309, p < 0.001$) and prior history of symptoms ($t(1,228) = -6.886, p < 0.001$) in individuals who had ASD versus those who were non-spectrum. The main effect of group remained when including diagnosis as a covariate for current symptoms ($p < 0.032$) and was $p = 0.067$ for prior history of symptoms. These findings, coupled with the findings from study 1, suggest that previous reports of a lack of correlation between ADOS and ADI-R scores may be accounted for by strong effects of age, IQ and current versus abnormal/ever questions. When these factors are controlled for, correlations were consistent across these studies.

Study 3

Caregivers reported more severe symptoms retrospectively when recalling 5 year-old behaviors five years later ($F(1,92) = 66.472, p < 0.001, \eta^2 = 0.422$) than when they described the same symptoms when the child was actually 5 years of age (Figure 3). There was an interaction with NVIQ at T1 ($F(1,92) = 6.914, p < 0.01, \eta^2 = 0.07$) and T2 ($F(1,92) = 12.389, p < 0.001, \eta^2 = 0.119$). Caregivers whose children had lower NVIQs at T1 described more severe behaviors at ages 5 ($r = -0.51, p < 0.001$) and 10 ($r = -0.31, p < 0.002$) with similar findings for lower NVIQs at T2 (age 5: $r = -0.58, p < 0.001$; age 10: $r = -0.32, p < 0.001$). Results were identical when including VIQ instead of NVIQ as a covariate.

Caregivers reported more severe scores at T2 compared to T1 on 44% of the 4-5 most abnormal algorithm items, no change on 47% of items and less severe behaviors at T2

compared to T1 on 9% of items. The differences in reporting remained significant when controlling for changes in informant, ADOS severity scores, ASD vs. non-spectrum diagnosis and maternal education, with no significant interactions between these variables and reporting at T1 or T2 (p 's > 0.15).

As a control analysis, we examined whether caregivers consistently inflated all scores (both current and most abnormal 4-5) at age 10, but found no significant difference in *current* symptoms at T2 compared to T1 when including NVIQ and VIQ at both time points as a covariate (p 's > 0.11). Interestingly, caregivers reported more severe symptoms at age 10 on 'ever' questions compared to the current reports at age 5 (T2 M = 7.44; T1 M = 6.57) ($F(1,93) = 6.281, p < 0.015, \eta^2 = 0.062$), suggesting that once caregivers anchor their description around current behaviors, they inflate symptoms in the past both in general and for specific time points. It is possible that more severe behaviors may have manifested and then dissipated between the ages of 4-5 and 10 that were not captured by the 4-5 years questions, but unlikely.

Discussion

Traditionally, a history of early behaviors is considered crucial for an autism diagnosis (American Psychiatric Association, 1994; 2013). Simultaneously, for immediate clinical purposes, knowledge of a child or adult's current behaviors is essential and information from the two time points is neither the same nor independent. Overall, when asked about the present before or at the same time as being asked about the past, caregivers reported more severe autism behaviors about both time periods. Moreover, asking caregivers only prior history questions without the context of current questions or vice versa, resulted in different descriptions of symptom severity. Without the "anchor point" of current behaviors, descriptions of the past were less severe. The ways in which the present and past are considered (or not considered) have clinical and research implications for characterizing behaviors.

Analyzing two longitudinal time points for the same participants at ages 5 and 10, caregivers, retrospectively described more severe behavior problems for their children when they were age 5, than they had when their children were actually 5 years old. Caregivers' retrospective reports described more abnormal behavior on 44% of items than they had described at the time, with no change on 47% of items. One possibility was when caregivers described current behaviors, they used this description to cornerstone the past, often rating the past as worse. Caregivers were not simply saying 'everything' was worse, but they rarely remembered behaviors as milder than they were at the time, particularly for children with lower IQs (Hus et al., 2011). Without the "anchor point" of current behaviors, descriptions of the past were consistently less severe across all three studies.

Score differences were driven by disparities in reporting on the social and communication behaviors and not restricted and repetitive behaviors or unusual verbal communication in the three administrations of the ADI-R (*Current 1st* ; *History 1st* ; *Standard*). Question format for the non-verbal communication and social domains ask about most abnormal behaviors during the 4-5 age period, targeting the development of pro-social behaviors, or the lack of

such behaviors. In contrast, the verbal communication and restricted, repetitive behavior (RRBs) items ask about the presence of unusual behaviors that have “ever” occurred. It is possible that memories for RRBs and the presence of unusual behaviors are more precise, compared to the absence of typical behaviors.

Given the length of the ADI-R (Ward-King et al., 2010), researchers and clinicians sometimes ask caregivers only those questions that provide the information needed to score either the past diagnostic algorithm or current behaviors. This may contribute to “current” scores collected years later appearing to be considerably milder than “ever/most abnormal 4-5” scores from years before (Seltzer et al., 2003). Several investigators have proposed data reduction methods to determine “essential” ADI-R questions for diagnosis and to document the social-communication dimensions within ASD, consistent with research priorities. This may be possible but would require re-administering and re-norming a shorter instrument. Prior research has shown that minor alterations in questionnaire answer categories can influence scores (Goodman et al., 2007), suggesting that even seemingly trivial changes to a standardization procedure impact outcomes and may affect research based on these outcomes.

The three experiments have clear implications. First, researchers and clinicians using the ADI-R, as well as potentially other caregiver reports such as the Child and Adolescent Psychiatric Assessment (CAPA) (Angold and Costello, 2000) and the Schedule for Affective Disorders and Schizophrenia for School-aged Children Present and Lifetime version (K-SADS-PL) (Kaufman et al., 1997) need to be mindful that subtle changes in administration can influence scores. Second, time affects how caregivers recall symptom severity. This is particularly important for interviews with caregivers of older children when a parent is first asked about current behaviors and then asked about the past (which is typically then described as worse). In research with the ADI-R collected at a broad range of ages, particularly with older individuals, a covariate of age should be included in analyses with ‘ever/most abnormal 4-5’ items. Comparisons of “current” scores to “most abnormal/ever” scores may be misleading in terms of symptom severity and thus diagnoses; if the goal is to quantify change, “current” needs to be compared to “current” collected in the same context (e.g., with or without most abnormal/ever scores). One possibility is to plan longitudinal studies to administer current questions first in all evaluations and then, per DSM-5 diagnoses, if historical information is needed to meet criteria, add additional questions about the past.

Limitations

There was no examination of parental factors such as caregiver stress or depression that could influence how caregivers report about their children (Hock and Ahmedani, 2012, Waters et al., 2000). This data is available from the longitudinal sample at older ages and is an area for future investigation. In addition, all three studies use a single instrument, the ADI-R, which may have unique properties. Future studies utilizing alternative instruments and addressing other disorders will be important.

Conclusions

A caregiver's perspective and ability to recall information is impacted by multiple factors. Data from three studies demonstrate that it is not only the content of the questions but context - how and when questions are asked - that provides a complete picture of autism behaviors. Caregiver descriptions are critical in informing us about the emergence of many psychiatric disorders, however the method by which clinicians or researchers obtain information can affect clinical severity, potentially impacting attempts to define trajectories over time and provide links to neurobiological mechanisms.

Acknowledgements

This study was funded by the National Institute of Mental Health (NIMH RO1 MH081873-04 and 1R01HD073975-01), Autism Speaks Meixner Fellowship (7608), and the Health Resources Services Administration (HRSA UA3-MC-11055).

References

- ANGOLD A, COSTELLO EJ. The Child and Adolescent Psychiatric Assessment (CAPA). *J Am Acad Child Adolesc Psychiatry*. 2000; 39:39–48. [PubMed: 10638066]
- AMERICAN PSYCHIATRIC ASSOCIATION. Diagnostic and statistical manual of mental disorders. American Psychiatric Association; Washington DC: 1994.
- AMERICAN PSYCHIATRIC ASSOCIATION. Diagnostic and statistical manual of mental disorders. 5th ed.. American Psychiatric Publishing; Arlington VA: 2013.
- DUNN, L.; DUNN, L. The peabody picture vocabulary test. 3rd Edition. American Guidance Service; Circle Pines, MN: 1997.
- ELLIOTT, CD. Differential Ability Scales. 2nd Ed.. Psychological Corporation; Antonio, TX: 1990.
- GOODMAN R, IERVOLINO AC, COLLISHAW S, PICKLES A, MAUGHAN B. Seemingly minor changes to a questionnaire can make a big difference to mean scores: a cautionary tale. *Soc Psychiatry Psychiatr Epidemiol*. 2007; 42:322–327. [PubMed: 17334898]
- GOTHAM K, PICKLES A, LORD C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord*. 2009; 39:693–705. [PubMed: 19082876]
- HOCK R, AHMEDANI BK. Parent perceptions of autism severity: exploring the social ecological context. *Disabil Health J*. 2012; 5:298–304. [PubMed: 23021742]
- HUS V, LORD C. Effects of child characteristics on the Autism Diagnostic Interview-Revised: implications for use of scores as a measure of ASD severity. *J Autism Dev Disord*. 2012; 54:216–224.
- HUS V, LORD C. The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *J Autism Dev Disord*. 2014; 44:1996–2012. [PubMed: 24590409]
- HUS V, TAYLOR A, LORD C. Telescoping of caregiver report on the Autism Diagnostic Interview--Revised. *J Child Psychol Psychiatry*. 2011; 52:753–760. [PubMed: 21410473]
- KAUFMAN J, BIRMAHER B, BRENT D, RAO U, FLYNN C, MORECI P, RYAN N. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*. 1997; 36:980–988. [PubMed: 9204677]
- LORD C, RISI S, DILAVORE PS, SHULMAN C, THURM A, PICKLES A. Autism from 2 to 9 years of age. *Arch Gen Psychiatry*. 2006; 63:694–701. [PubMed: 16754843]
- LORD, C.; RUTTER, M.; DILAVORE, PC.; RISI, S.; GOTHAM, K.; BISHOP, SL. Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1-4. Western Psychological Services; Torrance, CA: 2012.
- MILLER CJ, NEWCORN JH, HALPERIN JM. Fading memories: retrospective recall inaccuracies in ADHD. *J Atten Disord*. 2010; 14:7–14. [PubMed: 19794136]

- MULLEN, EM. Mullen Scales of Early Learning. American Guidance Services, Inc.; Circle Pines, MN: 1995.
- RAVEN, J.; RAVEN, JC.; COURT, JH. Manual for Raven's Progressive Matrices and Vocabulary Scales. Harcourt Assessment; San Antonio, TX: 2003.
- RICHLER J, HUERTA M, BISHOP SL, LORD C. Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Dev Psychopathol.* 2010; 22:55–69. [PubMed: 20102647]
- RISI S, LORD C, GOTHAM K, CORSELLO C, CHRYSLER C, SZATMARI P, PICKLES A. Combining information from multiple sources in the diagnosis of autism spectrum disorders. *J Am Acad Child Adolesc Psychiatry.* 2006; 45:1094–1103. [PubMed: 16926617]
- RUTTER, M.; LECOUEUR, A.; LORD, C. Autism Diagnostic Interview- Revised (ADI-R) manual. Western Psychological Services; Los Angeles: 2003.
- SELTZER MM, KRAUSS MW, SHATTUCK PT, ORSMOND G, SWE A, LORD C. The symptoms of autism spectrum disorders in adolescence and adulthood. *J Autism Dev Disord.* 2003; 33:565–581. [PubMed: 14714927]
- SNOW AV, LECAVALIER L, HOUTS C. The structure of the Autism Diagnostic Interview-Revised: diagnostic and phenotypic implications. *J Child Psychol Psychiatry.* 2009; 50:734–742. [PubMed: 19207624]
- TAYLOR CM, VEHORN A, NOBLE H, WEITLAUF AS, WARREN ZE. Brief Report: Can Metrics of Reporting Bias Enhance Early Autism Screening Measures? *J Autism Dev Disord.* 2014; 44:2375–2380. [PubMed: 24682706]
- WARD-KING J, COHEN IL, PENNING H, HOLDEN JJ. Brief report: telephone administration of the autism diagnostic interview--revised: reliability and suitability for use in research. *J Autism Dev Disord.* 2010; 40:1285–1290. [PubMed: 20195733]
- WATERS E, DOYLE J, WOLFE R, WRIGHT M, WAKE M, SALMON L. Influence of parental gender and self-reported health and illness on parent-reported child health. *Pediatrics.* 2000; 106:1422–1428. [PubMed: 11099598]
- WECHSLER, D. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation; San Antonio, TX: 1999.
- WECHSLER, D. Wechsler Intelligence Scale for Children. 4th ed.. Psychological Corporation; San Antonio, TX: 2003.

Key points

- Caregiver report is crucial for the diagnosis of childhood onset psychiatric disorders, particularly autism.
- Caregivers retrospectively reported more severe behaviors for age 5 when their children were age 10 than they had when their children were age 5.
- Manipulating the format of the ADI-R demonstrated that caregivers anchor their perception of the children's past behaviors on the basis of their current functioning.
- Methods of caregiver reporting may influence clinical diagnoses, prognoses, classifications and research data.

Figure 1.

(a) ADI-R total algorithm scores for current and past autism behaviors were compared for caregivers who reported about all current behaviors followed by reporting about behaviors in the past (Current 1st) versus caregivers who reported about prior history of behaviors and then answered questions about current behaviors (History 1st). (b) ADI-R scores for the social communication domains (Soc & Comm) were compared to totals for restricted and repetitive behavior domain (RRB). * represents a significant difference ($p < 0.05$), specific p values reported in text.

Figure 2.

(a) ADI-R total algorithm scores for current and past autism behaviors were compared for caregivers who received the ADI-R in its typical format, each item is queried about current and prior history of behaviors together (Standard), to caregivers who received current questions first (Current 1st) or history questions first (History 1st). (b) Three groups were compared on scores for the social and communication domains and restricted and repetitive behaviors domain. (RRB). * represents a significant difference ($p < 0.05$), specific p values reported in text.

Figure 3.

At age 10 (T2), caregivers recalled the child's behavior at age 5 as being more severe than the descriptions they had given when the child was age 5 (T1).

Table 1

Demographics of child participants for Study 1 and 2.

	Study 1			Study 2
Group	Current 1st	History 1st	<i>p</i> value	Standard ADI-R Format
N	39	39		152
Male %	77	72	ns	70
Age in years M (SD)	11.36 (8.02)	12.13 (5.74)	ns	11.03 (5.90)
ADOS CSS M (SD)	5.31 (2.903)	4.72 (3.244)	ns	5.20 (2.905)
ASD %	67.00	56.00	ns	60.00
VIQ M (SD)	77.28 (34.79)	93.82 (29.71)	.03	87.05 (28.34)
NVIQ M (SD)	78.77 (32.61)	92.13 (28.72)	.07	88.70 (26.40)
Ethnicity *				
Caucasian	84.6	89.7	ns	83.6
African American	10.3	na	ns	2.6
Asian	na	na		3.9
American Indian	na	na		1.3
Biracial	5.1	2.6	ns	4.6
Hispanic/Other	na	7.7	ns	2.0

* N = 3 Ethnicity data missing *Standard* Group. ns = $p > 0.1$.

Table 2

Demographics of child participants for Study 3.

Time Point	N	Male %	Age in years (SD)	ADOS CSS M (SD)	ASD %	NVIQ M (SD)	VIQ M (SD)	Ethnicity %
5 years (T1)	94	80.00	4.70 (0.69)	5.28 (2.77)	69.00	67.35 (28.41)	52.49 (28.04)	63.8 Caucasian 35.1 African Am. 1.1 Biracial
10 years (T2)	94	80.00	10.01 (0.83)	6.37 (2.73)	70.00	62.84 (34.82)	54.59 (37.60)	