



Published in final edited form as:

J Child Psychol Psychiatry. 2011 July ; 52(7): 753–760. doi:10.1111/j.1469-7610.2011.02398.x.

Telescoping of caregiver report on the Autism Diagnostic Interview – Revised

Vanessa Hus^{1,2}, Amanda Taylor³, and Catherine Lord^{1,2}

¹University of Michigan Autism & Communication Disorders Center, USA

²Department of Psychology, University of Michigan, USA

³Human Development and Family Sciences, Oregon State University, USA

Abstract

Background—Delays in development are a fundamental feature in diagnosing autism spectrum disorders (ASD). Age of language acquisition, usually obtained through retrospective caregiver report, is currently used to distinguish between categories within ASD. Research has shown that caregivers often report children as having acquired developmental milestones earlier or later than they were actually achieved. The current study examines the extent to which this phenomenon, referred to as ‘telescoping,’ impacts retrospective reports provided by caregivers of children with ASD.

Methods—Participants were 127 caregivers of children referred for possible ASD or non-spectrum developmental delay. Caregivers were interviewed when children were 2, 3, 5, and 9 years of age. Caregiver-reported ages of first concern, language and non-diagnostic developmental milestones and interviewer-estimated age of onset were compared over time using linear models.

Results—Significant telescoping of language milestones resulted in more children meeting language delay criteria as they grew older, in spite of original reports that their language was not delayed. There was little evidence of consistent telescoping of caregiver-reported ages of first concern, daytime bladder control, and independent walking. With time, the interviewers’ judged ages of symptom onset increased, but remained prior to age three.

Conclusions—Telescoping of caregiver-reported ages of language acquisition has implications for both clinical diagnosis and genetic studies using these milestones to increase homogeneity of samples. Results support proposals to remove specific age-based criteria in the diagnosis of ASD. Telescoping should be considered when working with any clinical population in which retrospectively recalled events are used in diagnosis.

Keywords

Autism spectrum disorders; telescoping; retrospective caregiver-report; language milestones

© 2011 The Authors.

Correspondence to: Vanessa Hus, UMACC, 1111 E. Catherine Street, Ann Arbor, MI 48109, USA; Tel: 734.936.8600; Fax: 734.936.0068; vhus@umich.edu.

Conflict of interest statement: CL receives royalties for the ADIR; profits from this study were donated to charity.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Supplemental Tables S1–S3 and Figures S1 and S2.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Since autism was introduced to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) in 1980 and the *International Classification of Diseases* (ICD) in 1977, ages of developmental milestones and onset of symptoms have been fundamental to its diagnosis. This information is often obtained retrospectively from caregiver report. Methodological issues associated with retrospective report have significant implications for diagnostic practices. The current study investigates how forward telescoping (i.e., the tendency to report events as having occurred more recently in time than they actually occurred; Sudman & Bradburn, 1973) may impact retrospective reports made by caregivers of children with autism spectrum disorders (ASD).

Current DSM-IV-TR (American Psychiatric Association, 2000) criteria for a diagnosis of autistic disorder (AD) or pervasive developmental disorder – not otherwise specified (PDD-NOS) require that delays or difficulties in social, communication, or play skills are present before 3 years of age. Delay of spoken language (e.g., use of single words after 2 or phrases after 3 years) is also used to rule out Asperger's syndrome (AS), which requires that a child not have a clinically significant language delay. Recommended revisions for DSM-5 call for a single category of ASD to subsume previously differentiated AD, PDD-NOS and AS and for the replacement of the specific age of onset criterion with a general statement that 'symptoms must be present in early childhood' (American Psychiatric Association, 2010).

A substantial body of research has questioned the degree to which clinical presentation differs between AD, PDD-NOS and AS and whether language delay should be used to differentiate these diagnoses (see Bennett et al., 2008). Studies of young children suggest that those with a history of language delay or deviance have more severe autism symptoms (Eisenmajer et al., 1998; Gilchrist et al., 2001; Szatmari, Archer, Fisman, Streiner, & Wilson, 1995). However, when current language level is controlled, most studies including older children, adolescents and adults do not support the association between language history and symptom severity (Eisenmajer et al., 1998; Gilchrist et al., 2001; Howlin, 2003; Prior et al., 1998). Recently, directly assessed grammatical abilities at 6–8 years of age were found to be more predictive of later outcomes than diagnoses based upon caregiver-reported milestones at ages 4–6 (Bennett et al., 2008).

Inconsistent results from genetic studies of ASD also call into question whether distinct categories of ASD truly exist and whether language delay is a relevant marker for making such distinctions. Some genetic studies have attempted to reduce biological heterogeneity by limiting samples to individuals with narrowly defined AD, excluding 'milder' cases with PDD-NOS or AS (e.g., Shao et al., 2002). Others have used language milestones to stratify samples or as endophenotypes to identify quantitative trait loci (see Losh, Childress, Lam, & Piven, 2008 for review). These approaches have yielded some promising results; however, replication in independent samples is rare (Abrahams & Geschwind, 2008). This could be due to associations between retrospectively reported ages of language acquisition and other psychometric and demographic variables, such as IQ (Hus, Pickles, Cook, Risi, & Lord, 2007), or methodological issues, such as biases associated with caregiver recall.

Caregiver report is known to be influenced by many factors. For example, after a child has received a diagnosis of ASD, caregivers may tend to endorse behaviors, such as language delay, that are consistent with their child's diagnosis (Zwaigenbaum et al., 2007). Retrospective events may also be impacted by telescoping (Sudman & Bradburn, 1973). Telescoping is influenced by the amount of time between the event and the time of reporting, the interviewer's wording of a question, as well as the informant's emotional state (Barsky, 2002). Some suggest that telescoping results from a compression of the timescale, in which informants systematically 'move' events forward in time when asked to recall when they occurred (e.g., Neter & Waksberg, 1964). Others show that telescoping can also

result from more distant events being recalled with greater error than more recent events (i.e., heteroscedastic random measurement error; Pickles, Pickering, & Taylor, 1996). How onset of an event is operationalized may also play a role (Pickles et al., 1998). For example, although the child's first spoken word is considered a discrete event occurring on a specific date, what constitutes a word to caregivers may vary across time. When children are very young, approximations supported by context may be interpreted as first words (e.g., 'ba' while reaching for a bottle). In older children, however, first words may be recalled as the first time each syllable of a word was clearly enunciated (e.g., 'bottle'). Furthermore, delays in other areas of the child's development may make it difficult for caregivers to accurately recall the true ages at which specific skills were acquired.

Telescoping impacts both parent- and self-report and has been widely documented in retrospective recall of many types of events. Caregivers of children with externalizing or depressive disorders reported that their child's symptom onset occurred 6–18 months later than indicated when interviewed five years before (Todd, Huang, & Henderson, 2008). Comparisons of retrospectively reported developmental milestones in typical children were often recalled by caregivers as having occurred *earlier* than originally reported (i.e., backward telescoping; Donoghue & Shakespeare, 1967; McGraw & Molloy, 1941). Majnemer and Rosenblatt (1994) suggested that telescoping varied depending on the milestone the parent was asked to recall. In their study, 91% of caregivers recalled age of first steps within two months of the originally reported age. However, only 59% reported word acquisition within two months of the original report; 20% of caregivers reported ages that were between 6 and 25 months *later* than first indicated (i.e., evidence for forward telescoping). Although all children were originally reported to have acquired words prior to 16 months, based on ages of language acquisition provided when the children were 3 and 5 years old, many would have been suspected of having had early language delays. Consistent with other studies (Neligan & Prudham, 1969), greater differences were associated with lower developmental and cognitive test scores. Earlier informal analyses of children with ASD or developmental delays (a subset of the present sample) suggested evidence of telescoping of language milestones from age 2 to 5 (Lord, Shulman, & DiLavore, 2004).

The present study explores the potential effects of telescoping on reported ages of developmental milestones provided by caregivers of children referred for possible ASD or developmental delays prior to 3 years of age. It was hypothesized that, over time, caregivers would report later ages of concern and milestone acquisition (i.e., forward telescoping) and that telescoping of language milestones would result in an increase in the proportion of children meeting criteria for language delay. Because a clinician's overall impressions of the presence of early developmental abnormalities is pertinent to diagnosis and relies on caregiver report, interviewers' judgments of age of symptom onset were also examined.

Method

Participants

The original sample included 214 children seen as part of a longitudinal study of ASD (see Anderson et al., 2007; Lord et al., 2006). The Autism Referral (AR) group included 192 consecutive referrals to autism clinics in North Carolina and Chicago for evaluation of possible autism before 3 years of age. The non-autism, developmentally delayed (DD) control group included 22 children who had never been referred for or diagnosed with autism. Diagnoses for this group included non-specific intellectual disability, language disorders, Down syndrome or other known genetic or physical conditions, and attention deficit/hyperactivity disorder (ADHD). Individuals with moderate to severe sensory impairments, cerebral palsy, or poorly controlled seizures were excluded from both groups.

The AR group was assessed at approximately 2, 3, 5, and 9 years old and the DD group was seen at 2, 5, and 9. Hereafter, these time points will be referred to as T2, T3, T5, and T9.

Because loss of skills may make it difficult to accurately recall original milestone acquisition, children whose caregivers reported a possible or definite language regression were excluded from these analyses. Data presented here represent 127 children who had no history of language regression and had been administered the ADI-R at two or more time points. See Table 1 for description of this sample.

Data collection

Each participant was evaluated using the Autism Diagnostic Interview–Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), Autism Diagnostic Observation Schedule (Lord, Rutter, DiLavore, & Risi, 1999), Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984), and a cognitive assessment. At T2, T5, and T9, best estimate diagnoses were made by a psychologist and/or child psychiatrist, who was blind to previous diagnoses (see Risi et al., 2006). A caregiver or guardian provided informed consent in accordance with institutional review boards at both sites.

Measures

The ADI-R is a semi-structured, standardized caregiver interview designed to assess behaviors related to ASD. Caregivers are asked to recall how old their child was when they first became concerned about his/her development, as well as the ages at which s/he walked independently, used single words and phrases, and acquired bladder control. Near the end of the ADI-R, the interviewer is asked to estimate the age at which the child's developmental abnormalities first manifested, based upon information obtained during the interview. Caregivers were administered a toddler version of the ADI-R at T2 and T3 (Lord et al., 2004) and the ADI-R at T5 and T9. Relevant questions are consistent across interviews. All interviews were administered by trained examiners who met standard reliability criteria (Rutter et al., 2003).

As a measure of cognitive or developmental level, children were administered the Differential Ability Scales (Elliott, 1990), Mullen Scales of Early Learning (Mullen, 1995) or Wechsler Intelligence Scale for Children, Third Edition (Wechsler, 1991). For the purposes of this study, the most recent verbal IQ was used for each child; for 103 children, cognitive assessments from T9 were used; 17 from T5, and 7 from T3.

Design and analysis

Chosen for ability to model repeated measures data, as well as handle missing data, linear models allowing correlated and non-constant variance across time points were fitted using the MIXED procedure in SPSS 17.0. Models were fit including Time as a fixed effect. Referral group, child gender, verbal IQ, and maternal education were included as covariates. To facilitate interpretation, verbal IQ was centered at the mean for the sample (54.95). Although informants sometimes varied across interviews, because most were completed by mothers, the variation was not sufficient to test the effects of informant. Non-significant covariates and interactions were dropped from each model. Effects from the most parsimonious model, retaining, at minimum, time and referral group, are reported below. If too few participants had acquired a milestone by T2 (e.g., phrase acquisition, bladder control), analyses comparing T9 and T5 to T3 (using only the AR group), or comparing T9 to T5 (using all available participants) were used.

McNemar's χ^2 test with correction for continuity was used to compare the proportions of children who changed delay classifications due to telescoping of ages of language

acquisition. When the number of children who changed classification was less than 25, the binomial distribution was used (Sheskin, 2004).

Results

Preliminary analyses

Mean child ages at ADI-R administration for each time point and mean times between interviews are presented in Table S1. The AR group was nearly 3 months older than the DD group at T2 and 7.5 months older at T5. The mean time between T2 and T5 interviews was 4.5 months longer for the AR group compared to the DD group, $t(44.04) = 3.10$, $p = .003$. In contrast, the mean time between T2 and T9 was 6 months shorter for the AR group, $t(95.79) = -2.78$, $p = .006$.

Age of first concern

To explore possible telescoping of caregivers' reported ages of first concern, 107 children were included in the analysis (20 children were excluded because caregivers reported no concerns until after being referred). There was no significant difference between ages of first concern reported by caregivers of children in the DD compared to the AR groups. Ages of concern at T5 and T9 were not significantly older than reported at T2; see Table 2.

Age of first words

At T2, 80 children had reportedly acquired words. Caregiver-reported language status at T2 was consistent with other measures of expressive language (see Table S2), suggesting that these reports were a valid measure for comparing later-reported word acquisition.

As shown in Table S3 and Figure 1, when controlling for most recent verbal IQ, age of first words reported at T9 and T5 were significantly older than those reported at T2. Children in the DD group reportedly began using words significantly earlier than children in the AR group. Discrepancies between T9 and T2 reports were significantly smaller for children with higher verbal IQs.

To investigate the impact of telescoping on classification of language delay, children were divided into groups based upon ADI-R criteria for word delay (i.e., children whose age of first word was reportedly greater than 24 months were classified as *delayed* and children with reported first words at/before 24 months were classified as *not delayed*). The proportion of children meeting word delay criteria at T2 was significantly lower than the proportion classified as *delayed* based upon reported ages of word acquisition at subsequent time points (see Figure S1). Only 24% of children who had acquired words were classified as *delayed* at T2. However, 52%, 42% and 67% met criteria for delay at T3 ($p = .002$), T5 ($p = .013$), and T9 ($p < .001$), respectively. Assuming that classification at T2 was the most accurate, 33–60% of children originally in the *not delayed* group changed classification to *delayed* when they were older. In contrast, 13–27% of children who met word delay criteria at T2 were classified as *not delayed* at later time points.

Age of first phrases

To explore telescoping of reported phrase acquisition, analyses focused on 50 AR children with reported age of first words by T3. By T5, caregivers reported phrase acquisition as occurring 5.12 months later; by T9, 8.04 months later than at T3 (see Table 2). There was no difference between age of phrase acquisition reported for children who had acquired phrases by T2 and those who had not acquired phrases until T3 (see Table S3).

To examine whether children's later acquisition of phrases impacted caregiver recall, children were divided into *early* acquisition (i.e., reportedly acquired phrases by T2 or T3, $N = 55$) and *late* acquisition groups (i.e. no phrases reported until T5, $N = 23$). There was a significant three-way interaction between time (T5, T9), early vs. late group, and most recent verbal IQ (see Table S3). As demonstrated in Figure 2, age of phrase acquisition reported at T9 was significantly older than reported at T5, with greater discrepancies in reported age of first phrases for children with lower IQs. This effect was significantly smaller for the *early* acquisition group.

Based upon the ADI-R decision point of 33 months, children were divided into *delayed* and *not delayed* phrase groups. Although differences were nonsignificant, possibly due to small sample size, the proportion of children classified as *delayed* based on reported age of first phrases at T5 (73%) and T9 (81%) was higher than that meeting delay criteria at T3 (61%; $p = .25$ and $p = .45$, respectively).

General developmental milestones

Age of independent walking and daytime bladder control were examined to determine whether telescoping impacted reports of developmental milestones not specifically associated with ASD. Because the DD group was only asked when their child first walked at T2 and T5, analyses were conducted separately for referral groups. As shown in Table 2, average ages of independent walking and bladder control did not differ significantly across time points.

Interviewer's judgment of onset

It was also of interest to investigate whether there was a significant change over time in interviewers' judgments of onset of ASD symptoms. As shown in Table 2, estimated ages of onset were significantly younger for the DD group at each time point. For the AR group, ages of onset judged by the interviewer at T9 and T5 were significantly older than at T2 (3.6 and 2.6 months, respectively). For the DD group, judged age of onset was significantly younger (.16 months) at T9 compared to T2. At each time point, the average age of onset was also judged to be somewhat later for children with higher verbal IQs, $F(1,121.29) = 4.41$, $p = .038$.

Post-hoc analysis of discrepancies in caregiver-report and interviewers' judgments

To better understand the telescoping of language ages, but not of non-diagnostic milestones, discrepancies between raw ages reported at T9 and original reports (T3 for phrases and bladder control, T2 for all others) were examined. These comparisons are provided in the Online Appendix (see Figure S2) and discussed below.

Discussion

As children grew older, caregiver-reported ages of word and phrase acquisition were forward telescoped. At ages 5 and 9, 58–75% of reports of language acquisition were more than three months later than indicated at age 2. As a result, 60% of children who did not meet criteria for language delay at age 2 changed classification to delayed at age 9. Only 7–8% of caregivers reported the same or earlier ages of acquisition.

Consistent with previous studies (Majnemer & Rosenblatt, 1994; Neligan & Prudham, 1969), discrepancies in caregiver report were greater for children with lower verbal IQs. Telescoping effects did not differ by referral group, gender, or maternal education. However, discrepancies ($M = 5.34$ months) observed between reported age of first words at 5 and 2 in our sample were greater than discrepancies observed at similar ages in Majnemer

and Rosenblatt's (1994) sample of non-delayed children ($M = 2.74$ months). Because all children had developmental delays, perhaps caregivers in our study were more apt to report early language delays consistent with their children's level of functioning. With time, interviewers' judgments of onset also tended to move to older ages, though to a lesser extent than observed with language milestones.

In contrast, caregiver-reported age of first concern was relatively stable, with average discrepancies of less than one month. Two-thirds of caregivers reported the same age of first concern at 2 and 9. The younger age of our sample may explain why discrepancies were smaller than the differences of 6–18 months in parent-reported age of symptom onset for children with externalizing or depressive disorders (Todd et al., 2008). It is also possible that onset of behavioral difficulties occurring in middle-childhood are more difficult for caregivers to pinpoint than onset of developmental concerns.

Similarly, there was little evidence of consistent telescoping of caregiver-reported ages of non-language milestones. The small number of participants available for some comparisons is a limitation of this study, but as demonstrated in previous research (McGraw & Molloy, 1941; Neligan & Prudham, 1969; Pyles, Stolz, & Macfarlane, 1935), most ages of independent walking remained within three months of original reports. In contrast, although linear models showed few *consistent* biases, 53% of discrepancies in daytime bladder control reported at 9 and 3 were *greater than* three months, with reported ages both earlier and later than originally indicated.

There are many possible reasons why telescoping was more prominent for language milestones. Responses may be influenced by caregivers' changing perceptions of what constitutes a word. Language development is a 'soft event' (i.e., a continuous process that is not marked by a discrete 'event'; Pickles et al., 1998). Consequently, it may be more difficult for caregivers to recall when their child met a specific definition of language onset (i.e., spontaneous, meaningful sound–meaning correspondences) compared to when discrete events (e.g., first steps) occurred. Furthermore, caregivers' knowledge of their child's diagnosis may bias recall of milestones (e.g., language) associated with the diagnosis more than other non-related milestones (Charman et al., 2005; Zwaigenbaum et al., 2007). Similar patterns of telescoping were observed in the DD group as for the children with ASD, suggesting that these influences on caregiver report are not specific to autism.

Implications

Telescoping of language milestones provides a caution against using retrospectively recalled language delay to differentiate subtypes of ASD. Telescoping effects were observed in the context of the ADI-R, where caregivers are asked to provide examples of first words and phrases and follow-up questions are used to clarify that these words were used meaningfully. Thus, the reliability of reports on this instrument may be expected to be higher than if a caregiver were asked to recall these milestones in an unstructured interview or questionnaire. The fact that telescoping was observed at age 5 in this sample suggests that telescoping may result in misclassification of ASDs in clinical settings where children are often referred at older ages (Shattuck et al., 2009).

The large proportion of children who were not language-delayed based on reports at age 2 but who later met criteria for delay due to telescoped age of first words presents a significant issue for genetic studies using these ages to stratify samples or as endophenotypes (e.g., Alarcon et al., 2005; Buxbaum et al., 2001; Wassink et al., 2004). If telescoping occurred in those samples, associations with language delay could be inaccurately interpreted or missed altogether. Depending on the age and cognitive-functioning level of the participants in a given sample, telescoping may be more or less of an issue. Because children in this study

were referred for possible autism at a time when relatively little was known about early signs, they may represent a more cognitively delayed group than clinic-referred samples today.

Discrepancies in caregivers' concerns and interviewers' judgments of onset may also influence studies seeking to understand patterns of onset in ASD. Retrospective studies suggest that ASD-onset is marked by an early emergence of symptoms or a loss of skills in the second year of life. However, in a recent prospective study, many children demonstrated both early parental concerns and loss of skills (Ozonoff et al., 2010). Given the relatively short time-frame during which symptom onset and/or regression may occur, discrepancies of just a few months could mask the separation of symptom onset and loss of skills, explaining why retrospective studies have suggested two distinct patterns of onset.

In situations where clinicians or researchers must rely on retrospective report, use of records (Sudman & Bradburn, 1973) or anchor points (e.g., 'was she using words on her first birthday?') may increase reliability of recall (Keller et al., 1987; Loftus & Marburger, 1983). Statistical techniques which account for measurement error due to telescoping should be used when analyzing these types of data (e.g., Pickles et al., 1994, 1996). Although these strategies may minimize telescoping effects, prospective studies provide unique, important information about patterns of onset (e.g., Ozonoff et al., 2010), as do direct assessments of language abilities, compared to retrospective reports (Bennett et al., 2008).

Conclusions

Telescoping often influences caregivers' retrospective reports of their children's developmental milestones, particularly for children with ASD and other developmental disorders who are older and have lower verbal abilities. Currently, diagnosticians often rely on such retrospective reports of early development to determine if a child meets criteria for a diagnosis of an ASD (see Bennett et al., 2008). Telescoping in these reports provides support for current proposals to remove specific age-based diagnostic criteria for ASDs in DSM-5. Although this study focuses primarily on caregivers of children with ASD, telescoping has been documented to affect caregiver and self-reports in several other populations. The potential effect of telescoping on diagnosis and classification of other disorders that depend on retrospective reports to determine whether an individual meets age of onset and other age-based diagnostic criteria should also be considered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by a graduate fellowship from the Simons Foundation to VH and NIMH grant R01MH081873-01A1 to CL. We gratefully acknowledge the families in the longitudinal study, Andrew Pickles and Brady West for statistical consultation, and Shanping Qiu for assistance compiling datasets.

References

- Abrahams BS, Geschwind DH. Advances in autism genetics: On the threshold of a new neurobiology. *Nature Review Genetics*. 2008; 9:341–355.10.1038/nrg2346
- Alarcon M, Yonan AL, Gilliam TC, Cantor RM, Geschwind DH. Quantitative genome scan and ordered-subsets analysis of autism endophenotypes support language QTLs. *Molecular Psychiatry*. 2005; 10:747–757. [PubMed: 15824743]

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: American Psychiatric Association; 2000. text revision: DSM-IV-TR
- American Psychiatric Association. DSM-5 Development. 2010. Retrieved from <http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=94#>
- Anderson DK, Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Welch K, et al. Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*. 2007; 75:594–604.10.1037/0022-006X.75.4.594 [PubMed: 17663613]
- Barsky AJ. Forgetting, fabricating, and telescoping: The instability of the medical history. *Archives of Internal Medicine*. 2002; 162:981–984.10.1001/archinte.162.9.981 [PubMed: 11996606]
- Bennett T, Szatmari P, Bryson S, Volden J, Zwaigenbaum L, Vaccarella L, Duku E, et al. Differentiating autism and Asperger syndrome on the basis of language delay or impairment. *Journal of Autism and Developmental Disorders*. 2008; 38:616–625.10.1007/s10803-007-0428-7 [PubMed: 17721697]
- Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, Lawlor BA, et al. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *The American Journal of Human Genetics*. 2001; 68:1514–1520.10.1086/320588
- Charman T, Taylor E, Drew A, Cockerill H, Brown JA, Baird G. Outcome at 7 years of children diagnosed with autism at age 2: Predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*. 2005; 46:500–513.10.1111/j.1469-7610.2004.00377.x [PubMed: 15845130]
- Donoghue EC, Shakespeare RA. The reliability of paediatric case-history milestones. *Developmental Medicine and Child Neurology*. 1967; 9:64–69.10.1111/j.1469- 8749.1967.tb02211.x [PubMed: 6031535]
- Eisenmajer R, Prior M, Leekam S, Wing L, Ong B, Gould J, Welham M. Delayed language onset as a predictor of clinical symptoms in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*. 1998; 28:527–533.10.1023/A:1026004212375 [PubMed: 9932239]
- Elliott, CD. *Differential Ability Scales*. San Antonio, TX: Psychological Corporation; 1990.
- Gilchrist A, Green J, Cox A, Burton D, Rutter M, Couteur AL. Development and current functioning in adolescents with Asperger syndrome: A comparative study. *Journal of Child Psychology and Psychiatry*. 2001; 42:227–240.10.1111/1469-7610.00714 [PubMed: 11280419]
- Howlin P. Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*. 2003; 33:3–13.10.1023/A:1022270118899 [PubMed: 12708575]
- Hus V, Pickles A, Cook EH, Risi S, Lord C. Using the Autism Diagnostic Interview–Revised to increase phenotypic homogeneity in genetic studies of autism. *Biological Psychiatry*. 2007; 61:438–448.10.1016/j.biopsych.2006.08.044 [PubMed: 17276746]
- Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, Andreasen NC. The Longitudinal Interval Follow-up Evaluation: A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry*. 1987; 44:540–548.10.1001/archpsyc.1987.01800180050009 [PubMed: 3579500]
- Loftus EF, Marburger W. Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory and Cognition*. 1983; 11:114–120.
- Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Archives of General Psychiatry*. 2006; 63:694–701.10.1001/archpsyc.63.6.694 [PubMed: 16754843]
- Lord, C.; Rutter, M.; DiLavore, PS.; Risi, S. *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, CA: Western Psychological Services; 1999.
- Lord C, Shulman C, DiLavore P. Regression and word loss in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry*. 2004; 45:936–955.10.1111/j.1469-7610.2004.t01-1-00287.x [PubMed: 15225337]
- Losh M, Childress D, Lam K, Piven J. Defining key features of the broad autism phenotype: A comparison across parents of multiple- and single-incidence autism families. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2008; 147B:424–433.10.1002/ajmg.b.30612

- Majnemer A, Rosenblatt B. Reliability of parental recall of developmental milestones. *Pediatric Neurology*. 1994; 10:304–308.10.1016/0887-8994(94)90126-0 [PubMed: 8068156]
- McGraw MB, Molloy LB. The pediatric anamnesis inaccuracies in eliciting developmental data. *Child Development*. 1941; 12:255–265.
- Mullen, E. *The Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Service; 1995.
- Neligan G, Prudham D. Norms for four standard developmental milestones by sex, social class and place in family. *Developmental Medicine and Child Neurology*. 1969; 11:413–422.10.1111/j.1469-8749.1969.tb01459.x [PubMed: 5805346]
- Neter J, Waksberg J. A study of response errors in expenditures data from household interviews. *Journal of the American Statistical Association*. 1964; 59:18–55.10.2307/2282857
- Ozonoff SP, Iosif AP, Baguio FB, Cook I, Hill MM, Hutman TP, Rogers SJP, et al. A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child*. 2010; 49:256–266.
- Pickles A, Neale M, Simonoff E, Rutter M, Hewitt J, Meyer J, Crouchley R, et al. A simple method for censored age-of-onset data subject to recall bias: Mothers' reports of age of puberty in male twins. *Behavior Genetics*. 1994; 24:457–468.10.1007/BF01076181 [PubMed: 7993322]
- Pickles A, Pickering K, Simonoff E, Silberg J, Meyer J, Maes H. Genetic 'clocks' and 'soft' events: A twin model for pubertal development and other recalled sequences of developmental milestones, transitions, or ages at onset. *Behavior Genetics*. 1998; 28:243–253.10.1023/A:1021615228995 [PubMed: 9803017]
- Pickles A, Pickering K, Taylor C. Reconciling recalled dates of developmental milestones, events and transitions: A mixed generalized linear model with random mean and variance functions. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*. 1996; 159:225–234.
- Prior M, Leekam S, Ong B, Eisenmajer R, Wing L, Gould J, Dove D. Are there subgroups within the autistic spectrum? A cluster analysis of a group of children with autistic spectrum disorders. *Journal of Child Psychology and Psychiatry*. 1998; 39:893–902.10.1111/1469-7610.00389 [PubMed: 9758197]
- Pyles MK, Stolz HR, Macfarlane JW. The accuracy of mothers' reports on birth and developmental data. *Child Development*. 1935; 6:165–176.
- Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, Cook EH, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2006; 45:1094–1103.10.1097/01.chi.0000227880.42780.0e [PubMed: 16926617]
- Rutter, M.; Le Couteur, A.; Lord, C. *Autism Diagnostic Interview–Revised*. Los Angeles, CA: Western Psychological Services; 2003.
- Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, Abramson RK, et al. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *The American Journal of Human Genetics*. 2002; 70:1058–1061.10.1086/339765
- Shattuck PT, Durkin M, Maenner M, Newschaffer C, Mandell DS, Wiggins L, Lee L, et al. Timing of identification among children with an autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2009; 48:474–483.10.1097/CHI.0-b013e31819b3848 [PubMed: 19318992]
- Sheskin, D. *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: Chapman & Hall/CRC Press; 2004.
- Sparrow, SS.; Balla, DA.; Cicchetti, DV. *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service; 1984.
- Sudman S, Bradburn NM. Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*. 1973; 68:805–815.
- Szatmari P, Archer L, Fisman S, Streiner DL, Wilson F. Asperger's syndrome and autism: Differences in behavior, cognition, and adaptive functioning. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1995; 34:1662–1672.10.1097/00004583-199512000-00017 [PubMed: 8543538]

- Todd RD, Huang H, Henderson CA. Poor utility of the age of onset criterion for DSM-IV attention deficit/hyperactivity disorder: Recommendations for DSM-V and ICD-11. *Journal of Child Psychology and Psychiatry*. 2008; 49:942–949.10.1111/j.1469-7610.2008.01892.x [PubMed: 18564071]
- Wassink TH, Piven J, Vieland VJ, Pietila J, Goedken RJ, Folstein SE, Sheffield VC. Examination of AVPR1a as an autism susceptibility gene. *Molecular Psychiatry*. 2004; 9:968–972. [PubMed: 15098001]
- Wechsler, D. Wechsler Intelligence Scale for Children. 3. New York: The Psychological Corporation; 1991. (WISC-III)
- Zwaigenbaum L, Thurm A, Stone W, Baranek G, Bryson S, Iverson J, Kau A, et al. Studying the emergence of autism spectrum disorders in high-risk infants: Methodological and practical issues. *Journal of Autism and Developmental Disorders*. 2007; 37:466–480.10.1007/s10803-006-0179-x [PubMed: 16897376]

\$watermark-text

\$watermark-text

\$watermark-text

Key points

- Telescoping of retrospective reports has been documented to impact caregiver- and self-report in several populations across different types of events, including age of language milestones and age of symptom onset.
- Significant telescoping of language milestones may influence the accuracy of diagnostic classification of different ASDs (e.g., Asperger's syndrome) and has implications for genetic studies using age of language acquisition as an endophenotype or to stratify samples.
- Telescoping of language acquisition and interviewer's judgment of onset support proposals to remove specific age-based diagnostic criteria for ASD from DSM-5.
- When relying on retrospective report of age of symptom onset or other age-based diagnostic criteria for other disorders, clinicians should be aware of the potential for telescoping.

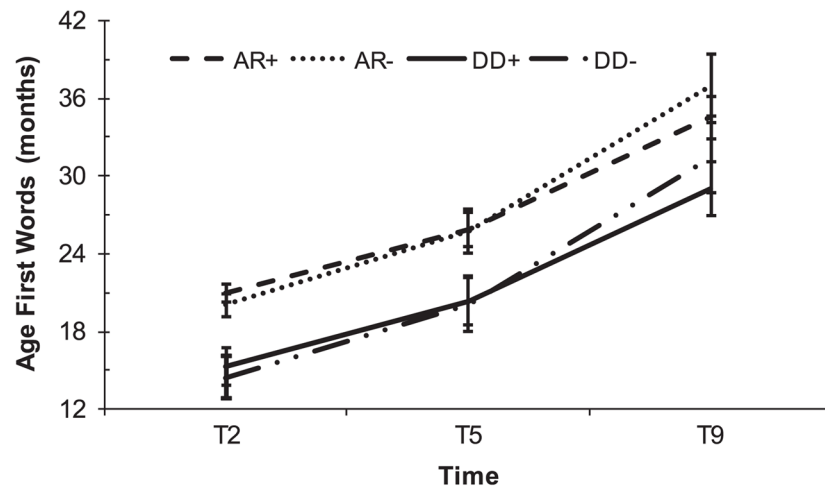


Figure 1. Caregiver-reported age of first words by referral group. Error bars represent standard errors. + = centered Verbal IQ+15; - = centered Verbal IQ-15

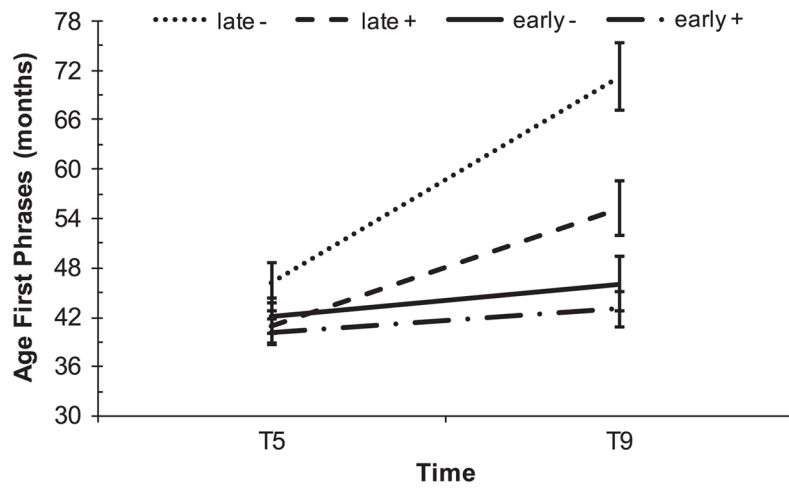


Figure 2. Age of first phrases by acquisition group and verbal IQ. Error bars represent standard errors. Late = Earliest reported phrases at T5. Early = Earliest reported phrases at T2 or T3. + = centered Verbal IQ+15; - = centered Verbal IQ-15

\$watermark-text

\$watermark-text

\$watermark-text

Table 1

Sample demographics

	AR (N = 109)	DD (N = 18)
Male (%)	82	44
Caucasian (%)	60	61
Maternal education (%)		
BA/Grad/Professional	34	28
Some college/assoc	27	22
High school or less	31	50
Not reported	8	0
Most recent diagnosis (%)		
ASD	80	0
Non-spectrum	20	100
Most recent verbal IQ [*]		
Mean (SD)	52.81 (33.94)	67.94 (23.63)
Most recent nonverbal IQ		
Mean (SD)	66.09 (31.93)	72.24 (28.25)

Grad = Graduate degree; BA = Bachelor's degree; Assoc = Associate's degree.

* AR vs. DD, $p < .05$.

Table 2

Caregiver-reported ages of developmental milestone acquisition

	Referral group	N	T2		T3		T5		T9	
			EM	(SE)	EM	(SE)	EM	(SE)	EM	(SE)
First concern	AR	98	13.90	(.70)	–	–	14.36	(.79)	13.10	(.66)
	DD	9	11.96	(1.91)	–	–	12.41	(1.88)	11.15	(1.90)
First words ^{a, b}	AR	65	14.89	(1.52)	–	–	20.24	(1.89)	30.21	(2.27)
	DD	15	20.49	(.74)	–	–	25.84	(1.38)	35.81	(1.87)
First phrases ^{a, b}	AR	50	–	–	34.41	(.94)	39.54	(1.25)	42.46	(1.90)
	AR	61	13.93	(.57)	14.24	(.59)	14.59	(.73)	14.02	(.82)
	DD	11	19.09	(1.95)	–	–	19.74	(1.96)	–	–
Bladder control (daytime)	AR	25	–	–	34.56	(.96)	34.76	(1.28)	35.69	(3.04)
Interviewer's judgment of onset ^b	AR ^c	109	9.33	(.54)	–	–	11.88	(.79)	12.96	(.67)
	DD ^d	18	5.05	(1.34)	–	–	6.39	(1.70)	4.89	(1.65)

Ages reported in months; EM = Estimated Marginal Mean; SE = Standard Error.

^aDD > AR, $p < .001$.

^bDD < AR, $p < .01$.

^cT9 > T2 and T5 > T2, $p < .001$.

^dT9 < T2, $p < .05$.