

# Restricted and Repetitive Behaviors in Toddlers and Preschoolers with Autism Spectrum Disorders Based on the Autism Diagnostic Observation Schedule (ADOS)

So Hyun Kim and Catherine Lord

Restricted and repetitive behaviors (RRBs) observed during the Autism Diagnostic Observation Schedule [ADOS; Lord et al., 2000] were examined in a longitudinal data set of 455 toddlers and preschoolers (age 8–56 months) with clinical diagnosis of Autism Spectrum Disorders (ASD; autism,  $n = 121$  and pervasive developmental disorders—not otherwise specified (PDD-NOS),  $n = 71$ ), a nonspectrum disorder (NS;  $n = 90$ ), or typical development (TD;  $n = 173$ ). Even in the relatively brief semi-structured observations, GEE analyses of the *severity* and *prevalence* of RRBs differentiated children with ASD from those with NS and TD across all ages. RRB total scores on the ADOS were stable over time for children with ASD and NS; however, typically developing preschoolers showed lower RRB scores than typically developing toddlers. Nonverbal IQ (NVIQ) was more strongly related to the prevalence of RRBs in older children with PDD-NOS, NS, and TD than younger children under 2 years and those with autism. Item analyses revealed different relationships between individual items and NVIQ, age, diagnosis, and gender. These findings are discussed in terms of their implications for the etiology and treatment of RRBs as well as for the framework of ASD diagnostic criteria in future diagnostic systems.

**Keywords:** restricted and repetitive behaviors (RRBs); autism spectrum disorders (ASD); Autism Diagnostic Observation Schedule (ADOS); toddlers; preschoolers

## Introduction

Restricted and repetitive behaviors (RRBs) have long been considered one of the core characteristics of autism [Kanner, 1943]. According to the Diagnostic and Statistical Manual [DSM-IV; American Psychiatric Association, 1994], RRBs include a very broad category of behaviors such as preoccupation with one or more restricted patterns of interest (e.g. having very specific knowledge about vacuum cleaners), adherence to specific, nonfunctional routines (e.g. insisting on taking a certain route to school), repetitive motor manners (e.g. hand flapping), and preoccupation with parts of objects (e.g. peering at the wheels of toy cars while spinning them).

In the past, RRBs were thought to be rare in preschoolers or toddlers with autism [Charman & Baird, 2002; Stone, Lee, & Weiss, 1999; Ventola et al., 2006]. This assumption has been challenged in recent studies that reported the presence of RRBs in preschoolers, toddlers, and even infants as young as 8 months later diagnosed with autism [Richler, Bishop, Kleinke, & Lord, 2007; Watson et al., 2007]. However, at young ages, RRBs are not unique to children with autism spectrum

disorders (ASD) but are also present in children with nonspectrum disorders (NSs), such as intellectual disabilities and language disorders as well as in typical development (TD) [Evans et al., 1997; Sallustro & Constance, 1978; Thelen, 1979]. Even though RRBs are not unique to ASD, RRBs are described by caregivers as more prevalent and severe in very young children later diagnosed with ASD than in children later diagnosed with NS or found to be typically developing [Richler et al., 2007; Watson et al., 2007].

Examining early diagnostic differences in RRBs has important implications for the revision of the ASD criteria in future diagnostic systems (e.g. DSM-V, Internal Classification of Diseases-11). Given previous findings [Charman & Baird, 2002; Stone et al., 1999; Ventola et al., 2006], one question is whether RRBs should be deliberately excluded as a requirement for the diagnosis of ASD in very young children. According to the current DSM-IV [APA, 1994], Pervasive Developmental Disorders—Not Otherwise Specified (PDD-NOS), a milder form of ASD, includes children who have social deficits that are similar to those in autism and difficulties in either communication or RRBs or both (but at a mild level). Thus, in the

From the University of Michigan Autism and Communication Disorders Center (UMACC), Ann Arbor, Michigan (S.H.K., C.L.)

Received August 24, 2009; revised March 3, 2010; accepted for publication May 6, 2010

*Conflict of interest for Catherine Lord:* I receive royalties from a publisher for the instrument on which this study was based, but I did not receive royalties related to this study (proceeds were donated to charity).

Address for correspondence and reprints: So Hyun Kim, UMACC, 1111 Catherine St., Ann Arbor, MI 48109-2054. E-mail: sohkim@umich.edu

Grant sponsors: National Institute of Mental Health (NIMH); National Institute of Child Health and Human Development (NICHD); Department of Education (DOE); Simons Foundation; Grant numbers: R01 MH 066496; R01 MH081873; H324C030112.

Published online 29 June 2010 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.142

© 2010 International Society for Autism Research, Wiley Periodicals, Inc.

current system, children without any RRBs can be diagnosed with PDD-NOS. However, previous studies have found that RRBs consistently occurred more in young children with PDD-NOS as well as autism compared to children with NS or TD [Richler et al., 2007; Watson et al., 2007]. In the current proposal for DSM V (www.apa.org), the current distinction between autism and PDD-NOS will be eliminated, creating one general category of ASD, within which different dimensions (e.g. social-affect, RRBs) may be quantified. If this is the case, all children with ASD could be required to have evidence of some kind of RRBs.

Because results of the past literature in this area have been less clear, further investigation of the possibility of subgroups within ASD is needed. For example, children classified with PDD-NOS have been found to show less severe levels of RRBs than those with autism [Georgiades et al., 2007; Loh et al., 2007]. Yet research on PDD-NOS has suggested that there is no single behavior or factor that differentiates it from autism or Asperger syndrome. The way the terms, autism, Asperger syndrome, and PDD-NOS are used colloquially, it often appears as if these are discrete groups that can be clearly differentiated by the *presence* (vs. *absence*) of particular symptoms including the entire domain of RRBs. However, the evidence to date suggests that these subgroups within ASD share common symptoms, such as RRBs, with differences among them primarily lying in the *severity* of such symptoms.

Compared to a number of studies examining RRBs based on parent reports, only a few studies have examined RRBs using observational assessments. Because many observations occur within a brief time period, they may not provide an optimal opportunity for the assessment of RRBs. For this reason, Lord et al. [2000] initially excluded RRB items from the diagnostic algorithm in the Autism Diagnostic Observation Schedule—Generic (ADOS-G). However, because the inclusion of RRBs resulted in stronger predictive validity, recently revised algorithms for the ADOS-G do contain RRB items, [Gotham, Risi, Pickles, & Lord, 2007]. DiLavore, Lord, and Rutter [1995] also included RRB items in the Pre-Linguistic ADOS (PL-ADOS) diagnostic algorithm because RRB scores significantly increased the ability to classify young children with autism, developmental delays, and TD and to predict diagnostic outcome years later [Lord et al., 2006].

Research using parent reports of RRBs in young children showed most subtypes of RRBs, but not all, were stable or increased over time [Lord et al., 2006; Moore & Goodson, 2003; Richler, Huerta, Bishop, & Lord, 2010]. Using longitudinal data based on the RRB items in the Autism Diagnostic Interview—Revised [ADI-R; Rutter, LeCouteur, & Lord, 2003], Richler et al. [2010] found that RRBs in children with ASD either remained relatively high or increased over time from the age of 2–9. Lord

et al. [2006] also found that RRB scores at age 2 predicted RRB scores at age 9 for children with autism and that their RRBs were stable over time. In another study, Moore and Goodson [2003] found that ADI-R scores for *unusual preoccupations, compulsions and rituals, hand and finger mannerisms*, and *repetitive use of objects* increased between 2 and 4–5 years while scores for *complex mannerisms* decreased over the same time.

Other clinical features such as children's age and intellectual functioning and the interaction between those features have been found to be associated with RRBs measured by the ADI-R [Bishop, Richler, & Lord, 2006]. Bishop et al. [2006] found an interaction effect of age and intellectual levels on RRBs among children with ASD under 12 years old; NVIQ was more closely associated with the presence of RRBs at older ages compared to younger ages. More recently, *motor stereotypies* in individuals with ASD and intellectual disability were found to show less improvement over time compared to individuals with ASD with higher intellectual functioning [Esbensen, Seltzer, Lam, & Bodfish, 2009]. These findings raise the possibility that level of cognitive functioning is associated with changes over time in RRBs; however, these findings are based on parent reports of relatively older children and adults with ASD. Thus, longitudinal data based on toddlers and preschoolers provides important additional information as to whether this claim would apply to younger children.

Associations between RRBs and other clinical features of ASD can also vary depending on the specific RRBs. For instance, Turner [1999] suggested that “lower-order” RRBs (e.g. *unusual sensory interests*) are negatively correlated with IQ and “higher-order” RRBs (e.g. *compulsions/rituals*) are positively correlated with IQ. Using data obtained from children from 2 to 11 years old, Militerni et al. [2002] found that *sensory behaviors* were more prevalent in children with lower IQ scores, whereas *complex motoric sequences* were more prevalent in those with higher IQ scores. All of these findings highlight the importance of distinguishing different types of RRBs in relation to their associations with age and IQ scores.

Previous studies have also examined heterogeneity in RRBs using factor analyses of ADI-R items and found support for two different RRB factors—*repetitive sensory-motor behaviors* (RSMB; including items such as *hand/finger mannerisms, unusual sensory interests, repetitive use of objects, complex mannerisms*, etc.) and *insistence on sameness* (IS; *difficulties with change in routine, compulsions/rituals, unusual attachment to objects*, etc.) [Bishop et al., 2006; Cuccaro et al., 2003; Richler et al., 2010; Szatmari et al., 2006], or three factors—*circumscribed interests* in addition to the two factors previously mentioned [Lam, Bodfish, & Piven, 2008]. These factors differed in their associations with cognitive levels and age. For example, Richler et al. [2010] found that RSMB were associated

with NVIQ and stable over time, whereas IS was relatively independent of NVIQ and closely associated with age. It is important to note that none of these studies using a factor analytic approach have used observational data with toddlers and preschoolers, partly due to the limited number of items capturing RRBs in some available observational measures such as the ADOS. In addition, for ADOS items in modules appropriate to very young children, RRBs that have a quality of IS (e.g. lining things up or placing objects in a particular ways) are not distinguished from other repetitive activities (e.g. spinning a top; opening and closing a door) so that the distinction between RSMB and IS behaviors cannot be made. However, even though a factor analytic approach is not feasible with a limited number of items in certain measures, heterogeneity in RRBs can be examined through observing differences in the associations between individual RRB item scores and other clinical features such as age and IQ.

The present paper adds to the emerging literature on RRBs by contributing longitudinal data from toddlers and preschoolers (from 8 to 56 months old) with autism, PDD-NOS, NS, and TD using scores from the Autism Diagnostic Observation Schedule (ADOS). RRBs that are coded in the ADOS are presented in Table I. These items result in a single RRB total score on the ADOS diagnostic algorithms. Using the observational data, we hypothesized that:

- (1) RRBs would be significantly more prevalent in children with ASD (autism and PDD-NOS) than in children with NS and TD; the percentage of children who show at least one RRB will be higher in children with ASD than other groups.
- (2) RRBs will be greater in severity (defined by frequency, degree of impairment and ease of interrupting in the ADOS) for children with ASD than children with NS and TD; RRB

**Table I. Examples of RRBs on the ADOS**

Items	Examples
Stereotyped language <sup>a</sup>	Repeating lines of characters in Disney movies
Intonation of vocalizations <sup>a</sup>	Markedly flat, toneless "mechanical" vocalization
Sensory interests	Repetitively peering at objects from the side
Hand and finger mannerisms	Flicking or twisting fingers
Complex mannerisms	Spinning in circles
Repetitive behaviors	Lining up toy cars

Note. RRB, restricted and repetitive behaviors; ADOS, Autism Diagnostic Observation Schedule.

<sup>a</sup>Intonation of vocalizations was only included in the *no words* algorithm and was substituted for *stereotyped language* in the *some words* algorithm.

total scores will be higher for children with ASD than other groups.

- (3) Consistent with the literature using parent reports showing the predictability of RRBs in ASD, RRB totals in individual toddlers will predict RRB totals in preschool children.
- (4) Based on the past studies suggesting the stability of RRBs, RRB totals will be stable or increase over time.
- (5) Previous studies with relatively smaller numbers of younger children have indicated that the role of NVIQ in the manifestation of RRBs became stronger with increasing age; thus, it is predicted that the association between NVIQ and RRBs in older children (over or equal to 25 months) will be stronger than in younger children (under 25 months).
- (6) Individual item analyses will show heterogeneity in RRB items; the patterns of associations between individual RRB items and participant characteristics (e.g. NVIQ, age, diagnosis, and gender) will differ for different types of RRBs.

## Methods

### Participants

Data for this study were primarily obtained from three research projects, First Words and Toddlers (FW/T), Early Diagnosis of Autism (EDX), and Word Learning (WL) at the University of Michigan Autism and Communication Disorders Center (UMACC). Remaining participants were seen through clinic evaluations at UMACC. Children in the FW/T and WL projects entered the study at around 12 months of age and were assessed every several months or every 6 months (based on availability for repeated assessments primarily due to geographic locations) with the Toddler module of the ADOS [Lord, Luyster, Gotham, & Guthrie, 2010]. Children in the EDX project were assessed at ages 2 and 3 with the PL-ADOS [DiLavore et al., 1995], the previous version of the ADOS Module 1 and the ADOS-Toddler Module (ADOS-T).

Participants were 347 males (76%) and 108 females (24%). The sample consisted of 121 children with autism, 71 children with PDD-NOS, 90 children with NS, and 173 children with TD based on *best estimate diagnoses* (See *measures*). The NS group consisted of children with a history of developmental delay including language impairment and/or intellectual disability without a diagnosis of ASD. Their diagnoses were primarily expressive language delay, mild intellectual disability, and nonspecific developmental delay. With regard to ethnicity, 73% of children were Caucasian, 20% were

African American, 4% were multiracial, 1% were Asian, and 2% of parents did not provide information regarding ethnicity. Of the participants, 60% were recruited from the EDX study, 14% from the FW/T study, 25% from the WL study, and 1% from the clinic at UMACC. Of the 455 children, 394 children were assessed at least twice. As a result, 635 observations of children 8–56 months of age, with a mean age of 28.37 months (SD = 0.42) were included in the present study. Of 635 cases (which includes the multiple observations), 501 cases were males and 134 cases were females. In regards to each diagnostic group, the percentage of males ranged from 66 to 89% (TD, 66%; Autism, 86%; PDD-NOS, 89%; NS 73%). The ASD groups (autism and PDD-NOS) had significantly higher ratios of males to females than TD and NS groups ( $F = 11.7, P < 0.001$ ). Nonverbal IQ scores ranged from 13 to 155 with scores in the different diagnostic groups as follows: Autism,  $M = 62.94, SD = 1.32$ ; PDD-NOS,  $M = 72.5, SD = 2.12$ ; NS,  $M = 77.18, SD = 2.26$ ; TD,  $M = 113.11, SD = 1.3$ . No significant difference between males and females emerged for NIVQ scores; however, NVIQ scores differed significantly by diagnosis ( $F = 211.2, P < 0.001$ ). In addition, even when children with TD were excluded, there were significant age differences by diagnosis ( $F = 129.5, P < 0.001$ ), which was one of the reasons why we divided children into different cohorts (See *Age Cohorts*). Furthermore, males were significantly older than females (Males,  $M = 29.1, SD = 10.8$ ; Females,  $M = 25.2, SD = 10.2$ ;  $F = 14.807, P < 0.001$ ). Because there were significant diagnostic group differences for gender, NVIQ, and age, we controlled for these factors in all of the analyses as covariates.

#### Age Cohorts

Because we were interested in looking at effects of age on RRBs, when sample size permitted, toddlers and preschoolers were further divided into even smaller chronological age cohorts. Thus, there were a total 6 age cohorts for each

diagnostic group (Table II). There were 27 cases under 12 months old of age and all of them were in the TD group with mental ages over 12 months. Even though they were younger than 12 months, we included these children to take advantage of the available data given the limited number of children in the TD group. When statistics were performed, children under 25 months (cohort 1 and 2) in the autism and PDD-NOS groups and 43–56-month olds in the NS group were combined to obtain sample sizes large enough for the analyses. Each cohort included no more than one observation per child. Data for the TD group were only available from 8 to 30 months. NVIQ scores were also examined across different cohorts as follows: cohort 1,  $M = 85.63, SD = 20.12, range = 27–141$ ; Cohort 2,  $M = 82.31, SD = 18.17, range = 47–128$ ; cohort 3,  $M = 71.14, SD = 21.55, range = 13–123$ ; cohort 4,  $M = 64.71, SD = 20.91, range = 22–118$ ; cohort 5,  $M = 66.65, SD = 22.62, range = 17–132$ ; cohort 6,  $M = 59.19, SD = 21.57, range = 19–120$ . The NVIQ means were higher for younger cohorts than older cohorts even when typically developing children were excluded ( $F = 14.26, P < 0.001$ ) even though the standard deviations and ranges for NVIQ scores were similar across the cohorts. This was why we controlled for NVIQ in all of the analyses along with gender and age as covariates.

#### Procedure

A standard assessment battery was administered for each child after the IRB was approved and informed consent was obtained. The battery included the ADOS-T or PL-ADOS, the ADI-R, and the Mullen Scales of Early Learning [MSEL; Mullen, 1995] or the Bayley Scales of Infant Development [BSID; Bayley, 1993]. Testing was usually completed within a period of 2 weeks.

#### Measures

**The Autism Diagnostic Observational Schedule [Lord et al., 2000].** The ADOS is a semi-structured, standardized assessment of communication, social

**Table II. Number of Children (N), Mean Age (in Months) and Standard Deviation (SD) for Children in six Cohorts**

		Age in months					
		< 18 <sup>a</sup>	19–24	25–30	31–36	37–42	43–56
Autism	<i>N</i>	7	18	50	62	27	52
	Mean (SD)	16.43 (0.72)	22.06 (0.33)	27.5 (0.24)	33.18 (0.2)	39.48 (0.38)	45.63 (0.36)
PDD-NOS	<i>N</i>	6	16	29	30	12	27
	Mean (SD)	14.83 (0.7)	21.12 (0.53)	27.79 (0.33)	33.63 (0.31)	39.42 (0.47)	46.26 (0.52)
Non spectrum	<i>N</i>	20	16	27	34	8	15
	Mean (SD)	15.2 (0.52)	21.88 (0.51)	27.48 (0.28)	33.5 (0.31)	39.62 (0.63)	46.4 (0.68)
Typical development	<i>N</i>	102	62	15			
	Mean (SD)	13.51 (0.27)	21.34 (0.2)	27.4 (0.57)			

*Note.* 8–24-month olds for autism and PDD-NOS groups, and 37–56-month olds for NS groups were combined to obtain sample sizes large enough for statistical analyses; Children with typical development were assessed up to 30 months.

<sup>a</sup>< 18 = all children with ASD and NS were over 12 months; 27 children with TD under 12 months were included.

interaction, and play for children who may have ASD. Children in this study were administered either the PL-ADOS; DiLavore et al., 1995] or an experimental version intended for children under 36 months of age [ADOS-T; Lord et al., in press; Luyster et al., 2009]. A standardized diagnostic algorithm can be calculated for each version, and established cut-off scores based on algorithm totals are used to differentiate children with autism, ASD, and NS or TD. For the present study, the focus was on the domain of RRBs.

Scores on the ADOS-T and PL-ADOS items range from 0 to 3. A score of 0 indicates that the particular behavior is not present, and ratings from 1 to 3 vary in severity based on both frequency of the behaviors measured and their interference with other behaviors. A higher score indicates more severe abnormality. The revised algorithms of the ADOS-Module 1 are different for children who use *no words* or *some words* during the observation [Gotham et al., 2007].

For the RRB algorithm totals, we used the *no words* algorithm for 235 cases (37% of the entire cases) when the children used no or fewer than five spontaneous words/word approximations during the ADOS. The *some words* algorithm was used for the rest of cases, all of whom used more than five different words. Children who used simple phrases who had been given the PL-ADOS were also included here. The items in the algorithms were: *stereotyped language*, *intonation of vocalizations*, *sensory interests*, *hand and finger mannerisms*, *complex mannerisms*, and *repetitive behaviors* (see Table I for specific examples for each item). *Stereotyped language* and *intonation of vocalizations* were included in these algorithm subtotals on the basis of factor analyses that indicated they most often grouped with other RRBs [Luyster et al., 2009]. In both *no words* and *some words* algorithms, scores of 3 on the ADOS protocols are converted to 2, and the highest of score of either *hand and finger mannerisms* or *complex mannerisms* is selected and then combined with the other three items. The *no words* algorithm included all the items mentioned above except *stereotyped language*. The *some words* algorithm included all the items except *intonation of vocalizations*. As a result, the maximum score for the RRB total was 8 (when a child received a “2” on either *hand and finger mannerisms* and/or *complex mannerisms* and on each of the three other items) and the minimum was 0.

Ratings for two items in the ADOS-T were modified to match the same items on the PL-ADOS; for *sensory interests*, a score of 2 was converted to 1 because the score for a 2 in the ADOS-T was equivalent of a rating of 1 in the PL-ADOS item; in the same way, ratings of 1 and 2 were converted to 0 and 1, respectively, for *complex mannerisms*. The present study used the RRB totals from the ADOS Module 1 algorithms since we were able to maximize the number of identical items across the ADOS-T and PL-ADOS using the

Module 1 algorithms. However, because of a concern that *unusual intonation* and *stereotyped phrases* are not typically considered RRBs, though they had loaded onto factors with RRBs, analyses were also carried out on raw RRB total scores without these items. This also could address another concern of *no words* and *some words* algorithms having a slightly different composition of RRB totals when the algorithm RRB totals are used (because *stereotyped language* is substituted by *unusual intonation* in the *no words* algorithm) because the raw RRB totals consist of the same items across children who received *no words* and *some words* algorithms. For the raw RRB totals, we added the scores of *sensory interests*, *hand and finger mannerisms*, *complex mannerisms*, and *repetitive behaviors*, which ranged from 0 to 8.

**Mullen Scales of Early Learning [Mullen, 1995].** The MSEL is a developmental test intended for children from birth to 68 months. The MSEL was designed to assess children from birth through 68 months old. Because the MSEL does not provide separate VIQ and NVIQ scores, ratios of the two verbal subtests divided chronological age and multiplied by 100 and ratios of the two nonverbal subtests divided by chronological age and multiplied by 100 were used [Richler et al., 2007]. NVIQ scores were used as a measure of the child’s overall intellectual ability because they tend to be more stable over time among children with autism than verbal or full scale IQ scores [Howlin, Goode, Hutton, & Rutter, 2004]. The MSEL was administered to all children except 60 cases in the TD group seen as part of the EDX study.

**Bayley Scales of Infant Development [BSID, 1993].** The BSID was used for 60 out of 188 cases in the TD group. The BSID was designed to measure the developmental functioning of infants and toddlers. Because the BSID does not provide separate NVIQ and VIQ scores, full scale IQs on the mental scale were used as a proxy for NVIQ scores.

**Best estimate diagnosis.** For children in the FW/T project, all available data, including research diagnosis history over the most recent months and chart notes, were used by two examiners to generate consensus best estimate “working diagnoses.” The most weight was given to most recent diagnosis and “blind diagnoses” made by an examiner not familiar with the child [Luyster et al., 2009]. For children in the EDX study, an experienced clinical researcher used the ADOS and ADI-R and observations during the full assessment to generate independent best estimate diagnoses of autism, PDD-NOS, or a NS [Lord et al., 2006]. All of the examiners were trained to meet the standard requirements for research reliability on both instruments. Training involved 5 days of didactic lectures and hands-on practice in small groups with the new examiner then practicing partial and then full administrations until he or she was reliable in terms of exact agreement on 80% of items for three consecutive scorings including one scoring of a standard videotape

and one administration. Consensus coding was conducted approximately every fifth administration to ensure maintenance of reliability. All of the ADOS administrators were blind to the diagnostic status of the children at each evaluation except 66 children from the FW/T study. These children were seen, on an alternating basis, by a combination of a familiar clinician and a new clinician who was blind to their previous performance and tentative diagnosis [Luyster et al., 2009].

### Statistical Analyses

First, with SPSS 16.0, a Generalized Estimating Equations (GEE) logistic regression for repeated measures was used to examine differences in the prevalence of RRBs (percentage of children who showed at least one RRB indicated by getting a score of more than 0 on any of 6 items) among different diagnostic groups. Another GEE analysis was used to assess differences in the severity of RRBs (RRB total scores on the algorithms) among different diagnostic groups.

To examine the predictability of RRB totals, a hierarchical regression analysis was carried out using children's earlier RRB totals (from 8 to 30 months) as one of the predictors for the same children's later RRB totals (from 31 to 56 months). For this analysis, a subsample of 72 children at least a RRB total score of one for either cohort 1, 2, or 3 (Time 1) and one score for either cohort 4, 5, or 6 (Time 2) were selected. Step #1 included NVIQ scores, age, gender, and diagnosis at Time 2. Step #2 included RRB totals at Time 1 predicting RRB totals at Time 2. The stability of RRBs was also examined by performing another GEE analysis to see if RRB totals were independent of age for children with ASD.

To examine how the association between RRB totals and NVIQ varied by age, a GEE model was performed for each age cohort independently. For all of these GEE models, rate ratios (RR) were calculated for each covariate in the model. Additionally, we also performed separate analyses for each diagnostic group independently. For the TD group, a hierarchical regression analysis was performed because all data were cross-sectional, and for the

rest of the groups, GEE analyses were performed because of repeated measures.

Last, based on the hypothesis that relationships between the prevalence of RRBs and diagnosis, age, and NVIQ might differ for each type of RRB, a GEE logistic regression for each individual item score was performed, and odds ratios (OR) were calculated for each covariate in the model. One of the results from the item analyses yielded a significant association between RRBs and gender (See Results); therefore, gender was included as one of the covariates in all of the analyses mentioned above.

## Results

### Diagnostic Differences in the Prevalence of RRBs

Not surprisingly, a GEE logistic regression indicated that there was a significant main effect of diagnosis ( $\chi^2 = 53.34, P < 0.001$ ) on RRB prevalence while controlling for age, NVIQ, and gender. As shown in Table III, RRBs in children with autism and PDD-NOS were significantly more prevalent than those in children with NS and TD at all ages (pairwise comparisons with  $P < 0.05$ ). Almost all of children with ASD scored more than 0 on at least one of the RRB items. Prevalence rates of at least one RRB ranged from 96–100 and 90–97% by age cohorts for autism and PDD-NOS groups, respectively (Table III).

Even though RRBs in children with ASD were significantly more prevalent than in children with NS and TD, RRBs were relatively common in the NS group as well. When prevalence was compared across all of the diagnostic groups, children with NS showed positive scores (score of 1, 2, or 3 vs. 0) on one RRB item on average (mean ( $M$ ) = 1.01, standard deviation (SD) = 1.04). The TD group showed positive scores on fewer than one item ( $M = 0.71, SD = 0.81$ ). Children with autism and PDD-NOS showed positive scores on more than two RRB items on average (Autism,  $M = 2.83, SD = 1.02$ ; PDD-NOS,  $M = 2.15, SD = 1.17$ ).

**Table III. Percentage of Children who Showed RRBs by Age Cohorts and Diagnosis**

	Age in months					
	<18**	19–24	25–30	31–36	37–42	43–56
Autism %( $n/N$ )*	100(25/25) <sup>a</sup>		96(48/50) <sup>a</sup>	100(62/62) <sup>a</sup>	100(27/27) <sup>a</sup>	98(51/52) <sup>a</sup>
PDD-NOS	95(21/22) <sup>b</sup>		90(26/29) <sup>b</sup>	97(29/30) <sup>b</sup>	92(11/12) <sup>b</sup>	93(25/27) <sup>b</sup>
Non-spectrum	70(14/20) <sup>a,b</sup>	69(11/16) <sup>a,b</sup>	63(17/27) <sup>a,b</sup>	53(18/34) <sup>a,b</sup>		65(15/23) <sup>a,b</sup>
Typical development	63(64/102) <sup>a,b</sup>	53(33/62) <sup>a,b</sup>	27(4/15) <sup>a,b</sup>	–	–	–

\* $N$ , number of cases in each cell;  $n$ , number of cases with scores other than 0.

\*\* <18 = all children with ASD and NS were over 12 months; 27 children with TD under 12 months were included.

<sup>a</sup>Autism > NS and TD ( $P < 0.05$ ).

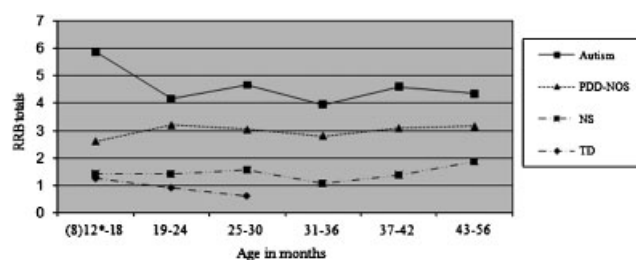
<sup>b</sup>PDD-NOS > NS and TD ( $P < 0.05$ ).

### Diagnostic Differences in the Severity of RRBs

When the severity of the RRBs in the four diagnostic groups was compared across age groups, the GEE analysis indicated that there was a significant main effect of diagnosis ( $\chi^2 = 97.24, P < 0.001$ ) while controlling for NVIQ, age, and gender. As hypothesized, consistent with the findings from the prevalence analyses, both groups with autism and PDD-NOS showed significantly higher RRB totals than children with NS and TD at all ages for which they were available (Fig. 1). When the raw RRB total scores without language items were used for these analyses, the results remained the same; there was a main effect of diagnosis ( $\chi^2 = 49.03, P < 0.001$ ) while controlling for NVIQ, age, and gender.

Because an interaction effect of diagnosis and age ( $\chi^2 = 38.8, P < 0.01$ ) emerged as well, we performed more analyses to examine the differences within the ASD groups (autism vs. PDD-NOS) for each age cohort. For this analysis, cohort 1 and 2 were combined to achieve sample sizes large enough for statistical analyses. Though the autism and PDD-NOS groups showed the biggest difference in RRB totals for cohort 1 (12–18 months) originally, when the difference was measured with cohort 1 and 2 (12–24 months) combined as one cohort, the difference was no longer significant. However, the differences in RRB totals between autism and PDD-NOS were significant for the remaining cohorts (25–56 months) confirming the age by diagnosis interaction (pairwise comparisons with  $P < 0.05$ ). However, when the raw totals were used without language items, the differences between these subgroups were no longer significant though the raw RRB totals for children with autism were still higher than those for children with PDD-NOS across age groups.

Across all ages, RRB totals for children with autism ( $M = 4.39; SD = 1.77$ ) were almost three times higher ( $RR = 2.93, P < 0.01$ ) than RRB totals for children with NS ( $M = 1.40, SD = 1.5$ ); RRB totals for children with PDD-NOS



**Figure 1.** RRB totals of the four diagnostic groups by age cohorts. *Note:* For statistical analyses, 8–24 month-olds for the autism and PDD-NOS groups and 37–56 months olds for the NS group were combined to obtain sample sizes large enough for the analyses (sample sizes indicated in Table II). (8)12\* = all children with ASD and NS were over 12 months; 27 children with TD under 12 months were included. The data represented in the graph were primarily longitudinal with a few cross sectional participants.

( $M = 3.03, SD = 1.84$ ) were more than two times higher ( $RR = 2.26, P < 0.01$ ) than RRB totals for children with NSs.

### The Predictability in the Severity of RRBs

As hypothesized, unstandardized regression coefficients indicated that RRB totals at Time 1 (12–30 months) were a significant predictor of RRB totals at Time 2 (31–56 months) for children with autism, PDD-NOS, and NS while controlling for NVIQ scores, diagnosis, age, and gender at Time 2 ( $\beta = 0.61, P < 0.001$ ). The model with all the predictors including the earlier RRB scores accounted for 46% of the variance in the later RRB totals ( $R^2_{adj} = 0.46, P < 0.001$ ). When the raw RRB totals without language items were used, results were identical.

### Stability in the Severity of RRBs

We expected that age would be independent of RRB scores because ADOS algorithm items were originally chosen specifically to distinguish ASD from other diagnoses at different language levels and age. As expected, the severity of RRBs for the autism and PDD-NOS groups was independent of age. RRB totals were similar for all ages in both groups, suggesting the stability of RRB scores over time. The NS group showed a similar pattern as well. However, age was significant for the TD group, indicating that RRBs in children with TD became less severe with increasing age ( $\beta = -2.81, P < 0.05$ ).

### The Association Between the Severity of RRBs and NVIQ Scores

As hypothesized, lower NVIQ scores predicted higher RRB scores for all of diagnostic groups combined across all ages ( $\chi^2 = 24.86, P < 0.01$ ). To examine how the association between RRBs and NVIQ differed by age, we performed separate analyses for three age groups: children under 25 months, 25–36 months, and 37–56 months. When the relationship between RRBs and NVIQ scores was examined by these groups, as predicted, the relationship was not significant for children under 25 months but it was for children over or equal to 25 months ( $\chi^2 = 7.55, P < 0.01$  for 25–36 months;  $\chi^2 = 13.32, P < 0.01$  for 37–56 months).

To our surprise, we also found that the impact of NVIQ on the RRB totals differed by diagnosis when each diagnostic group was analyzed separately. Interestingly, NVIQ was not a significant predictor of RRBs for children with autism at any age. However, NVIQ was a significant predictor of RRBs for children with PDD-NOS ( $\chi^2 = 6.53, P < 0.05$ ), NS ( $\chi^2 = 12.12, P < 0.01$ ), and TD ( $\beta = -0.17, P < 0.05$ ) over or equal to 25 months, but not for those under 25 months. Thus, children with higher NVIQ scores showed lower RRB scores for all children *except* for children with autism and children under 25 months in all diagnostic groups. In addition, the associations between

the raw RRB totals without language items and NVIQ scores by different age groups were consistent with the results mentioned above; raw RRB totals were not significantly related to NVIQ scores for the youngest cohorts, but were significantly related to NVIQ for the older cohorts ( $\chi^2 = 19.99, P < 0.001$  for those from 25 to 36 months,  $\chi^2 = 38.18, P < 0.001$  for those from 37 to 56 months). The associations between the raw RRB totals and NVIQ scores by diagnosis were nearly identical to the results using the algorithm RRB scores with significant associations emerging only between the raw RRB totals without language items and NVIQ scores for children with PDD-NOS ( $\chi^2 = 15.03, P < 0.001$ ), NS ( $\chi^2 = 11.41, P < 0.01$ ) and TD ( $\beta = -0.2, P < 0.05$ ) and not for children with autism.

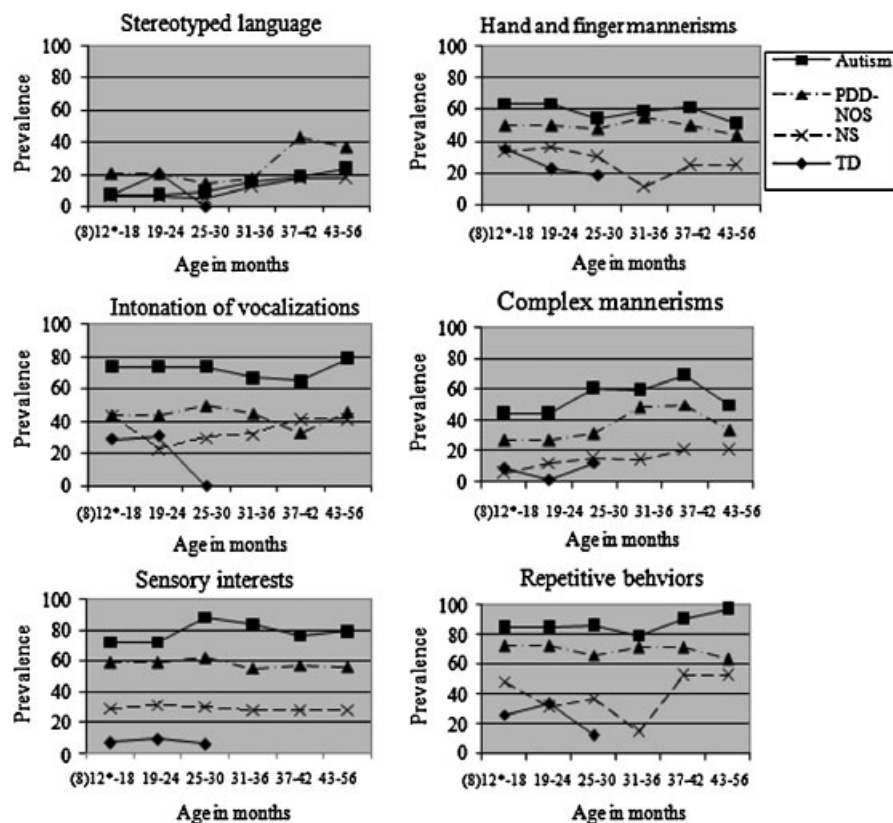
#### Individual Item Analyses in the RRB Domain

GEE logistic regression was run for each item separately to examine the heterogeneity in the associations between

the prevalence of each RRB item with NVIQ and age while controlling for diagnosis and gender. For the analysis of *stereotyped language*, only those children who received "Some Words" algorithms (thus, using phrase speech during the ADOS) were included.

Gender was a significant predictor for *repetitive behaviors*; girls were more likely to show these behaviors than boys. ( $OR = 0.7, P < 0.01$ ). No other gender effects or interactions with gender were significant.

As expected, there were differences in the associations between each RRB and age. Age was *not* associated with the prevalence of *intonation of vocalizations*, *sensory interests*, *hand and finger mannerisms*, and *complex mannerisms* across any diagnosis, suggesting stability over time (Fig. 2). However, age was a significant predictor of prevalence rates for *repetitive behaviors* and *stereotyped language*; the prevalence rates of these items for the children with autism, PDD-NOS, and NS increased as age increased, while the rates of the TD group decreased over time (Fig. 2 and Table IV).



Note. Prevalence rates for 8-24 month-olds in the autism and PDD-NOS groups and 37-56 month-olds for the NS group are the same since they were combined for the prevalence analyses (sample sizes indicated in Table 2); (8)12\* = all children with ASD and NS were over 12 months; 27 children with TD under 12 months were included. The data represented in the graph were primarily longitudinal with a few cross sectional participants.

**Figure 2.** Prevalence rates for each item by age cohorts. Note: Prevalence rates for 8-24 month-olds in the autism and PDD-NOS groups and 37-56 month-olds for the NS group are the same since they were combined for the prevalence analyses (sample sizes indicated in Table II); (8)12\* = all children with ASD and NS were over 12 months; 27 children with TD under 12 months were included. The data represented in the graph were primarily longitudinal with a few cross sectional participants.



**Table IV. Predictors for Prevalence of Individual RRB Items**

	Predictors	$\chi^2$	<i>P</i>
Stereotyped language <sup>a</sup>	Age	24.84	<0.001
	NVIQ scores	9.05	<0.01
Intonation of vocalizations <sup>b</sup>	Diagnosis	36.78	<0.001
	NVIQ scores	30.87	<0.001
Sensory interests <sup>c</sup>	Diagnosis	74.73	<0.001
	NVIQ scores	27.97	<0.001
Hand and finger mannerisms <sup>c</sup>	Diagnosis	27.97	<0.001
	NVIQ scores	17.85	<0.001
Complex mannerisms <sup>c</sup>	Diagnosis	39.2	<0.001
	NVIQ scores	13.61	<0.001
Repetitive behaviors <sup>c</sup>	Diagnosis	80.26	<0.001
	Gender	5.78	<0.05
	Age	12.82	<0.05

<sup>a</sup>Number of cases were 420.

<sup>b</sup>Number of cases were 235.

<sup>c</sup>Number of cases were 655.

As hypothesized, the association between RRBs and NVIQ also differed by types of RRBs. NVIQ was a significant predictor for *sensory interests*, *hand and finger mannerisms*, and *complex mannerisms* (Table IV). These three items were significantly more prevalent in children with lower NVIQ scores than children with higher NVIQ scores. For example, for each additional one point of NVIQ, the odds of showing *sensory interests* across cohorts diminished by 3% ( $OR = 0.97$ ,  $P < 0.01$ ). The prevalence rates of *repetitive behaviors* and *intonation of vocalizations* were independent of NVIQ. Even within the group of children who had phrases, *stereotyped language* was significantly more prevalent in children with higher NVIQ scores than in children with lower NVIQs.

Consistent with the results from RRB domain analyses, individual RRB item analyses showed that each type of RRBs for children with ASD were more prevalent than for children with NS and TD for all items except *stereotyped language*, which was independent of diagnosis (Table IV). There were diagnostic differences in the prevalence rates of all the other items with the autism group showing the highest rates consistently across all ages (Fig. 2). *Sensory interests*, *hand and finger mannerisms*, *complex mannerisms*, and *repetitive behaviors* were significantly more prevalent in children with ASD than children with NS and TD in all ages. For example, compared to the NS group, the autism group was nearly 13 times more likely ( $OR = 12.6$ ,  $P < 0.001$ ), and the children with PDD-NOS were about four times more likely to show *repetitive behaviors* ( $OR = 3.95$ ,  $P < 0.001$ ). The autism group was nearly 11 times more likely ( $OR = 10.8$ ,  $P < 0.001$ ), and the PDD-NOS group four times more likely than the NS group to show *sensory interests* ( $OR = 4.2$ ,  $P < 0.001$ ).

## Discussion

Using longitudinal data from brief observations, we were able to find stable diagnostic differences in both

*prevalence* and *severity* of RRBs measured during the ADOS among 655 observations of children (from 455 toddlers and preschoolers ranging from 8 to 56 months of age) with autism, PDD-NOS, NS and TD. The diagnostic differences in RRBs found in the current study suggest the importance of RRBs to early diagnosis of ASD, in line with previous studies that have shown similar results using parental interviews and questionnaires [Lord et al., 2006; Richler et al., 2007; Watson et al., 2007]. The results of this study also extend the findings of Morgan, Wetherby, and Barber [2008] who reported diagnostic differences in RRBs among children with ASD, developmental delays, and TD from 18 to 24 months based on systematic observations.

The present study adds to the growing body of literature showing that semi-structured observations which occur in a brief time period can successfully provide an opportunity for the assessment of RRBs. In fact, the diagnostic differences in RRBs found in the present study using the ADOS highlight the importance of using observational data with very young children because of evidence that parents may not notice RRBs in very young children [Chawarska, Klin, Paul, & Volkmar, 2007]. While parents might have more opportunities to observe RRBs in their children, they may also have more difficulties in judging the abnormality of such behaviors compared to clinicians. In fact, though our results indicated that most children with ASD showed RRBs during the ADOS, the prevalence analyses revealed that there were 11 cases in the PDD-NOS group who scored 0 on the RRB totals in the ADOS. However, it was found that all children with autism *and* those with PDD-NOS had at least one RRB at the time of assessment when RRB scores from both the ADI-R and ADOS, which were administered within the same week, were taken into account. All of the children with ASD who scored 0 on the RRB totals in the ADOS scored more than 0 on at least one of the *current* RRB item in the ADI-R.

Importantly, our findings with very young children suggest that when data from observations and parent interviews are combined, RRBs are almost always present in ASD, including in children given PDD-NOS diagnoses. This supports the idea of having RRBs as a requirement for the diagnosis of broader ASD (including PDD-NOS) in new diagnostic frameworks under development when information from parent report and observation are both taken into account.

The results of the present study on the differences in the *prevalence* and *severity* of RRBs between the ASD (autism and PDD-NOS) and NS groups are of practical significance. Even though RRBs in the ASD groups were consistently more prevalent and higher in severity than the NS and TD groups consistent with past studies [Bodfish, Symons, Parker, & Lewis, 2000; Richler et al., 2007; Watson et al., 2007], RRBs were relatively common

in the NS group. On average, children with NS showed at least one type of RRBs during an ADOS session while ASD groups showed 2–3 types of RRBs on average. Although current diagnostic guidelines, such as DSM-IV [APA, 1994], require the presence of a single RRB for a diagnosis of ASD, clinicians should be aware that the presence of any single RRB alone does not distinguish children with ASD from those with other NSs. Nevertheless, though the appearance of RRBs is not unique to ASD, particularly in very young children, the presence of a RRB coupled with social and communication deficits increases the likelihood of a stable diagnosis of ASD over years to come [Lord et al., 2006].

In the present study, the likelihood of having an RRB was the same for both autism and PDD-NOS groups; though the rated severity of these RRBs was higher in children over age 2 with autism than PDD-NOS diagnoses. This is consistent with the past literature that has shown that no single behavior or factor differentiated PDD-NOS from autism or Asperger's syndrome [Klin et al., 2005]. This is also in line with previous research using the ADI-R, indicating that individuals with autism had significantly higher scores than those with PDD-NOS in the *RSMB* factor [Georgiades et al., 2007]. These findings support the concept of a single category ASD differentiated by severity, rather than separate subgroups of ASD such as PDD-NOS or autism.

Consistent with previous findings on the trajectory of RRBs [Richler et al., 2010], the severity of RRBs was independent of age such that the RRB totals were stable over time in both ASD and NS groups. Even though these results were expected because ADOS items were originally selected to be differentially diagnostic at different language levels and ages, it is interesting that RRB totals for typically developing children were still associated with age such that their scores decreased over time, showing that differentiating children with ASD from typically developing children becomes easier as they get older. Early RRB scores in children with ASD and NS under 30 months predicted their later RRB scores indicating that RRBs that are severe in toddler years often remain severe over the course of development into preschool years.

The current study showed that IQ is more closely related to the manifestation of RRBs in older children with PDD-NOS, NS, and TD than very young or more severely affected children. Lower NVIQ scores were *not* associated with higher RRB scores in children with autism (regardless of their age) *and* children under 25 months (regardless of their diagnosis). This might be because of differences in the NVIQ distributions by cohort. In the present study, the older cohorts had the fewer children with lower NVIQ scores than younger cohorts, which might have minimized the effect of NVIQ scores on RRB totals for younger cohorts. However, another possibility is that older children

with higher IQs may have more interests and abilities that foster participation in less repetitive activities based on similar findings by Bishop et al. [2006]. In toddler years, these options may be limited since these very young children have not yet acquired the chance to develop creative play and more productive activities regardless of the level of their cognitive functioning. As children enter preschool years, they would be more likely to be exposed to novel environments and activities that would facilitate the development of more elaborate play. In the same way, since children with autism have more severe levels of impairment in their social and communication functioning compared to those with PDD-NOS, a milder form of ASD, they may have fewer opportunities to develop more productive, nonrepetitive activities through social interactions. These findings have important implications for treatment. One of the goals of early intervention may be to alter the course of developmental trajectories of RRBs by providing alternative behaviors that are equally motivating.

As we hypothesized, different patterns of association between the prevalence of RRBs and child characteristics were observed. Interestingly, three items, *sensory interests*, *hand and finger mannerisms*, and *complex mannerisms* showed similar patterns of associations with NVIQ, age, and diagnosis, distinct from the other RRB items. Surprisingly, the prevalence of these three items was associated with NVIQ but was stable over time, which is similar to findings from parent reports on *RSMB*) in past studies [Hus et al., 2007; Lam et al., 2008; Richler et al., 2010]. Turner [1999] described *RSMBs* as “low-order” behaviors related to developmental delays, which was also true for these three items that were more prevalent in children with lower NVIQ scores. Other items, *stereotyped language*, *intonation of vocalizations*, and *repetitive behaviors* differed in terms of their associations with NVIQ, age, diagnosis, and gender. *Repetitive behaviors* and *intonation of vocalizations* were independent of NVIQ, and *stereotyped language* and *repetitive behaviors* became more prevalent with increasing age, which suggests that these items in the ADOS might capture more “higher-order” behaviors [Turner, 1999]. This can be because, unlike in the ADI-R and more advanced Modules of the ADOS where the item, *repetitive behaviors*, is distinguished from *compulsions* and *rituals*, *repetitive behaviors* in the earlier Modules of the ADOS encompass a very broad range of behaviors including both *Repetitive Sensory-Motor* and *Insistence of Sameness* behaviors (e.g. repetitive nonfunctional use of toys; insistence on unusual routines). These results support the idea of different classes of RRBs and heterogeneity among different types of RRBs in associations with intellectual abilities and age [Turner, 1999]. The results also support that examining RRB subtypes and their relationships with other characteristics (e.g. NVIQ, age, gender, etc) can help us to better identify specific types of RRBs predictive of diagnosis. *Stereotyped*

*language* did not differentiate between diagnostic groups likely because of the still limited language levels of the children with ASD. A main effect of gender emerged for *repetitive behaviors*. However, since there was no significant effects of gender on NVIQ scores and on the other RRB subtypes as well as on the RRB total scores, the possibility of gender differences in subtypes of RRBs needs further exploration in future studies.

### Limitations/Future Directions

The present study assessed children with TD only up to 30 months in comparison to the other groups of children who were assessed up to 56 months. Had the TD group been followed at older ages, we would have been able to compare the rest of the groups to the TD group from age 31 to 56 months. Furthermore, children in the present study were divided into six cohorts, and the time interval for each cohort was about 5 months. Since rapid developmental changes occur during toddler and preschool years, it will be important for further studies to examine RRBs in those early years with shorter time intervals such as 1 or to 2 months to capture the rapid developmental changes in more detail.

One other limitation was that, in order to maintain a sufficiently large sample, we combined samples who received two different algorithms even though there was a slight difference in the composition of the RRB totals between these two algorithms (*intonation of vocalizations* was substituted for *stereotyped language* for the *no words* algorithm). In addition, one could argue that the language items can be considered not as central as the other items to the concept of RRBs. To address these concerns, we performed the same set of analyses using the raw RRB totals without language items for the *severity* of RRBs. However, when the raw totals were used, all of the results remained the same although the differences in the RRB raw totals between subgroups of ASD were not significant anymore. This confirms our belief that RRB algorithm totals can validly represent the severity of RRBs to test our hypotheses. On the other hand, it will be interesting for the future research to examine nonverbal and verbal samples separately if possible to investigate the role of the child's verbal level on RRBs in regards to items related to language.

Furthermore, not all aspects of RRBs can be assessed in brief observations. Distinctions between *IS* and other RRBs was not possible in these young because frequency, content or quality of a behavior were not coded sufficiently specifically. Parent and other caregiver (e.g. teacher, therapist) reports remain critical in order to capture broader aspects of RRBs. Semi-structured observations with very young children suspected of having ASDs can provide more information than even we originally

assumed, but still must be complemented by detailed information from people who know them well.

### Conclusions

This heterogeneity in RRBs found even in these very young children during relatively brief observations in the current study is consistent with past studies based on parent interview, suggesting researchers should attend to emerging differences among RRB items that could easily be missed by grouping multiple items under one single domain [Cuccaro et al., 2003; Lam et al., 2008; South, Ozonoff, & McMahon, 2005; Szatmari et al., 2006]. Further evidence about the heterogeneity in RRBs and their developmental trajectories may hold important clues for etiology, pathological mechanisms, and treatment of RRBs.

### Acknowledgments

We gratefully acknowledge the help of Somer Bishop, Kaite Gotham, Whitney Guthrie, Rhiannon Luyster, and Jen Richler as well as the families who participated in this research. We also acknowledge the supports from NRSA, NIMH, Autism Speaks, and Simons Foundation.

*Disclosure:* C. Lord receives royalties for the ADOS; profits related to this study were donated to charity.

### References

- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders, 4e. Washington, DC: American Psychiatric Association.
- Bayley, N. (1993). Bayley scales of infant development 2. San Antonio: The Psychological Cooperation.
- Bishop, S.L., Richler, J., & Lord, C. (2006). Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychology*, 12, 247–267.
- Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry*, 43, 289–305.
- Chawarska, K., Klin, A., Paul, R., & Volkmar, F. (2007). Autism spectrum disorder in the second year: Stability and change in syndrome expression. *Journal of Child Psychology and Psychiatry*, 48, 128–138.
- Cuccaro, M.L., Shao, Y., Grubber, J., Slifer, M., Wolpert, C.M., et al. (2003). Factor analysis of restricted and repetitive behaviors in autism using the Autism Diagnostic Interview-R. *Child Psychiatry and Human Development*, 34, 3–17.
- DiLavore, P., Lord, C., & Rutter, M. (1995) Pre-linguistic Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, 25, 355–379.
- Evans, D.W., Leckman, J.F., Carter, A.J., Reznick, S., Henshaw, D., et al. (1997). Rituals, habit, and perfectionism: The prevalence

- and development of compulsive-like behavior in normal young children. *Child Development*, 68, 58–68.
- Esbensen, A.J., Seltzer, M.M., Lam, K.S.L., & Bodfish, J.W. (2009). Age-related differences in restricted and repetitive behaviors in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 57–66.
- Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., et al. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of American Academy of Child Adolescent Psychiatry*, 46, 188–196.
- Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorder*, 37, 613–627.
- Howlin, P., Goode, S., Hutton, G., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, 45, 212–229.
- Hus, V., Pickles, A., Cook, E., Risi, S., & Lord, C. (2007). Using the Autism Diagnostic Interview—revised to increase phenotypic homogeneity in genetic studies of autism. *Biological Psychiatry*, 61, 438–448.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- Klin, A., Pauls, D., Schultz, R., & Volkmar, F. (2005). Three diagnostic approaches to Asperger syndrome: Implications for research. *Journal of Autism and Developmental Disorders*, 35, 221–234.
- Lam, K., Bodfish, J., & Piven, J. (2008). Evidence of three subtypes of repetitive behavior in autism that differ in familiarity and association with other symptoms. *Journal of Child Psychology and Psychiatry*, 49, 1193–1200.
- Loh, A., Soman, T., Brian, J., Bryson, S., Roberts, W., et al. (2007). Stereotyped motor behavior associated with autism in high-risk infants: A pilot videotape analysis of a sibling sample. *Journal of Autism and Developmental Disorder*, 37, 25–36.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., et al. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism & Developmental Disorders*, 30, 205–223.
- Lord, C., Risi, S., DiLavore, P., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, 63, 694–701.
- Lord, C.E., Luyster, R., Gotham, K., & Guthrie, W.J. (2010). Autism Diagnostic Observation Schedule—Toddler Module. Los Angeles, CA: Western Psychological Services.
- Luyster, R., Gotham, G., Guthrie, W., Coffing, M., Petrak, R., et al. (2009). The Autism Diagnostic Observation Schedule—Toddler Module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 1305–1320.
- Militeri, R., Bravaccio, C., Falco, C., Fico, C., & Palermo, M.T. (2002). Repetitive behaviors in autistic disorder. *European Child & Adolescent Psychiatry*, 11, 210–218.
- Moore, V., & Goodson, S. (2003). How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. *Autism*, 7, 47–63.
- Morgan, L., Wetherby, A., & Barber, A. (2008). Repetitive and stereotyped movements in children with autism spectrum disorders late in the second year of life. *Journal of Child Psychology and Psychiatry*, 49, 826–837.
- Mullen, E. (1995). The Mullen scales of early learning. Circle Pines, MN: American Guidance.
- Richler, J., Bishop, S.L., Kleinke, J.R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37, 73–85.
- Richler, J., Huerta, M., Bishop, S., & Lord, C. (2010). Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. *Development and Psychopathology*, 22, 55–69.
- Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism Diagnostic Interview—Revised—WPS, WPS ed. Los Angeles: Western Psychological Services.
- Sallustro, F., & Constance, W.A. (1978). Body rocking, head banging, and head rolling in normal children. *The Journal of Pediatrics*, 93, 704–708.
- South, M., Ozonoff, S., & McMahon, W.M. (2005). Repetitive behavior profiles in Asperger syndrome and high-functioning autism. *Journal of Autism and Developmental Disorders*, 35, 145–158.
- Stone, W.L., Lee, E.E., & Weiss, B.H. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychiatry*, 40, 219–226.
- Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., et al. (2006). Investigating the structure of the restricted, repetitive behaviors and interests domain of autism. *Journal of Child Psychology and Psychiatry*, 47, 582–590.
- Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behavior*, 27, 699–715.
- Turner, M. (1999). Annotation: Repetitive behavior in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry*, 40, 839–849.
- Ventola, P.E., Kleinman, J., Pandey, J., Barton, M., Allen, S., et al. (2006). Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders*, 36, 839–847.
- Watson, L.R., Baranek, G.T., Crais, E.J., Reznick, J.S., Dykstra, J., & Perryman, T. (2007). The first year inventory: Retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. *Journal of Autism and Developmental Disorders*, 37, 49–61.
- World Health Organization. (1990). International classification of diseases (10th revision). Geneva: World Health Organisation.