Research in Autism Spectrum Disorders

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Relationship between executive functions and motor stereotypies in children with Autistic Disorder

Brittany C. LeMond a, Roee Holtzer a,b, Sylvie Goldman b,c,*

a Ferkauf Graduate School of Psychology, Yeshiva University, 1165 Morris Park Road, Bronx, NY 10461, United States
b Department of Neurology, Albert Einstein College of Medicine, Jack and Pearl Resnick Campus, Yeshiva University, 1300 Morris Park Avenue, Bronx, NY 10461, United States
c Department of Pediatrics, Albert Einstein College of Medicine, Jack and Pearl Resnick Campus, Yeshiva University, 1300 Morris Park Avenue, Bronx, NY 10461, United States

ARTICLE INFO

Article history:
Received 29 February 2012
Accepted 1 March 2012

Keywords:
Executive functions
Videotape coding
Motor stereotypies
Developmental disabilities
Autism disorder

ABSTRACT

This study reports on the relationship between motor stereotypies and impairments in executive functions (EF) in children with Autistic Disorder (AD) and in children with Developmental Language Disorders (DLD). We hypothesized that low EF performance would predict higher frequency and longer durations of stereotypies in the AD group only. Twenty-two children (age range = 7–9 years, 6 months, girls = 5) with AD were recruited from a longitudinal multi-site study and compared to twenty-two non-autistic children with DLD (age range = 7–9 years, 6 months, girls = 5). The two groups were matched on non-verbal IQ and demographic characteristics. Frequency and duration of stereotypies were coded from videotaped semi-structured play sessions. EF measures included the Wisconsin Card Sorting Task (WCST) Categories, Wechsler Intelligence Scale for Children-Revised (WISC-R) Mazes, and Stanford-Binet Fourth Edition (SB-IV) Matrices. The scores for frequency and duration of stereotypies were higher in the AD group. Separate linear regressions revealed that group status, EF, and their interactions predict stereotypies. Specifically, lower EF scores predicted higher frequencies and longer durations of stereotypies in the AD group only. Analyses controlled for age, gender, and parent education. Findings suggest that in AD, EF impairments and stereotypies may be linked to shared brain pathways.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Restricted and repetitive behaviors (RRBs) represent the third core criteria for a diagnosis of Autistic Disorder (AD) in the DSM-IV-TR (American Psychiatric Association [DSM-IV-TR], 2000). RRBs are present in children and adults with Autism Spectrum Disorder (ASD) at higher rates than in any other developmental disability (Matson, Dempsey, & Fodstad, 2009; Morgan, Wetherby, & Barber, 2008) and hamper normal development, learning, and social adaption (Leekam, Prior, & Uljarevic, 2011).

Among the different subtypes of RRBs, motor stereotypies are defined as patterned, repetitive, and purposeless movements (e.g., hand flapping, finger flicking, rocking). A recent review of the assessment and treatment of motor
stereotypic behaviors in children with autism and other pervasive developmental disabilities, found that 25 unique categories of stereotypies have been identified in the literature (Di Gennaro Reed, Hirst, & Hyman, 2012). In ASD, motor stereotypies are prevalent, enduring, and present in all age groups (Bishop, Richler, & Lord, 2006; Bodfish, Symons, Parker, & Lewis, 2000; Richler, Huerta, Bishop, & Lord, 2010; Zandt, Prior, & Kyrios, 2007). Additionally, infant and high-risk sibling studies suggest that stereotypies represent valid early markers of ASD (Baranek, 1999; Matson, Wilkins, et al., 2009; Wetherby et al., 2004).

Emerging evidence has demonstrated that motor stereotypies occur not only in children with ASD, and other developmental disabilities but also in normally developing children and adolescents (Barry, Baird, Lascelles, Bunton, & Hedderly, 2011; Harris, Mahone, & Singer, 2008; Mahone, Bridges, Frahmme, & Singer, 2004). The nature and developmental course of stereotypies among individuals with ASD is not well understood (Symons, Sperry, Dropik, & Bodfish, 2005). Moreover, observational research has revealed that motor stereotypies may be more prevalent during times of emotional or social demand (e.g., times of excitement or stress, situations with increased social engagement) and in times of boredom (Schlaggar & Mink, 2003).

Repetitive motor behaviors have not only been associated with the severity of ASD diagnosis (Bodfish et al., 2000), but also with deficits in executive functions (EF) (Liss et al., 2001; South, Ozonoﬀ, & McMahon, 2007). EF are deﬁned as higher order cognitive processes that include planning, initiating and inhibiting actions, selecting relevant sensory information from one’s surroundings and general cognitive ﬂexibility (e.g., Shallice & Burgess, 1991). The Pre-Frontal Cortex (PFC) and frontal-basal ganglia circuits have been identiﬁed as the core brain substrates associated with EF (Stuss, 1992; Stuss & Alexander, 2000).

Poor EF performance has been found in individuals with ASD on a variety of tasks including, those involving inhibition of prepotent responses (Stroop Test), planning (Tower of London), generativity (Verbal Fluency Test), cognitive ﬂexibility (Wisconsin Card Sorting Test), and memory (Bennetto, Pennington, & Rogers, 1996; Robinson, Goddard, Dritschel, Wisley, & Howlin, 2009). Impairments on EF tasks are thought to directly relate to one’s ability to respond, plan, and adapt behaviors based on environmental and social cues (Ridley, 1994). The executive dysfunction hypothesis, which states that the types of symptoms present in ASD are due to deﬁcits in the executive control of behaviors, has been linked to the prevalence of RRBs seen in ASD (Hughes, Russell, & Robbins, 1994; Sayers, Oliver, Ruddick, & Wallis, 2011).

Only within the past decade have subtypes of RRBs in ASD been examined systematically (Leekam et al., 2011; Richler et al., 2010). One identiﬁable subtype is motor stereotypies, which are typically considered lower-level repetitive behaviors and usually associated with younger and lower functioning children, while restricted interests, nonfunctional routines, and rituals are considered higher-level repetitive behaviors, and are typically associated with older and higher-functioning individuals (Leekam et al., 2011; Turner, 1999).

Most validated diagnostic measures for ASD contain subsections devoted to assessing RRBs, such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) the Autism Observation Scale for Infants (AOI; Bryson, Zwaiangen, McDermott, Rombough, & Brian, 2008), the Baby and Infant Screen for Children with Autism Traits (BISCUIT; Matson, Wilkins, et al., 2009), and the Communication and Symbolic Behavior Scales Developmental Proﬁle (CSBS–DP; Wetherby, Allen, Cleary, Kubin, & Goldstein, 2002). Other widely used measures include questionnaires, such as the Repetitive Behavior Scale–Revised (RBS-R; Lam & Aman, 2007) and paper–and–pencil rating scales, which typically capture global impressions of stereotypy frequency and severity, which may be subjective and biased (Pyles, Riordan, & Bailey, 1997).

While one-time observational measures aid in capturing the breadth of RRBs in autism, several factors limit their reliability, such as (1) a reduced ability to accurately document rapid stereotypies, (2) diﬃcultly observing simultaneous motor movements, and (3) a limited capacity to identify when movements begin and end (Sprague & Newell, 1996). In contrast, the coding of videotapes allows for systematic observation and assessment of repetitive movements, so as to offer better characterization, a capability to re-watch rapid and concomitant movements, as well as the power to determine duration of these movements.

A recent study using alternative methods to precisely measure motor stereotypies utilized accelerometers and pattern recognition algorithms to automatically detect these movements in children with ASD (Goodwin, Intille, Albiniali, & Velicer, 2011). Findings revealed that approximately 90% of the time these automated methods were able to correctly identify stereotypical movements. Videotape recordings and automated detection of motor stereotypies may allow for a better classiﬁcation and understanding of this RRB subtype in ASD.

Motor stereotypies, a subtype of RRBs, are thought to be overt and acute forms of perseveration, which may be related to a lack of generativity (Turner, 1999). Based on this view, we hypothesize that deﬁcits in EF in autism may underlie motor stereotypies seen in this population. In line with this expectation, positive relationships between the prevalence of RRBs and deﬁcits of speciﬁc EF tasks, such as response inhibition, generativity, ﬂexibility, and working memory, in individuals with ASD but not in matched-controls (Lopez, Lincoln, Ozonoﬀ, & Lai, 2005; South et al., 2007). Similarly, a recent study using self-report measures of adolescents with ASD, found greater stereotypic behaviors and restricted interests scores and poorer EF performance, especially in boys (Bo¨lte, Duketis, Poustka, & Holtmann, 2011). Another study found that executive deﬁcits, as measured by an informant-based rating scale, were correlated with repetitive behaviors, assessed by the RBS-R in children with high functioning autism and healthy, typical controls (Boyd, McBee, Holtzclaw, Baranek, & Bodﬁsh, 2009). Findings also revealed that ASD diagnosis, younger age, sensory processing issues, and behavioral regulation deﬁcits predicted more repetitive behaviors (Boyd et al., 2009).

On the contrary, other studies have failed to ﬁnd any correlation between RRBs and several common neuropsychological EF tasks in school-aged children with ASD (Joseph & Tager-Flusberg, 2004), and another study did not ﬁnd an association between RRBs, measured by the ADOS (Lord et al., 2000) and the ADI-R (Lord, Rutter, Le Couteur, 1994), and inhibition and
working memory (Ozonoff et al., 2004). The above divergent results on the association between EF and RRBs may stem from the lack of focus on the distinct subtypes of repetitive behaviors in ASD such as motor behaviors versus rituals.

Accordingly, we sought to revisit the question of the relationship of RRBs and EF deficits. Specifically, we examined repetitive motor behaviors (i.e., stereotypies) in a group of children diagnosed with Autistic Disorder (AD) and a matched control group of non-autistic children with Developmental Language Disorders (DLD). We used quantitative objective measures of the frequency and duration of motor stereotypies collected during videotaped semi-structured play sessions. In regard to EF, we selected three tasks from the fixed research battery of neuropsychological tests, WCST Categories, WISC-R Mazes, and SB-IV Matrices. These tasks measure cognitive domains such as: inhibition, planning, and mental flexibility, as well as perseverations, which are linked to RRBs in autism.

We hypothesized greater frequencies and longer durations of stereotypies in the AD group compared to controls, replicating prior work on RRBs and autism. We expected EF measures to be related to stereotypy frequency and duration; however, AD diagnosis was expected to moderate the EF-stereotypy association such that it would be stronger (or only evidenced) in the AD group compared to a control group of children with DLD. This work advances on the literature in two key ways: (1) we utilize video-recorded quantitative measures of motor stereotypies rather than broad measures of RRBs that collapse routines, rituals, and motor behaviors, and (2) we focus on a narrow age range of children at a time in development when motor stereotypes are prevalent.

2. Methods

2.1. Participants

Children were selected from a longitudinal, multi-centric (Boston MA, the Bronx NY, Cleveland OH, and Trenton, NJ), nosological study conducted between 1985 and 1992 involving comprehensive behavioral, neurological, neuropsychological, and psychiatric evaluations that placed them into one of four diagnostic groups based on cognitive level and autism diagnosis (Rapin, 1996). The children were assessed at age three by trained psychiatrists according to the DSM-III-R (American Psychiatric Association [DSM-III-R], 1987) and the diagnosis of AD was validated with parents’ response to the Wing Autistic Disorder Interview Checklist (WADIC; Wing, 1985). Children diagnosed with AD in the present study according to the DSM-III-R criteria would also meet, with no exception, current diagnostic criteria for AD. Children who did not reach DSM-III-R criteria for AD were screened for Developmental Language Disorder (DLD) based on the Test of Early Language Development (TELD; Hresko, Reid, & Hammil, 1981) and the Sequenced Inventory of Communication Development, Revised Edition (Hendrick, Prather, & Tobin, 1984).

Twenty-two, seven-to-nine year old children with AD were recruited from the larger study and matched to a group of children with DLD on age, gender, Hollingshead parent education (Hollingshead, 1975), and nonverbal IQ based on the Stanford-Binet Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986). Children under psychotropic medication and/or with an overall IQ < 70 were excluded. Written consent was obtained from all children's parents and the study was approved by the Institutional Review Board. More detailed information regarding diagnostic and testing procedures for this cohort has been previously published (Rapin, 1996).

2.2. Measures of executive functions

Children’s cognitive functioning was assessed based on a standardized battery of neuropsychological tests. For the purposes of the present study, we focused on tests of EF, including, the Wisconsin Card Sorting Task (WCST), Wechsler Intelligence Scale for Children-Revised (WISC-R) Mazes, and Stanford-Binet Fourth Edition (SB-IV) Matrices. The Wisconsin Card Sorting Task (WCST; Heaton, Chelune, Talley, Kay, & Curtis, 1993) involves sorting cards based on number, shape, and color and utilizing problem solving strategies when provided with minimal feedback from the examiner. The WCST measures mental flexibility, inhibition, and perseveration (Heaton et al., 1993). In children and adolescents, test–retest reliability for the WCST ranges from .37 to .72 (Heaton et al., 1993; Ozonoff, 1995). In the present study, number of categories achieved was used to assess performance on this measure.

The Mazes subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) is a paper and pencil–timed task that involves completing a series of nine complex mazes of increasing difficulty. WISC-R Mazes involves abilities such as visuo–spatial processing and trial-and-error learning (Wechsler, 1974). Test–retest reliabilities for the Mazes subtest for ages 7½, 8½ and 9½ were .81, .77 and .71, respectively (Wechsler, 1974). The correlations of the WISC-R with other intelligence scales (e.g., WPPSI, SB-IV) range from .66 to .94 (Wechsler, 1974) and are indicative of good external validity. In the present study, the scaled score of the Mazes subtest was used to assess performance.

The Matrices subtest of the Stanford-Binet Fourth Edition (SB-IV; Thorndike et al., 1986) requires a subject to select a picture from a set of options that would complete a matrix puzzle. SB-IV Matrices measures EF, such as planning, inhibition, adaptive reasoning skills, and perseveration (Thorndike et al., 1986). The complexity of the puzzles is gradually increased for a total of 18 trials. The Matrices subtest has an internal consistency of .90 for ages 7, 8, and 9 (Thorndike et al., 1986). The SB-IV has strong external validity and associations with intelligence scales (e.g., Wechsler Intelligence Scale for Children-Revised) with correlations ranging from .27 to .91 (Thorndike et al., 1986). In the present study, T-scores of the Matrices subtest were used to assess performance.
2.3. Measure of stereotypies

Stereotypies were coded from 30-min videotaped semi-structured play sessions. At the beginning of each session a trained experimenter instructed the child to complete on his/her own a complex three-level Plexiglas maze, wherein the child had to utilize different size tools to get prizes through the maze. The children had to incorporate different problem-solving strategies in order to complete the task. After explaining the task, the examiner sat away from the child, to allow the child to solve the problem alone and rejoined the child after 5 min. Throughout the remainder of the session, the experimenter followed a structured play protocol. The toys were age-appropriate and geared towards symbolic play (e.g., dollhouse, telephone) and problem-solving (e.g., Plexiglass maze with tools). The amount of time the child and experimenter played with each toy varied slightly, depending on the child’s social abilities, cognitive functioning, and play skills. All sessions were videotaped from an adjacent observation room through a one-way mirror. Sound was recorded through a microphone hanging from the ceiling inside the testing room.

In the present study, the first 10 min of the child and experimenter interaction was coded. We quantified frequency (total number of episodes) and duration (seconds) of each stereotypy episode. A stereotypy episode was defined as an isolated and discrete period of a continuous repetitive movement or stereotypy (i.e., purposeless rhythmic patterned repetitive movement). A sensitive DVD computer software program (DVD Player Version 5.4, Manufactured by Dolby Laboratories) was used to code the videos. Coders were able to start, stop, and slow the videos as needed, so as to carefully score in minutes and seconds the beginning and end of each stereotypy episode in order to calculate its duration.

The first author was trained by an experienced psychologist with expertise in autism and movement disorders. The first author then trained two raters, who were blind to the children’s diagnosis. Both blind raters were trained using specific parameters outlined and coded videos of children not part of the present study alongside the first author until 100% agreement was reached for both frequency and duration of stereotypies. Once this level of agreement was attained, the first and second raters independently coded frequency and duration of stereotypy episodes from 15 randomly chosen videos (nine AD, six DLD). Both simple correlations and Intraclass Correlation Coefficient (ICC) analyses were used to test agreement between the two raters on the 15 videos. ICC estimates the proportion of variance in the data due to differences in the subjects, rather than differences in the raters. Cohen’s Kappa was also calculated to determine inter-rater reliability for the presence of stereotypy episodes.

The correlations between raters were high for both frequency (r = .68, p < .01) and duration (r = .78, p < .01) of stereotypy episodes. The ICCs were also high for both the frequency (ρI = .75, p < .001) and duration (ρI = .86, p < .001) of stereotypy episodes. The inter-rater reliability for the two raters as determined by Cohen’s Kappa was also very high (κ = 0.88, p < .001, 95% CI = 0.78–0.98).

2.4. Data analysis

A total of six separate hierarchical regressions were conducted, each with two blocks, to assess whether AD diagnosis, EF and their interactions predicted the duration and frequency of stereotypies. In block one age, gender, parent education, and individual EF task scores (WCST Categories, WISC-R Mazes, and SB-IV Matrices) and AD diagnosis served as predictors. Within the first block the predictors were entered simultaneously. Hence, the regression coefficients of the EF tasks were adjusted for the covariates and group status. Separate regressions were conducted for each of the EF measures due to their interaction term. The frequency and duration of stereotypy episodes served as the outcome measures in separate regression models. In the six linear regressions all vital assumptions were met. For all the statistical tests the null hypothesis was rejected at the 0.05 level.

3. Results

The AD and DLD groups were comparable for key demographic characteristics (Table 1).

As expected, the frequency (3.82, ±2.50) and duration (8.41, ±4.15) of stereotypy episodes in the AD group were significantly greater than the frequency (2.50, ±1.78) and duration (1.17, ±0.48) of stereotypy episodes in the DLD control group. No group differences were found on the EF tasks.

The key demographic variables, AD diagnosis, the three EF tasks (WCST Categories, WISC-R Mazes, and SB-IV Matrices) and the AD diagnosis-by-EF task interaction terms were used as predictors in six separate regression analyses, with stereotypy frequency and stereotypy duration serving as the dependent measures. Summary of the results are presented in Tables 2 and 3.

In block one, the key demographic characteristics, AD diagnosis, and EF task (WCST Categories, WISC-R Mazes, or SB-IV Matrices) were predictors of stereotypy frequency explaining 32, 38, and 33% of the variance, respectively (see Table 2). In
block one, the key demographic characteristics, AD diagnosis, and EF task (WCST Categories, WISC-R Mazes, or SB-IV Matrices) were also significant predictors of stereotypy duration explaining 33, 40, and 33% of the variance, respectively (see Table 2). The addition of the AD-by-WCST Categories and the AD-by-SB-IV Matrices interaction terms in the second block accounted for an additional incremental increase of approximately 11 and 13% in the amount of the variance explained in

Table 1
Demographic information, mean EF scores and stereotypy frequency and duration in the Autistic Disorder (AD) and Matched Developmental Language Disorder (DLD) groups.

<table>
<thead>
<tr>
<th>Variable</th>
<th>AD (n = 22)</th>
<th>DLD (n = 22)</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>%ile</td>
<td>Range</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>8.09 (1.06)</td>
<td>–</td>
<td>7.00–9.50</td>
<td>8.04 (1.01)</td>
</tr>
<tr>
<td>Father education</td>
<td>15.54 (3.45)</td>
<td>–</td>
<td>12–19</td>
<td>15.72 (4.02)</td>
</tr>
<tr>
<td>Mother education</td>
<td>13.32 (2.55)</td>
<td>–</td>
<td>10–18</td>
<td>13.49 (2.99)</td>
</tr>
<tr>
<td>Duration (seconds)</td>
<td>8.41 (4.15)</td>
<td>–</td>
<td>0–43</td>
<td>1.17 (0.48)</td>
</tr>
<tr>
<td>Frequency</td>
<td>3.82 (2.50)</td>
<td>–</td>
<td>0–11</td>
<td>2.50 (1.78)</td>
</tr>
<tr>
<td>SB-IV NVIQ</td>
<td>97.85 (14.39)</td>
<td>39</td>
<td>72–122</td>
<td>98.88 (16.02)</td>
</tr>
<tr>
<td>WCST Categories</td>
<td>2.96 (1.66)</td>
<td>16</td>
<td>0–6</td>
<td>3.35 (2.17)</td>
</tr>
<tr>
<td>WISC-R Mazes</td>
<td>9.35 (3.85)</td>
<td>37</td>
<td>1–16</td>
<td>9.84 (2.60)</td>
</tr>
<tr>
<td>SB-IV Matrices</td>
<td>50.78 (6.04)</td>
<td>50</td>
<td>20–42</td>
<td>50.46 (7.63)</td>
</tr>
</tbody>
</table>

Note: WCST, Wisconsin Card Sorting Task. WISC-R, Wechsler Intelligence Scale for Children-Revised. SB-IV NVIQ, Stanford-Binet Fourth Edition Nonverbal IQ. T-tests for independent samples (df = 42) were used to determine differences between the AD and DLD groups.

Table 2
Summary of regression analyses predicting frequency and duration in AD and DLD children.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>R²</td>
</tr>
<tr>
<td>R</td>
<td>R²</td>
</tr>
</tbody>
</table>

Model 1
Block 1
WCST Categories: –0.25, 0.22, –0.19, 0.251, –1.15, 0.61, –0.26, 0.066
AD diagnosis: 1.55, 0.90, 0.32, 0.059, 5.88, 2.27, 0.35, 0.013
Block 2
AD × WCST Categories: –1.06, 0.41, –0.48, 0.013, –3.51, 1.17, –0.48, 0.004

Model 2
Block 1
WISC-R Mazes: –0.24, 0.11, –0.34, 0.030, –0.99, 0.37, –0.41, 0.010
AD diagnosis: 1.67, 0.75, 0.32, 0.032, 6.66, 2.54, 0.37, 0.012
Block 2
AD × WISC-R Mazes: –0.22, 0.23, –0.26, 0.356, –1.56, 0.75, –0.55, 0.046

Model 3
Block 1
SB-IV Matrices: –0.08, 0.06, –0.21, 0.189, –0.34, 0.20, –0.26, 0.099
AD diagnosis: 1.70, 0.78, 0.32, 0.036, 6.77, 2.69, 0.38, 0.016
Block 2
AD × SB-IV Matrices: –0.29, 0.10, –0.47, 0.006, –1.24, 0.32, –0.59, <.001


Table 3
Fully adjusted models using AD, EF task, and AD-by-EF task interaction terms as predictors of frequency and duration of stereotypy.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>SE b</td>
</tr>
<tr>
<td>b</td>
<td>SE b</td>
</tr>
</tbody>
</table>

Model 1
Block 1
WCST Categories: –0.25, 0.22, –0.19, 0.251, –1.15, 0.61, –0.26, 0.066
AD diagnosis: 1.55, 0.90, 0.32, 0.059, 5.88, 2.27, 0.35, 0.013
Block 2
AD × WCST Categories: –1.06, 0.41, –0.48, 0.013, –3.51, 1.17, –0.48, 0.004

Model 2
Block 1
WISC-R Mazes: –0.24, 0.11, –0.34, 0.030, –0.99, 0.37, –0.41, 0.010
AD diagnosis: 1.67, 0.75, 0.32, 0.032, 6.66, 2.54, 0.37, 0.012
Block 2
AD × WISC-R Mazes: –0.22, 0.23, –0.26, 0.356, –1.56, 0.75, –0.55, 0.046

Model 3
Block 1
SB-IV Matrices: –0.08, 0.06, –0.21, 0.189, –0.34, 0.20, –0.26, 0.099
AD diagnosis: 1.70, 0.78, 0.32, 0.036, 6.77, 2.69, 0.38, 0.016
Block 2
AD × SB-IV Matrices: –0.29, 0.10, –0.47, 0.006, –1.24, 0.32, –0.59, <.001

Note: AD, Autistic Disorder; DLD, Developmental Language Disorder. Analysis adjusted for age, gender, and parent education.
stereotypy frequency (see Table 2). On the contrary, the addition of the AD-by-WISC-R Mazes interaction term in block two did not add to the amount of variance explained by the model in predicting stereotypy frequency. The addition of the AD-by-EF Task interaction terms (WCST Categories, WISC-R Mazes, and SB-IV Matrices) in the second block accounted for an additional incremental increase of approximately 12, 7, and 20% of the variance explained in stereotypy duration, respectively (see Table 2).

Table 3 reveals that in block one AD diagnosis and EF task performance were related to stereotypy frequency and duration. The interaction terms were related to stereotypy duration and frequency. The standardized β coefficients provide an estimate of the strength of the relationship between the AD-by-EF task interaction terms and stereotypy duration and frequency. Following inspection of the coefficient size, all three measures of EF had comparable strong associations with stereotypy duration. With respect to frequency, the coefficients of the AD-by-WCST Categories and AD-by-SB-IV Matrices factors were also comparable. However, the AD-by-WISC-R Mazes interaction did not significantly contribute to the prediction of stereotypy frequency.

4. Discussion

This study was designed to evaluate the relationship between frequency and duration of motor stereotypies and executive functions in 22 children with AD and 22 matched control children with DLD. The findings are consistent with our hypothesis, stating that lower EF performance on tasks of inhibition, planning, and mental flexibility would predict higher frequency and longer durations of stereotypies only in the AD group, even when adjusting for age, gender, and parent education.

Our results align with past studies that have indicated links between EF and repetitive, restricted behaviors in children with ASD (Bishop et al., 2006; Hill, 2004), but stand in contrast to studies which have found no links to specific ASD symptom clusters (e.g., Ozonoff et al., 2004). One reason for the discrepancies among results lies in the conceptualization of repetitive behaviors, rituals, and routine adherence as a single construct instead of discrete categories of abnormal behaviors with their particular neuropathophysiology. The present study utilizes objective assessment of motor stereotypy episodes, which is likely more sensitive and provides a more meaningful range of stereotypies (i.e., based on both frequency of stereotypy episodes and total duration). Another reason for such discrepancies refers to differences in study designs, such as variable age range and IQ levels, which are known to be related to stereotypies (e.g., hand/finger mannerisms) (Bishop et al., 2006; Ebenschen, Selzter, Lam, & Bodfish, 2009). In an attempt to overcome these issues, the present study focused on motor stereotypies and included children with AD and DLD within a limited age range and an IQ > 70. Future research should also focus on the relationship between motor stereotypies and other subtypes of repetitive behaviors, such as verbal stereotypies, and ritual and routine adherence. Findings from the present study suggest that links between motor stereotypies and EF in children with ASD may be attributed to shared brain substrates. However, this hypothesis remains speculative without support of functional brain imaging data.

Additionally, children in the present study were diagnosed with AD – Autistic Disorder, which is the most stringent definition of autism and is consistent with the most current criteria for a diagnosis of autism. The robust results from the present study may reflect the fact that children in the autism group are more severely autistic than children with autism spectrum disorders, typically used in more recent research.

Reliable measures for the coding of stereotypy severity are scarce. Current measures that rely on interview and observation (e.g., Lord et al., 2000; Wing, Leekam, Libby, Gould, & Larcombe, 2002) tend to assess the prevalence of behaviors rather than quantifying characteristics (e.g., duration) of stereotypies. A recent study began to address this limitation by coding videotaped motor stereotypies in children with autism using both quantitative and qualitative methods; stereotypies were characterized and classified based on body part (Goldman et al., 2009). The present study builds on this work by objectively assessing the frequency and duration of motor stereotypies.

Three separate measures of EF were used in this study. Although some overlap exists between the tasks, each measure highlights different domains of EF. WCST Categories assesses set shifting and flexibility; SB-IV Matrices assesses fluid reasoning, planning, and inhibition, and WISC-R Mazes measures planning, attentional focus, and spatial abilities. We found that all three measures were related to either frequency and/or duration of stereotypy episodes within range. The interaction was not significant for the WISC-R Mazes subtest, when predicting frequency of stereotypies, and only marginally significant when predicting stereotypy duration. While it appears that a range of higher-order cognitive processes subsumed under EF are related to motor stereotypies, the results of the present study allow us to begin to differentiate which EF are most significant in predicting motor stereotypies. That is, abilities such as planning and inhibition, which are more specific to the WCST and SBIV Matrices, were more relevant in predicting stereotypies than visuo-spatial abilities involved in WISC-R Mazes. Future research should replicate these findings and provide additional information that will allow us to distinguish which EF are most relevant to the prediction of stereotypies.

Inspection of Tables 2 and 3 suggests that frequency and duration have significant overlap in both the AD (r = .87) and DLD (r = .82) groups. However, it is noteworthy that the WISC-R Mazes predicted the duration but not frequency of stereotypy episodes in children with AD. This is suggestive of an incremental value in assessing both aspects of stereotypy episodes. Thus, future research should further clarify potential clinical utility of using both the frequency and duration of these behaviors.
4.1. Limitations

Several limitations to the present study need to be considered. The sample was selected based on availability of the videotaped play sessions and scores for all three EF tasks. Therefore, the sample represents children who were able to successfully complete the three EF tasks. The lack of a relationship between stereotypies and EF task performance in the DLD matched control group may be due, in part, to the restricted range and relatively low frequency of stereotypies. It is noteworthy that only six of the 22 children in the control group displayed stereotypies.

Additionally, data were collected in a single standardized setting. In the future, it would be important to assess stereotypies across different settings, such as when the child is alone versus when the child is interacting with a partner, in a familiar versus novel setting and/or with a caretaker versus with an experimenter. Furthermore, it may be that particular tasks or objects lend themselves to increased stereotypic motor behaviors. Research in this area would benefit from assessing the severity of stereotypy across a wide range of tasks and objects that may foster increased stereotypies.

4.2. Clinical and diagnostic implications

Findings from this work have practical implications for proposed revisions to diagnosis of ASD outlined by the DSM-V Task Force. The proposed ASD criteria will focus on two main areas: (1) deficits in both social communication and interactions, and (2) restricted, repetitive patterns of behavior, interests and activities, both of which must be present in early childhood (American Psychiatric Association [DSM-IV-TR], 2000). This movement toward greater emphasis on restricted, repetitive patterns of behavior for a diagnosis of ASD must be met with more objective, reliable methods of observation as advanced in this paper. In the future, research should examine the relationship between parent report and objective assessment of motor stereotypes and other repetitive behaviors to determine their contributions to improved diagnostic specificity and sensitivity.

Financial disclosure

Sylvie Goldman was supported by the Einstein/Montefiore Autism Center and a LEND grant-Leadership Education in Neurodevelopmental and Related Disabilities from the Bureau of Maternal and Child Health in the Department of Health and Human Services.

Conflicts of interest

None declared.

Acknowledgements

We thank the children and their parents for their participation. We thank Dr. Isabelle Rapin, Principal Investigator of this original study, supported by NINDS program Project NS 20489, for her insightful comments. We also thank the two blind raters who devoted their time and effort to code the videotapes used in this study. Portions of this manuscript were presented at the 39th annual North America meeting of the International Neuropsychological Society in Boston, Massachusetts on February 2nd, 2011.

References
