Published in final edited form as:

J Autism Dev Disord. 2007 July ; 37(6): 1068–1079. doi:10.1007/s10803-006-0250-7.

Executive Function in Preschoolers with Autism: Evidence Consistent with a Secondary Deficit

Benjamin E. Yerys,
Department of Psychology, University of Denver, 2155 S. Race Street, Denver, CO 80208, USA

Susan L. Hepburn,
Department of Psychiatry, Health Sciences Center, University of Colorado, Denver, CO, USA

Bruce F. Pennington, and
Department of Psychology, University of Denver, 2155 S. Race Street, Denver, CO 80208, USA

Sally J. Rogers
M.I.N.D. Institute, University of California - Davis. Sacramento, CA, USA

Abstract

Recent research on executive function (EF) deficits in autism has led investigators to conclude that EF deficits are secondary to the disorder. The current study has two major goals: (1) Examine whether specific EF deficits are present in the youngest autism group to date (mean = 2.9 years), and (2) examine whether such deficits are secondary to autism, or act as an early non-specific cognitive risk factor for autism by comparing EF abilities of this autism group to a CA-matched typically developing group. Results from Experiment 1 suggest no specific EF deficits in autism relative to MA-matched controls, while results from Experiment 2 are consistent with the hypothesis that EF deficits may emerge as a secondary deficit in autism. Alternative hypotheses are also considered.

Keywords

Executive function; Children; Autism; Cognitive flexibility

Introduction

There is accumulating evidence that late preschoolaged children (3.5–4 years) with autism fail to demonstrate specific deficits in Executive Function (EF), suggesting that such deficits are secondary to autism. However, EF has not been examined in very young children with autism. The current study has two goals: Examine (1) whether EF deficits are present at even younger ages in children with autism (3 year olds), and (2) determine whether such deficits are secondary to autism or act as an early non-specific cognitive risk factor for autism.

byerys@nova.psy.du.edu.

In what follows, we briefly define EF, and present evidence to suggest why researchers have examined EF as the primary deficit in autism. We then report a study that examines EF across three tasks in the youngest age group of children with autism to date.

EF is a blanket term referring to a set of abilities that allow individuals to achieve a particular goal (Welsh & Pennington, 1988). These abilities include working memory, inhibition, set-shifting/cognitive flexibility, self-monitoring, and generativity.

A plethora of research has examined EF in individuals with autism, and two seminal papers cogently argued how symptoms observed in autism are consistent with breakdowns in EF related to frontal lobe damage (Damasio & Maurer, 1978; Ozonoff, Pennington, & Rogers, 1991). Research with adult, adolescent, and school-aged samples of autism yielded significant EF deficits relative to a variety of control groups, leading to the primary EF deficit hypothesis (see Hill, 2004; Pennington & Ozonoff, 1996 for comprehensive reviews). Recent research has examined EF deficits as core to autism by examining young children's performance relative to children with other forms of developmental disorders (DD) matched on chronological age (CA) and mental age (MA) and typically developing children matched on MA (Daw son, Meltzoff, Osterling, & Rinaldi, 1998; Dawson et al., 2002; Griffith, Pennington, Wehner, & Rogers, 1999; McEvoy, Rogers, & Pennington, 1993). Initial studies with early school-age children (mean = 5.4-years for both studies) demonstrated significant specific deficits in children with autism relative to controls on well-validated EF tasks for children, such as A-Not-B, Spatial Reversal, Delayed Non-Matching to Sample, Delayed Response, and Delayed Alternation (Dawson et al., 1998; McEvoy et al., 1993).

However, more recent follow-up studies using even younger samples (mean = 3.9 years and 3.5 years) on similar tasks have failed to demonstrate specific EF deficits in children with autism relative to MA-matched controls (Dawson et al., 2002; Griffith et. al., 1999). Younger children with autism perform similarly to MA-matched controls on EF tasks, while older children and adults tend to perform significantly worse than MA-matched controls. The need for independent replication of such findings is critical with such clinical populations given the relative low-incidence rates and small Ns in many of the studies. Previous studies utilized comparison groups that do not effectively examine the “absolute level” of EF skills. The most effective method is to utilize a typically developing chronologically age-matched control group to highlight the overall amount of delay in EF skills.

In addition to replicating these findings in younger children, this dissociation in EF performance across age groups needs to be examined more directly. In particular, this study examines two hypotheses for this apparent dissociation: (1) EF deficits are secondary deficits to living with autism that emerge over the course of development (Dawson et al., 1998; Dawson et al., 2002; Griffith et al., 1999), or (2) EF deficits represent a non-specific cognitive risk factor that combines with other more specific risk factors to produce the disorder and is present from the beginning of the child's life.

In this study, we present two experiments to test these hypotheses. *Experiment 1* attempts to replicate previous findings by administering an EF battery of three measures (Windows,

Spatial Reversal, and A-Not-B) to the youngest sample of autism to date, and compare their performance to a DD group matched on CA and MA, as well as a typically developing group matched on MA.

Experiment 2 examines two hypotheses that attempt to explain the previous research and the results of Experiment 1. If the "secondary deficit "hypothesis is correct, then this assumes that children with autism should perform similarly to typically developing CA-matched controls at some point early in their development, and then, after living with autism for several years, deviate in the development of their EF abilities. However, if the "early non-specific cognitive risk-factor "hypothesis is correct, then children with autism will always demonstrate an EF deficit relative to typically developing CA-matched children, but not relative to MA-matched controls. Notably, this second hypothesis does not address *why* EF abilities are below MA level at later ages. However, a third hybrid hypothesis may also exist. In this hybrid or reciprocal causation hypothesis, there are early non-specific EF deficits (relative to CA controls) that contribute to the development of autism. But once autism develops, it has secondary effects on the later development of EF abilities, such that they are eventually below MA level. Because this hybrid hypothesis is less parsimonious than the only secondary hypothesis, it is important to test if the simpler only secondary hypothesis can be rejected.

Experiment 1

Method

Participants—A total of 54 children were included in Experiment 1, comprising three groups as a function of diagnosis: Autistic Disorder (n = 18; AUT), developmental delay of mixed etiology (n = 18; DD; six with idiopathic DD, four with Down syndrome, four with fragile X, and four with known chromosomal abnormalities), and typically developing children (n = 18; TYP). Two comparison groups were chosen to examine unique developmental delays specific to autism, relative to children with and without developmental delays (Burack, larocci. Bowler, & Mottron, 2002). Consistent with recent methodological debate, the clinical comparison group most appropriate for the questions in this study was determined to be a group of children with developmental delays of mixed etiology, matched on both CA and overall MA, (Seltzer, Abbeduto, Krauss, Greenberg, & Swe, 2004). See Table 1 for participant characteristics. Clinical groups were recruited from specialty clinics serving families with children with developmental disorders, parents/ advocacy groups (e.g.. Autism Society of Colorado, Mile High Down Syndrome Society, and National Fragile X Foundation), and community based service providers. The TYP group was recruited from the University Developmental Participant Pool and through word of mouth.

Children in the clinical groups were between the ages of 24–45 months (mean age = 35 months; standard deviations ranged from 3.8 to 5.6 months across groups). Children in the TYP group were recruited in an effort to be comparable on mental age to the clinical groups, and thus, they were significantly younger (mean = 22 months). Mental ages were measured with the Mullen Scales of Early Learning (MSEL; Mullen, 1992,1997). There were no significant group differences for overall mental age, $F(2, 53) = 1.51$, $p = .23$, non verbal

mental age, $F(2, 53) = 0.76$, $p = .47$. There was, however, significant group differences on verbal mental age, $F(2, 53) = 6.68$, $p < .01$. Post-hoc analyses with Tukey's test revealed that the TYP group had higher verbal mental ages than both clinical groups ($p < .05$), but not a significant difference between the two clinical groups. The groups did differ significantly on gender, $\chi^2(N = 51) = 6.08$, $p < .05$. The AUT group had significantly fewer girls than the DD or TYP groups.

At the time of enrollment none of the children in the AUT group presented with a known significant medical condition, history of illness, or acquired head injury. However, during the course of the longitudinal study one child from our AUT group was identified as having an abnormal reading on an EEC and another child had an abnormal MRI. Neither child was identified as having seizures, or any known medical condition. The profile of the skills of these children were comparable to the rest of the group, therefore these children were not excluded.

Nine children in the DD group presented with an additional medical condition. Five (28%) children experienced a prenatal insult, and/or toxic exposure and prematurity, 2 (11%) children experienced heart problems, 1 (5.5%) child presented with Cochayne Syndrome, and 1 (5.5%) child presented with Hypothyroidism. Children with delays of different etiology (e.g., fragile X vs. idiopathic) were not different from each other in overall mental age; therefore, they were combined into one developmental delay group. Overall, no etiological subgroup was over-represented.

Inclusion criteria for the current study were conducted in three separate waves: 1. diagnostic criteria, 2. cognitive criteria, and 3. EE criteria.

Diagnostic Criteria: Inclusion criteria for each of the groups were applied in a strict manner. Every child participated in a diagnostic assessment battery designed to identify symptoms of autism in young children. A diagnosis of autism was based upon the child meeting 4/5 of the following criteria: (1) Previous clinical diagnosis of autism, (2) scores exceeding the autism cut-off on the Autism Diagnostic Interview—Revised (Lord, Rutter, & LeCouteur, 1994), (3) scores exceeding the autism cutoff on the Autism Diagnostic Observation Schedule—Generic (Lord, Rutter, DiLavore, & Risi, 1999), (4) endorsements on a DSM-IV checklist, and (5) current clinical diagnosis of autism. Psychologists with extensive experience with autism formulated the clinical diagnoses. All children in the autism group had normal hearing and vision corrected to the normal range. The DD group comprised four subgroups of developmental delays: Idiopathic developmental delays, Down syndrome, fragile X syndrome, and other known chromosomal disorders. All children had normal hearing and vision corrected to the normal range, and had DNA verifica tion of fragile X, Down syndrome status, or other known chromosomal disorders.

Children were included with a diagnosis of idiopathic DD if they met the following criteria: (1) developmental delay similar to that observed in the autism group using a composite age equivalence score from the Mullen Scales of Early Learning (Mullen, 1992, 1997); (2) absence of fragile X or down's syndrome diagnosis; (3) no past or current diagnosis of

autism; and (4) not meeting criteria for autism on two or more of the autism diagnostic measures (e.g., ADOS, ADI-R, DSM-IV).

Children were included in the TYP group if they met the following criteria: (1) no presence of devel opmental delays; (2) absence of a known chromosomal disorder; (3) no past or current diagnosis of autism, and (4) not meeting criteria for autism on two or more of the autism diagnostic measures (e.g., ADOS, ADI-R, DSM-IV). All children in the TYP group had normal hearing and vision and did not present with any significant medical conditions.

Cognitive Criteria: Cognitive inclusion criteria required a minimum overall MA of 18 months on the MSEL to insure that performance was not at floor on the EF measures.

EF Criteria: EF inclusion criteria required children to complete at least two of three tasks in the battery in order to be included in the study.

Measures

Executive Function (EF) Battery: The EF battery consists of three tasks used in an earlier study by Griffith and colleagues (1999). This battery was designed to challenge children's working memory, inhibitory abilities and set-shifting abilities. All tasks involved the retrieval or search for a toy or food reward.

Windows Task: The current task (Rogers & Wehner, 1997) was a simplified version of the original Windows task developed by Russell and colleagues (Hughes & Russell, 1993; Russell, Mauthner, Sharpe, & Tidswell, 1991). Children were presented with two clear boxes showing the boxes' contents. If children selected the box with a reward inside, then they were not rewarded. If children selected the empty box, then they were rewarded. This adaptation challenged children to (1) infer a rule(s) (reach for empty box and/or ignore box with prize) and (2) maintain the rule(s) in working memory to overcome their bias (to reach for the box with the reward inside).

Children received a total of 21 trials (five training and 16 experimental; see Rogers & Wehner, 1997 for information on training trials). During the 16 experimental trials, the examiner pulled both boxes under the table, baited one (with pre-selected random order of side of baiting) and presented both. The children's points/reaches were considered either correct (i.e., pointing toward or reaching to the empty box) or incorrect (i.e., pointing to or reaching to the box with the reward). If children pointed incorrectly for more than one trial consecutively, then the additional errors were coded as perseverative.

Previous studies (Hughes & Russell, 1993; Russell et al., 1991) established two criteria for estimating children's success on the Windows task because of the bimodal nature of the data: (1) a “*conservative* ”criterion which required no more than three errors over 20 experimental trials with the child answering correctly on Trial 1 and (2) a “*liberal* ”criterion which required no more than three errors over any of the 20 experimental trials. Given that our version has only 16 experimental trials we altered the criterion to a cut-off of ±15 for the liberal criterion, but maintained the requirement for answering Trial 1 correctly to pass on the conservative criterion.

Spatial Reversal: This task (Kaufmann, Leckman, & Ort, 1989) was administered as described in previous research (Griffith et al., 1999; McEvoy et al., 1993). In this task, the experimenter sits in front of the child with a screen in between them. The examiner hides a reward behind the screen in one of two containers placed to the left and right of the child's midline, and tells the child that a reward is being hidden (e.g., “I'm hiding an M&M”). Then, the screen is removed and children see two cups (a reward is under both cups on trial 1). After children find a reward, the experimenter continues to hide the reward at that location until the child has 4 consecutive correct searches, then the reward location is switched without a cue. This task challenged children to (1) maintain the previous location of the reward in working memory, (2) to flexibly shift reward association for two locations.

Children's responses were coded as correct (i.e., finding the reward) or incorrect (i.e., not finding the reward). After the first switch, a child's responses were coded as correct (i.e., adjusting to the change, and choosing the correct cup after feedback), failure to maintain set errors (i.e., children switched sorting locations before completing a set of 4) or perseverative (i.e., children searching the previous location after receiving feedback on the previous trial that the location had changed). Each child received a total of 23 trials, and therefore had the opportunity to make four switches. Raw scores of correct searches and perseverative responses were used in these analyses.

A-not-B: This task was given as described in Griffith et al. (1999). In this version, children must retrieve a toy after watching the experimenter hide the toy in one of two identical locations (“A”). After two correct searches, children watched the experimenter hide the toy in a new location (“B”). Initially, the delay be tween the hiding of the toy and search was 8 s. If the child correctly searched both times at location A but incorrectly on the B trial, then the experimenter continued additional trials with a delay of 8 s. However, if the child correctly searched on both A trials *and* the B trial, then the next set of trials was conducted with a longer delay (i.e., 12 or 15). If the child searched correctly for six trials at 15 s, then the task was ended. By systematically manipulating the delay we were able to insure the following: No ceiling or floor effects, and that the task's working memory demands were sufficiently difficult—i.e., insuring perseveration—for most children.

If the children were unable to search correctly for both A trials with an 8 s delay, then the next set of trials was conducted with a shorter delay (5, 3, or 0 s) until the they were able to search correctly for both A trials. If any children were unable to search correctly for both A trials with a 0 s delay, then the task was ended. Children received a maximum of 24 trials, but the task ends if children searched correctly for 2 trials after the 5th reversal.

This task challenged children to (1) maintain the current location of the reward in working memory across a substantial delay, and (2) on B trials, maintain the location of the reward to overcome a bias to search the previous hiding location of the toy/reward, and finally (3) to flexibly shift reward association for two locations.

For this task we coded each child's delay length as a “proxy ”for their correct searches. Delay length seemed an appropriate proxy because the delay was systematically varied based on children's correct searches (Griffith et al., 1999 also examined delay length).

Moreover, delay length provides an index of children's working memory capacity. Error coding was broken into three main categories: Error after correct search, error after a reversal, and error after an error. Each category captures a different breakdown in children's performance. Error after correct search captures children's failures to maintain the location of the hidden reward. Error after a reversal captures children's failures to use the visual in formation that the reward's location has switched. Error after an error captures children's failures to use visual and feedback information that the reward's location has switched. Raw scores of delay length, errors, and the percentage of delay length were used in the analyses.

Procedure—This study is part of a larger longitudinal study and includes measures not reported on here. The entire study was carried out under IRB approval from the University of Colorado Health Sciences Ctenter. Consent forms were reviewed with each family and all questions were answered before consent was obtained and before any measures were gathered. Each child's mother was interviewed about her child's development and behavior using the ADI-R and other measures not reported here, usually during a home visit. The ADOSG, Mullen's, and the Executive Function Battery were administered in the lab over several visits, along with other measures not reported here. Children were given numerous breaks throughout testing.

Results

Preliminary Analyses—Before conducting our proposed inferential analyses, we examined our data for: (1) significant kurtosis, (2) skew, (3) significant outliers, and (4) whether children's performance was significantly different from chance. The Windows task was found to exhibit a non-normal distribution with 66% of all children exhibiting floor (0 correct) or ceiling effects (15 or 16 correct) (see Table 2 for sample distributions). This bimodal distri bution is consistent with previous studies using the Windows task (Hughes & Russell, 1993; Russell et al., 1991); thus we used non-parametric analyses on these data.

All the other variables met the assumptions of normality, allowing us to conduct our parametric analyses: Multiple Analysis of Variance tests for group differences.

We conducted additional analyses for each task to assess whether the children's responses could be accounted for by chance.

Windows Task: As shown in Table 2a and b, there were no significant group differences in the number of children passing and failing the Windows task, regardless of the scoring criterion (conservative: $\chi^2(2, N = 48) = .171$, $p > .05$; liberal: $\chi^2(2, N = 50) = .838$, $p > .05$).

Spatial Reversal and A-Not-B: To examine group differences (three groups) among the remaining dependent variables of interest (8 across 2 tasks) and control for inflation of a Type I error, we used a multivariate analysis of variance where group was entered as a fixed factor and the dependent variables of interest were entered as dependent variables. All groups performed similarly on the correct response measures across all tasks (see Table 3). The main effect of diagnostic group was non-significant, $F(2, 38) = 1.02$, $p = .446$. Despite the non-significant omnibus test between groups, we conducted exploratory univariate analyses on the dependent variables to confirm that no individual test reached significance.

There were significant differences between groups in errors on the perseverative responses on Spatial Reversal ($F(2, 38) = 3.61$, $p = .04$) and a trend for significant differences between groups on errors after an error on A-not-B ($F(2, 38) = 3.16$, $p = .05$). Tukey's test revealed that the TYP group made significantly more perseverative errors on Spatial Reversal and more errors after an error on A-not-B than the AUT group, but not significantly more than the DD group.

To address the question of whether children performed at chance in Spatial Reversal, we first examined whether children achieved multiple sets in the task. This metric appeared the most appropriate for demonstrating non-random performance by these age groups on Spatial Reversal, because using a simple proportion correct metric (requiring the groups demonstrate > .5 proportion correct) assumes that children are equally likely to search in either location, which is not the case when they are rewarded for searching in either location on trial 1. If children were searching in a completely random fashion, then they would be predicted to achieve no sets. Or, perhaps children could achieve one set because simple behavioral principles suggest that children would return to a previously rewarded location, but they would have been unable to infer the rule and flexibly shift locations. Thus, if children reliably established more than one set on Spatial Reversal, then we could reasonably conclude their performance was not at chance. All three groups achieved significantly more than 1 set on average on the Spatial Reversal task (AUT: $t(16) = 3.23$, $p < .01$; DD: $t(16) = 3.95$, $p < .01$; TYP: $t(17) = 3.6$, $p < .01$). Therefore, children of this mental age were not likely to be responding by chance, and all groups performed similarly.

With respect to the A-Not-B task, there is no concern for children performing at chance because the task was conducted in such a way as to elicit correct performance on “A ”searches and failures on “B” searches. Even children who received a delay of Os were able to correctly search on “A ”trials. Thus, no children in the study performed at chance on this task.

Discussion

Very young children with autism *do not* demonstrate specific EF deficits relative to either control group on a variety of EF measures. This null finding independently replicates previous research with other pre-school-aged children with autism (Dawson et al., 2002; Griffith et al., 1999). At a first pass, the results of Experiment 1 appear to support a secondary deficit hypothesis; however, if this hypothesis is true, then children with autism should perform similarly to CA matched typically developing children at some early point in development. If the autism group exhibits deficits relative to this control group, then these results would support the non-specific early cognitive risk factor hypothesis of EF in autism.

Thus, Experiment 2 will compare the current sample of children with autism to a CA-matched, typically developing sample of children. Because our groups were recruited through a longitudinal study on the behavioral phenotype of autism and DD, Experiment 2 will have two major limitations: (1) the CA-matched TYP group used in Experiment 2 is almost identical to the MA-matched TYP group used in Experiment 1 (this data was collected at our Time 2 assessment), thus most (66% or 12/18 children) in the CA-TYP group have previous experience with the EF battery, whereas this was the first experience in

the EF battery for all of the autism group (Time 1 assessment); (2) Because the CATYP group was assessed during Time 2 in the study, only two of the three EF tasks (Windows and Spatial Reversal) were retained for the Time 2 assessment phase in the longitudinal study. Thus, Experiment 2 only compares the groups on this subset of EF tasks.

Experiment 2

Method

Participants—A total of 36 children consisting of two groups (AUT, CA-TYP controls) of 18 children participated in this study. The AUT group was the same used in Experiment 1. The CA-TYP control group was almost the same group of TYP (12/18 children) used as MA matches in Experiment 1 tested one year later as part of the larger longitudinal study. The six new children in the CA-TYP group had normal hearing and vision and did not present with any developmental delays, medical diagnoses, or meet criteria for autism on any measure. As shown in Table 1, there were no significant differences in CA, $F(l, 34) = 2.47$, $p = .13$, SES $F(l, 32) = .002$, $p = .95$, or Ethnicity between groups the four groups, $\chi^2(3, N = 71) = 7.26$, $p = .610$. However, there were significant differences in gender between groups, $\chi^2(3, N = 36) = 4.13$, $p = .25$. The AUT group had significantly more boys than the CA-TYP group, $\chi^2(3, N = 36) = 9.26$, $p < .005$. As expected, the ANOVA revealed significant differences in NVMA, $F(l, 34) = 20.07$, $p < .001$, VMA, $F(l, 34) = 91.59$, $p < .001$, and overall MA $F(l, 34) = 62.91$, $p < .001$. In all cases, the CA-TYP achieved significantly higher MA scores than the AUT group.

Measures—The Windows task and Spatial Reversal were given as described in Experiment 1.

Procedure—The procedure was identical as described in Experi ment 1 for all groups.

Results

Preliminary Analyses—Identical to Experiment 1, all groups performance in the Windows task was bi-modally distributed with 75% exhibiting floor (0 correct) or ceiling effects (15 or 16 correct). Thus, non-parametric analyses (χ^2) were again conducted on this task using the same criteria as established in Experiment 1, however the 2 × 2 nature of this study allowed for Yates' Continuity Correction.

Both groups' performance was normally distributed without significant kurtosis or skew on the Spatial Reversal task. Thus, a one-way ANOVA was conducted with group assignment as the between-subjects factor and the key variables of interest from Spatial Reversal as the dependent variables (correct trials, sets achieved, perseverative errors, and failure to maintain set errors).

On the Windows task, χ^2-test yielded no significant differences among the groups using either the conservative criteria, $p = .15$, or the liberal criteria, $p = .16$ (see Table 2). Despite the non-significant statistic there is approximately 30% difference between the groups with the CA-TYP group (75% passing in conservative and liberal scoring) performing better than the AUT group (44% passing in conservative and 47% passing in liberal scoring).

On the Spatial Reversal task, there were no significant differences between groups on the number of correct searches, $F(1, 33) = 1.65$, $p = .21$, sets achieved, $F(1, 33) = 2.61$, $p = .12$, perseverative responses, $F(1, 33) = 0.12$, $p = .73$, and failure to maintain set errors, $F(1, 33) = 1.75$, $p = .20$ (see Table 4).

Correlations between EF measures, MA, and CA—The results of Experiment 2 posited an interesting paradox; If IQ and EF tend to have a reliable positive correlation (see Liss et al., 2001; Welsh, Pennington, Ozonoff, Rouse, & McClabe, 1990), then how can we explain the finding that our CA-matched typically developing group demonstrated a significantly higher level of overall MA compared to the autism group, but the groups perform similarly on both EF tasks.

Examining the interrelationship between EF performance with MA and CA in this age group became a point of interest. Unfortunately, due to the small sample size all four groups from both experiments were collapsed in analyses to increase statistical power. This also limited analyses to the Windows and Spatial Reversal tasks because the CA-matched TYP group did not participate in A-not-B. These analyses yielded significant correlations between overall MA and EF performance in predicted directions for five of the six dependent variables in the EF battery (only failure to maintain set in Spatial Reversal was non-significant) (see Table 5). With respect to EF performance and CA, the analyses yielded significant correlations in predicted directions for three of the six dependent variables in the EF battery (number correct, sets achieved, and perseverative responses in Spatial Reversal) and a trend for a fourth variable (Windows task using liberal scoring). Additional correlation analyses were not conducted within each group because of inadequate power to detect differences.

Discussion

Overall, very young children with autism demonstrated almost no EF deficits relative to CA-matched typically developing children on both EF tasks. Notably, the CA-matched group outperformed the autism group by approximately 30% in the Windows task and the associated effect size was moderate-to-strong ((Cohen's $d = .67$), which would likely yield a significant differ ence with a larger sample. However, because most of the CA-matched group participated in Windows a year earlier, it is possible that the difference in performance and the associated effect size may have been less if the control group had no prior experience with the task. Clearly, future research would need to run a larger sample with an independent CA-matched group to answer this question.

Nevertheless, these null results posited an interesting paradox in that there were significant differences in MA between the groups, but no significant differences in EF abilities as measured by the Windows and Spatial Reversal tasks. Moreover, the general pattern observed in Experiment 2 suggests that the children with autism have similar EF abilities compared to their CA-matched typically developing peers. Follow-up correlational analyses collapsing across all four groups in both experiments clearly demonstrated significant relationships between cognitive and EF abilities. This significant relationship in the context of no significant group differences on EF tasks raises a new paradox: If children with autism have significantly lower MA abilities to same-aged peers, then why are they demonstrating

equivalent EF abilities? As elaborated below, one issue may be that the current battery of EF tasks may not be sensitive to the deficits that are observed in older ages.

Finally, the null results from both Spatial Reversal and Windows are consistent with the just secondary deficit hypothesis of EF in autism.

General Discussion

Overall, on EF measures children with autism *did not* have a specific deficit in EF relative to either MA-matched control group; moreover, they *did not* exhibit a delay in EF abilities relative to the CA-matched typically developing control group. That is, our results demonstrate children with autism perform similarly to CA and MA-matched children with DD and MA-matched typically developing children on almost all EF measures. Once again, early research on EF abilities in autism suggests that EF deficits are not unique and specific to children with autism, and thus an unUkely primary deficit for the disorder (Dawson et al., 1998; Dawson et al., 2002; Griffith et al., 1999). When comparing children with autism to CA-matched typically developing children, the children with autism exhibited no EF deficits, and this result was consistent with the only secondary deficit hypothesis.

One point of note is that the EF battery reported here mostly examined working memory, inhibition, and cognitive flexibility abilities using simple visual-search based tasks, and did not measure other components of EF (e.g., generativity, extra-dimensional shifting, or non-verbal reasoning tasks that do not rely on spatial location). There is some evidence to suggest that generativity, extra-dimensional shifting, and more complex spatial reasoning tasks may yield significant specific EF deficits for children with autism (Goldberg et al., 2005; Hughes, Russell, & Robbins, 1994; Rutherford & Rogers, 2003; Williams, Goldstein, & Minshew, 2006). In the case of generativity, more research is required to confirm the rehabiUty of such measures and their developmental appropriateness. In the cases of extra-dimensional shifting and complex non-verbal reasoning tasks (without using spatial location), current research has not established developmentally appropriate tasks for children under the age of 5.

These findings replicate other studies examining EF in preschoolers/ early school-age children with autism in comparison to MA-matched control groups (Dawson et al., 2002; Griffith et al., 1999), however the findings from these studies are not consistent with other research demonstrating unique EF deficits in school-age children and adults with autism (see Hill, 2004; Pennington & Ozonoff, 1996 for reviews).

Combining the results from Experiment 1 with previous research supports two major hypotheses regarding EF abilities in individuals with autism: (1) EF difficulties appear to increase with age in individ uals with autism and (2) EF difficulties are not a primary deficit causing autism. However, results from Experiment 1 did not provide direct support for the secondary deficit hypothesis of autism. As outlined in the introduction, in order to support the secondary deficit hypothesis, children with autism should demonstrate similar EF abilities to CA-matched typically developing children at some point early in develop ment, and then their EF abilities should deviate from CA-matched typically developing children over the course of development.

The overall results of Experiment 2 are consistent with the secondary deficit hypothesis of EF deficits in autism, while challenging the early non-specific cognitive risk factor hypothesis. More specifically, if EF was an early non-specific risk factor, then children with autism should always perform significantly worse than typically developing CA-matched control groups. As demonstrated in Experiment 2, the autism group performed similarly to the CA-matched control group. Thus, the data are consistent with the key assumption in a secondary deficit hypothesis of EF in autism: EF abilities appear to be developing typically at early stages of development. However, making strong conclusions from the current study are unwarranted as the secondary deficit hypothesis predicted null results in Experiment 2.

Although the general pattern reported here is consistent with a secondary deficit in autism, other hypotheses may also explain why such EF deficits are not observed at early ages. There are three general hypotheses attempting to explain the emergence of an EF deficit in autism across development: (1) the delay in EF deficits may result from a lifetime of living with autism which may provide a poverty of social experiences (Dawson et al., 2002; Griffith et al., 1999) to develop and hone EF skills; (2) the delay in EF deficits may arise from non-critical task demands in the EF tasks given to older children and adults with autism (Griffith, 2003; Ozonoff, 1995); (3) or the delay in EF deficits may result from a deficit in processing complex information which is not observable at early ages due to the types of tasks that are used with preschool-aged children.

The first hypothesis has not been systematically tested yet, but there is much indirect evidence to suggest that EF and social skills are connected. There are several studies connecting poor joint attention abilities and poor EF abilities in children with autism (e.g., Dawson et al., 2002; Griffith et al., 1999; McEvoy et al., 1993). One possible social mechanism for poor EF skills may be related to early deficits in orienting to social stimuli and decreased interest in social stimuli (Dawson et al., 1998, 2002; Swettenham et al., 1998). Moreover this lack of interest may contribute to children avoiding or missing group participation. Group participation provides multiple opportunities to practice EF skills. For example, when an individual participates in a group discussion, s/he must track multiple speakers' comments simultaneously, while generating and modifying his/her own contributions. Moreover, there is also a large literature, which will not be reviewed here, connecting 'theory of mind' deficits and EF (see Carlson, Moses, & Claxton, 2004; Hughes, 2001; Joseph & Tager-Flusberg, 2004 for recent developments in this area of research). Systematic research must be conducted to clarify the contribution of social experience throughout development on cognitive and neural mechanisms underlying EF, and how deviation from typical social experiences contributes to an emergent deficit as seen in autism.

The second hypothesis has been systematically tested in two studies using computerized and human administrations of the WCST. In both studies, individuals with autism were found to perform significantly better on computerized versions of the WCST than human administered versions (Griffith, 2003; Ozonoff, 1995). These findings have led to a "feedback hypothesis", which argues that individuals with autism may respond better to computerized feedback because of its non-variable nature (Griffith, 2003). This hypothesis is intriguing; however, it is unclear what specific aspects of human administration disrupt

the performance of individuals with autism. Moreover, it is unclear whether this hypothesis can explain the wide breath of EF deficits seen in individuals, because many EF tasks in daily life require planning and execution in addition to using feedback.

The third hypothesis, complex information processing deficit, argues that EF deficits present in autism are observable when individuals are presented with stimuli that requires them to abstract key information from a large, complex stimulus (Williams et al., 2006; also see Minshew & Goldstein, 2001; Williams, Goldstein, & Minshew, 2005). Several behavioral studies have demonstrated that adults and school-age children with autism demonstrate significant difficulties on working memory tasks that do not rely on memory for location, such as the Design Memory and Picture Memory subtests of the Wide Range Assessment of Memory and Learning (WRAML), but demonstrate intact abilities in simple spatial EF tasks that do rely on memory for location, such as Visual Learning. Clearly, the data collected in the current study is consistent with this hypothesis, because all of our tasks can be simplified to different variants of “memory for location” tasks with a variety of rule manipulations. Thus, neither experiment could clearly test this hypothesis.

Limitations of the Current Study

There are several limitations of the current study. First, all tasks in this study could be categorized as “memory for location ”tasks, and thus fail to address at least one alternative hypothesis that has gathered empirical evidence recently. Second, Experiment 2 is limited by a lack of independence in the CA-matched control group, and this significantly hindered our ability to draw strong inferences on at least one of the tasks (Windows). Because this data was drawn from an ongoing longitudinal study, this design flaw was unavoidable. Another major limitation of the current study is a lack of longitudinal data for all groups. Having such data would allow us to map out EF development across time and examine developmental trajectories of EF across the different diagnostic groups. Tracking developmental trajectories is clearly the next step in examining the validity of the secondary deficit hypothesis of EF deficits in autism. Conducting such analyses would allow this hypothesis to make clear predictions regarding group differences—a major limitation in the construction of Experiment 2. Such analyses are underway as we complete our longitudinal testing, however this analysis will also be limited by the EF tasks selected for this study.

Another limitation of the current study is a selection bias. In our group selection process we eliminated children with overall MA's below 18 months of age, and this may have skewed our results such that we selected higher-functioning children with autism who were better at EF. This is a distinct possibility and future research should widen the sample MA range to include lower-functioning children; however, it is likely that other, simpler EF tasks may be necessary to examine children's abilities, and this may increase the difficulty of testing alternative hypotheses, such as the “complex information processing deficit.”

The current study raises several questions that require future explorations of EF in children with autism. First, future research should pursue to examine the relationship of EF and MA abilities in typically developing and developmentally delayed samples large enough to conduct meaningful analyses. This will provide clear insight as to whether the correlation is driven by both samples or by typically developing children. Second, future studies need to

directly examine whether EF deficits arise from the consequence of living with autism, or relate to non-EF task factors, such as the feedback hypothesis (which could be expanded into more of a social motivation hypothesis) or the complex information processing deficit hypothesis. If these other hypotheses can be ruled out or controlled for and EF deficits are still observed, then the secondary deficit hypothesis will require additional direct tests.

Finally, a much larger question for the field is to develop a standardized EF metric for infants and toddlers to aid clinicians and researchers in identifying delayed EF development. For example, even though the research presented here suggests no EF deficits in our autism sample relative to three different types of control groups, no standardized EF metric exists as of yet to state whether their absolute EF abilities are developing typically or not. Clearly, this is a much larger issue than the current paper could address, but nonetheless an issue that would aid future EF research in such young samples.

Acknowledgment

The research was supported by two National Institute of Child Health and Human Development grants: The Collaborative Programs of Excellence in Autism (HD3S468), and the Mental Retardation and Developmental Disabilities Center grant (HD04024). We would like to thank all the children and their families who participated in this study.

References

Burack JA, larocci G, Bowler D, Mottron L. Benefits and pitfalls in the merging of disciplines: The example of developmental psychopathology and the study of persons with autism. Development and Psychopathology. 2002; 14:225–237. [PubMed: 12030689]

Carlson SM, Moses LJ, Claxton LJ. Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. Journal of Experimental Child Psychology. 2004; 87:299–319. [PubMed: 15050456]

Damasio A, Maurer R. A neurological model for childhood autism. Archives of Neurology. 1978; 35

Dawson G, Meltzoff AN, Osterling J, Rindaldi J. Neuropsychological correlates of early symptoms of autism. Child Development. 1998; 69:1276–1285. [PubMed: 9839415]

Dawson G, Munson J, Estes A, Osterling J, McPartland J, Toth K, Carver L, Abbott R. Neurocognitive function and joint attention ability in young children with autism spectrum disorder versus developmental delay. Child Development. 2002; 73:345–358. [PubMed: 11949896]

DiLavore PC, Lord C. The pre-linguistic autism diagnostic observation schedule. Journal of Autism and Developmental Disorders. 1995; 25:355–379. [PubMed: 7592249]

Goldberg MC, Mostofsky SH, Cutting LE, Mahone EM, Astor BC, Denckla MB, Landa RJ. Subtle executive impairment in children with autism and children with ADHD. Journal of Autism and Developmental Disorders. 2005; 35:279–293. [PubMed: 16119469]

Griffith EM. Examining the feedback hypothesis of the executive function deficit in autism. Dissertation Abstracts International: Section B: The Sciences & Engineering. 2003; 63:3473.

Griffith EM, Pennington BF, Wehner EA, Rogers SJ. Executive function in young children with autism. Child Development. 1999; 70:817–832. [PubMed: 10446722]

Hill EL. Evaluating the theory of executive dysfunction in autism. Developmental Review. 2004; 24:189–233.

Hughes, C. Executive dysfunction in autism: Its nature and implications for the everyday problems experienced by individuals with autism.. In: Burack, JA.; Charman, T.; Yirmiya, N.; Zelazo, PR., editors. The development of autism: Perspectives from theory and research. Lawrence Erlbaum Associates; Mahwah, NJ: 2001.

Hughes C, Russell J. Autistic children's difficulty with mental disengagement from an object: Its implications for theories of autism. Developmental Psychology. 1993; 29:498–510.

Hughes C, Russell J, Robbins TW. Evidence for executive dysfunction in autism. Neuropsychologia. 1994; 32:477–492. [PubMed: 8047253]

Joseph RM, Tager-Flusberg H. The relationship of theory of mind and executive functions to symptom type and severity in children with autism. Development and Psycho pathology. 2004; 16:137–155.

Kaufman P, Leckman J, Ort S. Delayed response performance in males with Fragile X. Journal of Clinical and Experimental Neuropsychology. 1989; 12:69.

Lord C, Rutter M, Le Couteur A. Autism Diag nostic Interview—Revised: A revised version of a diagnostic interview for caregivers of individuals with possible perva sive developmental disorders. Journal of Autism and Developmental Disorders. 1994; 24:659–685. [PubMed: 7814313]

Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S. Autism Diagnostic Observation Scale—WPS (ADOS-WFS). Western Psychological Services; Los Angeles: 1999.

Liss M, Fein D, Allen D, Dunn M, Feinstein C, Morris R, Waterhouse L, Rapin I. Executive functioning in high-functioning children with autism. Journal of Child Psychology and Psychiatry. 2001; 42:261–270. [PubMed: 11280422]

McEvoy R, Rogers SJ, Pennington BF. Executive function and social communication deficits in young autistic children. Journal of Child Psychology and Psychiatry. 1993; 34:563–578. [PubMed: 7685360]

Minshew NJ, Goldstein JG. The pattern of intact and impaired memory functions in autism. Journal of Child Psychology and Psychiatry. 2001; 42:1095–1101. [PubMed: 11806691]

Mullen, EM. Mullen scales of early learning. T.O.T.A.L. Child, Inc.; Cranston, RI: 1992.

Mullen, EM. Manual for the Mullen scales of early learning: AGS edition. American Guid ance Service, Inc.; Circle Pines, MN: 1997.

Ozonoff S. Reliability and validity of the Wisconsin Card Sorting Test in autism. Neuropsychology. 1995; 4:491–500.

Ozonoff S, Pennington BF, Rogers SJ. Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. Journal of Child Psychology and Psychiatry. 1991; 31:1081–1105. [PubMed: 1787138]

Pennington BF, Ozonoff S. Executive functions and developmental psychopathologies. Journal of Child Psychology & Psychiatry. 1996; 37:51–87. [PubMed: 8655658]

Rogers SJ, Wehner EA. Modification of the Windows task for young children. 1997 Unpublished data.

Russell J, Mauthner N, Sharpe S, Tidswell T. The ‘Windows task’ as a measure of strategic deception in preschoolers and in autistic subjects. British Journal of Developmental Psychology. 1991; 9:331–349.

Rutherford MD, Rogers SJ. Cognitive underpin nings of pretend play in autism. Journal of Autism and Developmental Disorders. 2003; 33:289–302. [PubMed: 12908832]

Seltzer MM, Abbeduto L, Krauss MW, Greenberg J, Swe A. Comparison groups in autism family research: Down syndrome, fragile x syndrome, and schizophrenia. Journal of Autism and Developmental Disorders. 2004; 34:41–48. [PubMed: 15098956]

Swettenham J, Baron-Cohen S, Charman T, Cox A, Baird G, Drew A, Rees L, Wheelwright S. The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. Journal of Child Psychology and Psychiatry. 1998; 39:741–753.

Williams DL, Goldstein G, Minshew NJ. Impaired memory for faces and social scenes in autism: Clinical implications of the memory disorder. Archives of Clinical Neuropsychology. 2005; 20:1–15. [PubMed: 15620811]

Williams DL, Goldstein G, Minshew NJ. The profile of memory function in children with autism. Neuro psychology. 2006; 20:21–29.

Welsh M, Pennington BF. Assessing frontal lobe functioning in children: Views from developmental psychology. Developmental Neuropsychology. 1988; 4:199–230.

Welsh MC, Pennington BF, Ozonoff S, Rouse B, McCabe ERB. Neuropsychology of early-treated phenylketonuria: Specific executive function deficits. Child Development. 1990; 61:1697–1713. [PubMed: 2083493]

Table 1

Participant demographics by diagnosis for both experiments

	Autistic disorder (*n* = 18)	Mixed DD (*n* = 18)	Typical development (*n* = 18)	CA-typical development (*n* = 18)
Chronological age (months)				
M (SD)	34.8 (3.8)	35.5 (5.6)	22.2*a* (4.5)	32.6 (4.6)
Range	26-41	24-45	15-35	25-43
Nonverbal MA (months)				
M (SD)	26.8 (6.2)	24.8 (4.3)	24.9 (5.3)	36.0_{a} (6.1)
Range	19-47	19-35	18-41	25-47
Verbal MA (months)				
M (SD)	21.1_{c} (6.2)	22.1 (4.1)	27.3_{b} (5.9)	40.1_{a} (5.7)
Range	13-37	14-30	19-43	27-53
Overall MA (months)				
M (SD)	23.9 (5.7)	21.8 (5.2)	22.1 (5.7)	38.2_{a} (5.1)
Range	18-42	19-30	18-42	26-46
Socioeconomic status				
***M* (SD)**	49.8 (13.5)	51.8 (12.5)	48.7 (14.0)	50.0 (9.3)
Range	22-66	15-66	22-66	35-66
Gender				
Male	15*a*	10	8	6
Female	3	8	10	12
Ethnicity				
European	15	14	16	17
African-American	1	0	0	0
Hispanic	0	2	1	1
Multiple ethnicities	2	1	1	0
Unreported	0	1	0	0

[a]Subscript indicates significant differences between groups. Same subscript indicates two groups are not significantly different

Table 2

Group performance on the Windows task in both experiments

Windows performance	Group			
	AUT (*N* = 17)[a]	DD (*N* = 17)[a]	TYP (*N* = 16)[a]	CA-TYP (*N* = 16)[a]
Conservative criterion[b]				
# Pass	7	6	7	12
# Fail	9	10	9	4
Liberal criterion				
# Pass	8	10	7	12
# Fail	9	7	9	4

[a] One child from the AUT and DD group and two from the TYP and CA-TYP group failed to complete the Windows task due to frustration

[b] Individual trial data was lost for one child in AUT and DD groups, so their data was removed from the conservative criterion analysis

Table 3

Group performance on the Spatial Reversal and A-not-B tasks in Experiment 1

Tasks	Group				
	AUT(N per task)	DD	TYP	F	p
Spatial Reversal	(*N* = 16)[a]	(*N* = 17)[a]	(*N* = 15)[a]		
# Correct	12.12 (6.41)	12.88 (5.36)	11.67 (4.20)	0.21	.81
# Perseverations	1.81 (2.14)	3.18 (3.41)	5.00 (4.12)	3.61	.04
# Failure to maintain sets	0.37 (0.72)	0.24 (0.44)	0.33 (0.16)	0.24	.79
# Sets	2.06 (1.29)	2.35 (1.41)	1.80 (1.01)	0.77	.47
A-Not-B	(*N* = 16)[a]	(*N* = 17)[a]	(*N* = 15)[a]		
Delay (0-15 s)	7.00 (5.68)	6.35 (4.57)	5.80 (4.35)	0.23	.79
# Wrong following reversal trials	1.81 (1.52)	1.59 (1.33)	1.80 (1.21)	0.14	.87
# Wrong following errors	0.37 (0.72)	1.53 (1.77)	2.00 (2.65)	3.16	.05
# Wrong following correct	0.50 (0.73)	0.65 (0.86)	1.00 (1.20)	1.15	.33

[a]Two children from the AUT group, one from the DD group, and three from the TYP group failed to complete both Spatial Reversal and A-Not-B due to frustration with the tasks, so they were dropped from the MANOVA

Table 4

Group performance on Spatial Reversal Experiment 2

Task	Group			
	AUT (*N* per task)	**CA-TYP**	***F***	***p***
Spatial Reversal	(*N* = 17)[a]	(*N* = 18)		
# Correct	12.00 (6.23)	14.39 (4.72)	1.65	.21
# Perseverations	2.24 (2.71)	1.89 (3.20)	0.12	.73
# Failure to maintain sets	0.35 (0.70)	0.11 (0.32)	1.75	.20
# Sets	2.00 (1.28)	2.72 (1.36)	2.61	.12

[a]One child from the AUT group failed to complete the Spatial Reversal task due to frustration

Table 5

Correlations between EF measures, MA, and CA collapsing across all 4 groups in both experiments

Variable	MA	CA
Windows (Con.)	.44***	.12
Windows (Lib.)	.36**	.22***
Spatial Reversal: # Correct	.31**	.28*
Spatial Reversal: # Sets	.27*	.30*
Spatial Reversal: # Perseverations	−.32**	−.47***
Spatial Reversal: # Failure to maintain set errors	−.14	−.07

* $p < .05$

** $p < .01$

*** $p < .001$

† $p < .10$