Journal of the International Neuropsychological Society (2011), **17**, 522–530.
Copyright © INS. Published by Cambridge University Press, 2011.
doi:10.1017/S1355617711000312

Executive Function in Young Males with Klinefelter (XXY) Syndrome with and without Comorbid Attention-Deficit/Hyperactivity Disorder

Nancy Raitano Lee, Gregory L. Wallace, Liv S. Clasen, Rhoshel K. Lenroot, Jonathan D. Blumenthal, Samantha L. White, Mark J. Celano, and Jay N. Giedd

Child Psychiatry Branch, National Institute of Mental Health, NIH, Bethesda, Maryland

(Received April 5, 2010; Final Revision February 9, 2011; Accepted February 11, 2011)

Abstract

Deficits in executive function (EF) are reported to occur in individuals with Klinefelter syndrome (XXY). The degree of impairment, if any, is variable and the nature of these deficits has not been clearly elucidated in young males. In this report, we (a) examine EF skills using multiple tasks in a non-clinic referred group of youth with XXY, (b) describe the extent of EF weaknesses in XXY when this group is compared with typical males of a similar SES or typical males with similar verbal abilities, and (c) evaluate the contribution of comorbid attention-deficit/hyperactivity disorder (ADHD) to EF skills. The sample included 27 males with XXY (ages 9–25), 27 typically developing age- and vocabulary-matched males, and 22 age- and socioeconomic status-matched males. EF tasks included Verbal Fluency, the Trail Making Test, and the CANTAB Spatial Working Memory and Stockings of Cambridge tasks. Mixed model analysis of variance was used to compare the groups on EF tasks and revealed a main effect of group but no group by task interaction. Overall, the XXY group performed less well than both control groups, but performance did not differ significantly as a function of task. ADHD comorbidity in males with XXY was related to poorer EF skills. (*JINS*, 2011, *17*, 522–530)

Keywords: Sex chromosomes, Neuropsychology, Cognition, Attention Deficit Hyperactivity Disorder, Working memory, Problem solving

INTRODUCTION

The occurrence of an extra X chromosome in males, commonly referred to as Klinefelter syndrome or XXY, occurs in 1/500 to 1/1000 live male births (Bojesen, Juul, & Gravholt, 2003; Nielsen & Wohlert, 1990). Intellectual abilities of individuals with XXY generally fall within the average range (for a review, see Boada, Janusz, Hutaff-Lee, & Tartaglia, 2009; Geschwind, Boone, Miller, & Swerdloff, 2000), but are often lower than those of siblings or typically developing individuals matched on socioeconomic status (Bender, Linden, & Robinson, 1991; Rovet, Netley, Bailey, Keenan, & Stewart, 1995). Verbal IQ tends to be lower than performance IQ (Rovet, Netley, Keenan, Bailey, & Stewart, 1996), and both language delays (Stewart, Bailey, Netley, Rovet, & Park, 1986) and verbal learning disabilities are common (Bender, Linden, & Robinson, 1993; Pennington, Bender, Puck, Salbenblatt, & Robinson, 1982; Rovet et al., 1996).

Studies of executive function (EF), a collection of skills that includes working memory, planning, inhibition, and cognitive flexibility (Lezak, Howieson, & Loring, 2004; Miyake et al., 2000), are sparse in XXY, particularly in childhood, despite their importance in predicting academic outcomes in typical populations (Blair & Razza, 2007). Furthermore, research suggests that males with XXY experience heightened rates of attention-deficit/hyperactivity disorder (ADHD; Bruining, Swaab, Kas, & van Engeland, 2009), a disorder that is also associated with EF difficulties (Barkley, 1997). Thus, investigating the presence and nature of EF difficulties in males with XXY with and without ADHD was the focus of the current study.

Research on EF skills in adults with XXY has documented more significant EF difficulties on tasks with pronounced verbal demands. Deficits have been reported in verbal inhibition (DeLisi et al., 2005) and verbal working memory (Fales et al., 2003) as well as composites comprised of several different verbal EF tasks (Boone et al., 2001). Studies of verbal

Conflict of Interest Statement: No conflicts declared.

Correspondence and reprint requests to: Nancy Raitano Lee, National Institute of Mental Health; NIH, Bldg. 10, 4C110, MSC-1367, 10 Center Drive, Bethesda, MD 20892-1600. E-mail: lnancy@mail.nih.gov

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

fluency have been inconsistent, with one study documenting deficits (Bender, Linden, & Harmon, 2001) and one study failing to do so (DeLisi et al., 2005). For nonverbal EF tasks, such as the Wisconsin Card Sorting Test (WCST), two adult studies did not find deficits (Bender et al., 2001; DeLisi et al., 2005). In contrast, Boone et al. (2001) reported impaired performance on a composite of nonverbal EF tasks (including the WCST, design fluency, Rey Tangled Lines, and Emotional Situations). However, because performance on individual tasks included in the composite was not reported for the group as a whole (just XXY subgroups with different IQ profiles), it is unknown whether the XXY group was impaired on all tasks or only tasks with social–emotional and motor demands, two areas of documented weakness in XXY (Ross et al., 2008; van Rijn, Swaab, Aleman, & Kahn, 2006).

The few studies of EF in children with XXY report more consistent deficits on EF tasks with pronounced verbal and/or motor demands. Bender et al. (1993) examined EF skills in 11 adolescents with XXY and 25 mixed-sex (12 male) controls. They reported intact performance on the WCST, but impaired performance on the Trail Making test. In a study of three children with XXY (Temple & Sanfilippo, 2003), deficits on verbal inhibition were found; however, performance on other verbal EF tasks, including Verbal Fluency and Trail Making, was similar to controls. Performance on nonverbal EF tasks, including the Brixton, Tower of London, and Rey-Osterrieth Complex Figure Drawing, did not differ significantly from controls. A study of 50 boys with XXY (but no controls) suggested below average performance on a task of attention and low average to average performance on a task of verbal inhibition based on published norms (Ross et al., 2008). More recently, Ross, Zeger, Kushner, Zinn, and Roeltgen (2009) compared performance of 93 boys with XXY with an all male control group (and males with XYY) and reported no differences from controls on verbal inhibition, but higher rates of omission errors on a continuous performance task and poorer organization scores on the Rey-Osterrieth.

Given the scarcity of EF studies in young males with XXY, further characterization of EF skills using multiple tasks is warranted. Furthermore, choosing an appropriate control group and in particular, a sex-matched group, is an important consideration when describing a male-specific disorder. Because the literature suggests small male/female differences on certain neuropsychological tasks, including a male disadvantage on verbal fluency and male advantage on some visuospatial tasks (Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003; but see Hyde, 1990), comparing performance to an exclusively male control group is most appropriate. Moreover, given the language difficulties and lower verbal IQs associated with XXY and the fact that certain EF tasks have significant verbal demands (e.g., Verbal Fluency), including a control group matched on verbal abilities may help to control for linguistic contributions to (particularly verbal) EF deficits. Thus, comparing performance of males with XXY to typically developing males matched on verbal ability was another goal of the current research.

Abbreviations

XXY	Klinefelter syndrome
XY	Typically developing males
SES	Socioeconomic status
EF	Executive function
CANTAB	Cambridge Neuropsychological Test Automated Battery
SS	Standard Score

In addition, the current research sought to examine how the presence of comorbid ADHD contributed to performance on EF tasks, as several studies have suggested that males with XXY have heightened rates of ADHD or attentional difficulties (DeLisi et al., 2005; Theilgaard, 1984). One recent study of a self-selected sample of youth with XXY reported that 63% met criteria for ADHD (Bruining et al., 2009). Given the well-established links between ADHD and EF deficits (Barkley, 1997), males with XXY and comorbid ADHD may have more severe EF difficulties. Thus, we sought to test this possibility directly.

The current study sought to improve upon and add to the existing literature by (a) comparing the performance of males with XXY to two all-male control groups using multiple EF tasks, including two tasks that have not been previously studied in XXY (the Cambridge Neuropsychological Test Automated Battery Spatial Working Memory and Stockings of Cambridge tasks); and (b) examining the contribution of comorbid ADHD to EF skills in XXY.

Regarding the two control groups, the first was matched pairwise on sex, age, and vocabulary performance. We chose to match on verbal abilities to lessen the impact of possible verbal disparities between the XXY and control groups on (particularly verbal) EF task performance. This approach is similar to IQ matching that is often used in the literature examining EF deficits in clinical groups (Pennington & Ozonoff, 1996). The second group was matched on sex, age, and SES (group-wise) and served as a comparison to document how EF skills in males with XXY compare to typically developing peers who are likely to come from families in which the parents have similar levels of educational and occupational attainment (i.e., analogous to classmates). This group provided a clinical comparison that is relevant for families and professionals who serve children with XXY and also permitted an opportunity to demonstrate the extent of cognitive deficits associated with XXY, including the overall reduction in IQ that is associated with the syndrome. Furthermore, unlike the vocabulary-matched group, this group did not require participant matching based on a correlated aspect of the XXY cognitive phenotype. Matching on an aspect of a disorder's phenotype, such as IQ, has been discouraged by some researchers in the field (e.g., Dennis et al., 2009), and thus, we believed that it would be conservative to include an SES-matched group in addition to the vocabulary-matched group. Lastly, we examined how the presence of comorbid ADHD impacted EF performance by stratifying the XXY group into those with XXY only and those with

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

comorbid XXY + ADHD and comparing these groups to one another as well as the two control groups. We anticipated that the occurrence of comorbid ADHD would result in more significant EF skills in XXY; however, we sought to determine if this reduction in EF performance was task-specific or more general.

METHOD

Participants

A cohort of 59 individuals with non-mosaic XXY (confirmed by karyotyping) were recruited nationally with the help of two parent groups (Knowledge, Support & Action; American Association for Klinefelter Syndrome Information and Support) to participate in an ongoing study of brain development in XXY (Giedd et al., 2007). Of these participants, 33 had complete data on the EF tasks examined in this study and we were able to match 27 to a control group of males on vocabulary performance. These 27 participants were included in all primary analyses. In addition, follow-up analyses were completed with the full XXY sample ($n = 33$, including the 6 participants whose vocabulary scores were too low to match to XY males) to examine performance of the entire XXY sample regardless of verbal ability.

The two all-male control groups were recruited locally and nationally as a part of an ongoing longitudinal study of typical brain development in children from single and twin births (Giedd et al., 2009). The first group, called XY-VOC, included 27 XY males (12/27 were non-related twin pair members) who were matched pairwise on age (less than 2 years apart) and Vocabulary scores (within a standard deviation) to members of the XXY group. The second group, called XY-SES included 22 XY males (12/22 were non-related twin pair members) who were matched group-wise on age and SES to the XXY group. Because approximately half of the control participants were unrelated twins (co-twins and related family members were *not* included in control groups) and some research suggests that there is mild neurocognitive risk associated with twinning (Deary, Pattie, Wilson, & Whalley, 2005; but see Webbink, Posthuma, Boomsma, de Geus, & Visscher, 2008 suggesting a smaller or nonexistent effect), the following measures were taken. Twin and non-twin control participants were compared on IQ and EF tasks, and no statistically significant differences were found. However, a trend was noted for Verbal Fluency performance, such that twins received lower scores ($t[47] = 1.66$; $p = .10$). Thus, to control for any twin disadvantage in group comparisons, twin status was included as a covariate in secondary analyses.

Table 1 summarizes demographic information. For the XY-VOC and XXY comparisons, the groups did not differ on age, SES, or any of the IQ variables, as desired. For XY-SES and XXY comparisons, the groups did not differ on age or SES as desired; however, the XY-SES group did outperform the XXY group on all IQ measures with the exception of the Block Design subtest. The majority of participants in the sample were Caucasian (>90%), and the racial composition

Table 1. Means, *SD*s, Ranges, and Effect Sizes of Group Differences on Demographic Variables in the XXY and XY Control Groups

	XXY $n = 27$			XY-VOC group $n = 27$			XY-SES group $n = 22$			Group differences	Effect size (d)[c]	
	M	SD	Range	*M*	SD	Range	*M*	SD	Range	Statistic	XXY-VOC	XXY-SES
Age	16.08	4.25	9–25	16.16	4.35	9–26	16.43	4.00	10–24	$F(2,73) = .05, p > .90$	.02	.08
SES	53.19	23.14	20–114	52.81	17.22	20–89	50.68	18.26	20–89	$F(2,73) = .11, p > .80$	−.02	−.01
IQ	101.70	13.26	79–126	102.63	9.85	84–133	115.36	7.49	104–130	$F(2,73) = 12.20, p < .001$[a]	.08	1.32
VIQ	98.37	13.07	79–127	99.04	10.35	80–131	114.45	9.86	88–137	$F(2,73) = 15.29, p < .001$[a]	.05	1.40
PIQ	104.56	13.47	78–129	105.74	10.69	84–127	113.27	10.25	96–128	$F(2,73) = 3.88, p < .03$[a]	.10	.73
Vocabulary	9.26	2.74	5–15	9.44	2.53	6–16	13.14	1.91	8–16	$F(2,73) = 18.68, p < .001$[a]	.07	1.67
Block Design	11.22	2.61	6–15	11.11	2.36	7–15	12.59	2.82	8–19	$F(2,73) = 2.38, p > .09$	−.04	.50
Similarities	10.00	2.57	6–16	10.37	1.96	6–14	12.32	2.08	8–16	$F(2,73) = 7.34, p < .01$[a]	.16	1.00
Matrix Reasoning	10.56	2.90	2–15	11.19	2.11	6–15	12.33	1.48	10–15	$F(2,73) = 3.27, p < .05$[b]	.25	.81

[a]XXY, XY-VOC < XY-SES; $p < .05$.
[b]XXY < XY-SES; $p < .05$.
[c]Cohen (1988).

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

of the XXY and control groups did not differ significantly [χ^2 (2) = 1.66; $p > .43$]. With regard to SES, both the XXY and XY control groups came from families with relatively high levels of educational and occupational attainment. Seventy-six percent of the XXY sample, 72% of the XY-VOC sample, and 55% of the XY-SES sample had at least one parent with a college or advanced degree. Similarly, the vast majority of families had at least one parent who held a minimum of a managerial or semi-professional/professional position (most of which required at least a college education). This was the case for 73%, 64%, and 70% of the XXY, XY-VOC, and XY-SES groups, respectively. Consequently, many of the children in this study are likely to be more privileged than the average child living in the United States. Thus, it is not surprising that the IQ in the SES-matched control group (which was free to vary) was above average. Given that the XXY group's parents had similar (if not higher) levels of educational attainment, a similarly high IQ should be expected. However, consistent with other studies, the IQ of the XXY group was about a standard deviation below SES-matched typical peers (albeit in the average range). This IQ difference is consistent with earlier prospective studies of XXY in which participants were identified through consecutive newborn screenings (Bender, Puck, Salbenblatt, & Robinson, 1986; Ratcliffe, Butler, & Jones, 1990).

Informed consent was obtained from adult participants and parents of child participants, and verbal assent was obtained from child participants. The National Institute of Mental Health Institutional Review Board approved the protocol.

Measures

Intellectual Functioning was measured using the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). Because the participants were enrolled in a longitudinal study, IQ testing was not always completed at the same visit as the EF tests (mean difference ~2 years). However, IQ tests were generally completed within 4 years of EF testing (~88% of cases), and the XXY and control groups did not differ in the length of time between IQ and EF testing, $F(2,73) = 1.14$; $p > .33$.

Executive Functioning Tasks, including Verbal Fluency, Trail Making, Spatial Working Memory, and Stockings of Cambridge, are described in Table 2. The first two tasks were included because they have been examined in other studies of adults and children with XXY. However, it is important to note that they have significant verbal and motor demands, two areas of documented weakness in XXY. In contrast, the Spatial Working and Stockings of Cambridge tasks do not have significant verbal or motor requirements, and thus, may be better measures of EF skills in XXY.

Attention-Deficit/Hyperactivity Disorder (ADHD) Symptomatology was assessed using the Kiddie-SADS-Present and Lifetime Version (Kaufman, Brent, Rao, & Ryan, 1996), a semi-structured psychiatric interview. Ratings were made based on report of caregivers for children under 14 and by a combination of caregiver and self-report for adolescents/young adults. Children with XXY were characterized as meeting criteria for ADHD using the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria; 11 of the 27 children (41%) met criteria for either ADHD-Predominantly Inattentive type ($n = 10$) or ADHD-Combined Type ($n = 1$). The Kiddie-SADS was only completed with the XXY participants, as our control participants are screened to be free of psychiatric diagnoses. Thus, ADHD status was not free to vary in the control groups.

Socioeconomic Status (SES) was assessed using the Hollingshead Two-Factor Index (Hollingshead & Redlich, 1958), which provides a single score that is an index of the educational and occupational level of the primary breadwinner in the family.

Data Analyses

Data were inspected for normality and outliers and then the following analyses were completed. First, a 3×4 mixed-model analysis of variance (ANOVA) with one between subjects factor (Group) and one within subjects factor (EF task) was completed to examine group differences in the pattern of scores on EF tasks for the XXY group and the two control groups. Second, to evaluate the impact of having a subsample of unrelated twins in our control groups, a 3×4 mixed-model analysis of covariance (ANCOVA) was completed with twin status covaried. Third, follow-up ANOVA analyses were completed with the full XXY sample ($n = 33$, including the 6 participants whose vocabulary scores were too low to match to XY males) to compare the performance of the entire group to controls regardless of verbal ability. Fourth, males with XXY with and without comorbid ADHD were compared with control participants using a 4×4 mixed model ANOVA to examine the impact of comorbid ADHD on EF performance in males with XXY.

RESULTS

The results of the 3×4 mixed model ANOVA with one between-subjects factor (Group: XXY, XY-VOC, XY-SES) and one within-subjects factor (EF task: Verbal Fluency Standard Score [SS], Trail Making Part B SS, Stockings of Cambridge Perfect Solutions SS, Spatial Working Memory Forgetting Errors SS) revealed a main effect of EF task ($F[2,219] = 9.14$; $p < .001$) and a main effect of group ($F[2,73] = 10.18$; $p < .001$), but no group by task interaction ($F[6,219] = 0.12$; $p > .90$). Pairwise task comparisons (collapsed across group) revealed that participants received lower scores on both the Verbal Fluency and Trail Making tasks ($ps < .01$) than on the Stockings of Cambridge and Spatial Working Memory tasks overall. Pairwise group comparisons (collapsed across task) revealed that the XXY group performed less well than both the XY-VOC and XY-SES groups overall and the XY-VOC group performed less well than XY-SES group ($ps < .05$, FDR corrected; Benjamini & Hochberg, 1995). Means and *SD*s for the EF

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

Table 2. Executive Function Task Descriptions and Dependent Variables

Task	Task description	Dependent variables
Verbal Fluency	Participants are asked to name as many words as they can in 60 seconds according to phonemic (F, A, S) or semantic (animals) cues.	The phonemic fluency and semantic fluency raw scores were converted to age-based standard scores and were averaged to yield one score – Verbal Fluency SS[a].
Trail Making Test	Participants are asked to draw lines between encircled numbers in order (Part A) and to alternate between connecting encircled numbers and letters arranged on a page in order (Part B). Two forms were given: the child intermediate form (15 encircled characters; 9–14 years) and the adult form (25 encircled characters; 15+ years).	Trail Making Part B performance was used as the dependent variable. Raw time to complete Part B was converted to an age-based standard score called Trails B SS[b].
CANTAB Spatial Working Memory (SWM)	In this computerized self-ordered search test that assesses the ability to retain spatial information and to manipulate items in working memory, participants are required to search through colored boxes to find a blue token. When the participant touches the box, it opens to reveal its contents. The aim is to find the blue tokens and use them to fill an empty column on the right side of the screen. Participants are told that tokens are hidden under each box only once during a trial. Thus, returning to an already-sampled box during a trial is what we will call a "forgetting error," though this is referred to as a "between search" error by the CANTAB. Participants completed four practice trials with three boxes. Testing consisted of four trials with 4, 6, and 8 boxes.	The primary dependent variable utilized was the number of 'forgetting errors' on all trials. This score was converted to an age-based standard score called Spatial Working Memory SS[c].
CANTAB Stockings of Cambridge (SOC)	In this computerized task of spatial planning based upon the Tower of London, participants are shown two displays (as a top-bottom split screen) containing stacks of colored balls held in stockings and instructed to arrange the balls in the bottom display so that they look like the top display using the minimum number of "moves" (from 2–5). Six 1–2 move practice items are followed by test items: two 2-move items, two 3-move items, four 4-move items, and four 5-move items.	The primary dependent variable for this task was the number of items solved using an optimal solution. This score was converted to an age-based SS called Stockings of Cambridge SS[c].

[a]Age-based *Z*-scores were created based on means taken from Gaddes & Crockett (1975), Halperin, Healey, Zeitchik, Ludman, & Weinstein (1989), and Tombaugh, Kozak, and Rees (1999) and were converted to standard scores (SS) with a mean of 100 and *SD* of 15. For ages 13, 14, and 15, group means were interpolated based on available age norms.
[b]Age-based *Z*-scores were created based on means taken from Spreen and Gaddes (1969) and Tombaugh, Rees & McIntyre (1998) and were converted to SS.
[c]Cambridge Neuropsychological Test Automated Battery; Age-based *Z*-scores were created based on means provided by Cambridge Cognition Limited (1999) and were converted to SS.

tasks can be found in Table 3. A follow-up 3 × 4 ANCOVA was run with twin status covaried and the results were similar such that the main effect of group was maintained ($F[2,72] = 9.17$; $p < .001$). These results indicated that the XXY group performed less well than both control groups, but the pattern of scores on these tasks did not differ significantly as a function of group.

Next, analyses were run to (a) examine how the complete group of participants with XXY ($n = 33$), including those with the lowest vocabulary scores, compared with controls on the EF tasks, and (b) rule out the possibility that our findings for the XXY and XY-VOC groups were due to regression effects on the EF scores that can occur when matching two subgroups of participants with extreme scores on another measure (i.e., the most verbally talented XXY males and the least verbally talent XY males). These analyses were run with the complete sample of XXY males with all EF measures and the XY-SES group only. The 2 × 4 repeated measures ANOVA revealed a main effect of group ($F[1,53] = 21.33$; $p < .001$) but no group X task interaction ($F[3,159] = 0.42$; $p > .70$). (Means and *SD*s for the whole XXY group were quite similar to the smaller XXY subsample and are reported in the

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

Table 3. Means, Standard Deviations (*SD*), and Effect Sizes (*d*) of Group Differences on Executive Function Task Standard Scores

	XXY $n = 27$[a]		XY-VOC $n = 27$		XY-SES $n = 22$		Effect size (d)[b]	
	M	*SD*	*M*	*SD*	*M*	*SD*	XXY-IQ	XXY-SES
Verbal Fluency SS	88.81	12.74	95.82	14.15	102.39	8.74	0.52	1.26
Trail Making Part B SS	91.36	19.70	97.85	12.41	106.17	15.77	0.40	0.84
Stockings of Cambridge SS	97.99	19.60	105.32	19.26	112.62	16.18	0.38	0.82
Spatial Working Memory SS	99.27	17.61	105.24	14.69	110.16	13.34	0.37	0.70

[a]For the whole unmatched XXY sample ($n = 33$), *M*s (*SD*s) for Verbal Fluency, Trail Making Part B, Stockings of Cambridge, and Spatial Working Memory were as follows, respectively: 87.20 (12.85), 89.95 (20.47), 96.31 (19.84), 99.12 (19.34).
[b]Cohen (1988).

footnote of Table 3.) Follow-up 2 × 4 ANCOVA analyses revealed that the main effect of group was maintained when vocabulary was covaried (F[1,52] = 5.91; $p < .02$) and when Verbal and Performance IQ were covaried (F[1,51] = 8.03; $p < .01$). These results lessen concerns about regression effects accounting for our pattern of EF findings. Moreover, they provide further support that young males with XXY have EF deficits, and these deficits cannot fully be accounted for by verbal weaknesses or overall reductions in IQ.

Lastly, the effect of comorbid ADHD status was examined. Based upon the Kiddie-SADs interview, 41% of the XXY group met DSM-IV criteria for ADHD. Thus, the group was stratified into those with (XXY + ADHD; $n = 11$) and without (XXY only; $n = 16$) ADHD, and ANOVA analyses were re-run. The 4 × 4 mixed model ANOVA with one between subjects factor (Group: XXY + ADHD, XXY only, XY-VOC, XY-SES) and one within subjects factor (EF Task: Verbal Fluency SS, Trail Making Part B SS, Stockings of Cambridge Perfect Solutions SS, Spatial Working Memory Errors Standard Score) revealed a main effect of task (F[3,216] = 7.81; $p < .001$), a main effect of group (F[3,72] = 9.34; $p < .001$), but no group by task interaction (F[9,216] = 0.62; $p > .77$). Pairwise group comparisons (collapsed across all EF tasks) revealed that the XXY + ADHD group performed less well than the other three groups (including their peers with XXY without comorbid ADHD; all *ps* < .05; FDR corrected). Both the XXY only and XY-VOC groups performed less well than the XY-SES group but did not differ significantly from one another. See Figure 1. (Results were maintained after adding twin status as a covariate.)

Furthermore, when vocabulary was included as a covariate, the main effect of group was maintained, and the XXY + ADHD group continued to perform less well on EF tasks than the other three groups (*ps* < .05, uncorrected; with FDR correction, XXY + ADHD *vs.* XXY only comparison

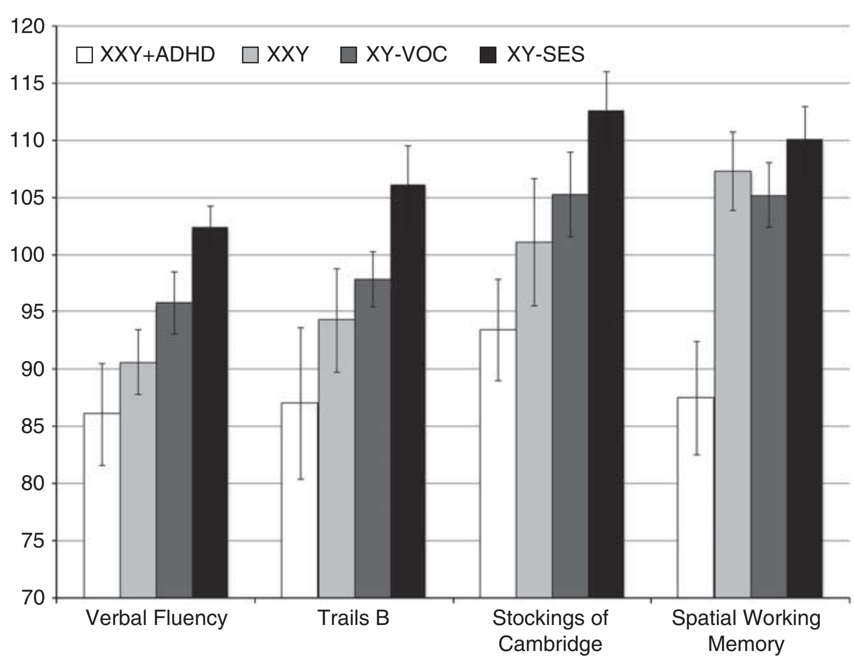

Fig. 1. Mean standard scores (with standard error bars) for executive function tasks for XXY + ADHD (white bars), XXY-only (light gray bars), XY-VOC (dark gray bars), and XY-SES (black bars groups). Standard scores have a mean of 100 and SD of 15. Please see manuscript text for results of statistical analyses.

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

falls short of statistical significance). Finally, when Verbal and Performance IQ were substituted as a covariates, the main effect of group was also maintained and the XXY + ADHD group continued to perform less well than the other three groups ($ps < .05$, uncorrected; with FDR correction, XXY + ADHD *vs.* XXY only comparison falls short of statistical significance).

DISCUSSION

This study examined the performance of males with XXY on several EF tasks, including Verbal Fluency, Trail Making, and the CANTAB Stockings of Cambridge and Spatial Working Memory tasks. The XXY group performed significantly less well than two typically developing male control groups, one matched on socioeconomic status (XY-SES group) and one matched on approximate verbal ability level (XY-VOC group), on EF tasks overall. Effect sizes of EF differences relative to the XY-SES group were medium to large and were similar in magnitude to those found for Verbal and Nonverbal IQ when comparing these two groups in the present study. When the XXY group was compared with the XY-VOC group, a significant XXY EF disadvantage was also found; however, the effect sizes for individual tasks were considerably smaller.

Comparisons of the XXY groups with and without ADHD to the control groups indicated that the comorbid group was the most impaired, even performing worse than their peers with XXY only. These findings suggest that it may be beneficial for parents and teachers to target children with XXY and comorbid attentional difficulties early to encourage the development of EF, as research suggests that early EF skills are malleable (Diamond, Barnett, Thomas, & Munro, 2007) and strong predictors of academic outcome (Blair & Razza, 2007). Furthermore, consultation with a psychiatrist or other medical professional may be considered to evaluate the appropriateness of a stimulant trial (or other appropriate medication). Research in chromosomally typical males with ADHD has demonstrated improvements in spatial working memory following the introduction of methylphenidate (Bedard, Martinussen, Ickowicz, & Tannock, 2004). Thus, treating ADHD symptoms may have implications for psychosocial, academic, and occupational outcomes for males with XXY.

This study adds to the literature on the cognitive phenotype in youth with XXY in several ways. First, it documents deficits in EF for young males with XXY using (a) a larger non-clinic referred sample than has been typically reported in the literature, and (b) two carefully selected all-male control groups. Second, this study used two CANTAB tasks that have not been used in studies of XXY before. These tasks may be useful for other investigators to consider when studying XXY, as they do not have significant verbal, motor, or processing speed demands. Lastly, this is the first study to examine the impact of comorbid ADHD on EF in youth with XXY, providing additional information about what factors influence variability in the XXY cognitive phenotype.

LIMITATIONS

Before closing, we will discuss the limitations of the current study. First, this study relied on recruitment of children from parent support groups rather than children who were prospectively identified through newborn screenings. As a consequence, this study included children who were diagnosed both prenatally ($n = 11$; 41%) and postnatally. Given that genetic screenings are not routinely completed in the United States and epidemiological studies suggest that many individuals with XXY go undiagnosed or are not diagnosed until adulthood (Bojesen et al., 2003), it may be that our study's results were influenced by an ascertainment bias, such that only parents who were concerned about their child's cognitive functioning participated. This appears unlikely for two reasons. First, we compared the effect size of IQ score differences between our XXY and XY-SES group to those included in four prospectively studied samples (Bender et al., 1986; Graham, Bashir, Stark, Silbert, & Walzer, 1988; Netley & Rovet, 1984; Ratcliffe, Masera, Pan, & McKie, 1994) using Cohen's d (Cohen, 1988). The current study's effect sizes of 1.40 and 0.73 for Verbal and Nonverbal IQ, respectively, fall within the 95% confidence interval of the mean effect sizes for prospective studies, which are 1.37 (95% confidence interval [CI] = 0.53–2.22) and 0.87 (95% CI = 0.06–.1.68), respectively. Second, we compared EF scores for the pre- and postnatally diagnosed XXY groups and found no significant group differences. Thus, we do not believe that our findings are driven by a subgroup of postnatally diagnosed participants with significant cognitive deficits.

Another limitation of this research is that both the XXY and control groups come from relatively privileged socioeconomic backgrounds and may not be representative of the average child with XXY or the average typically developing child in the United States. We have made an effort to deal with this limitation by comparing children with XXY to children who come from similarly privileged backgrounds. This at least evens the playing field, so to speak, and allows a comparison of males with XXY to males without an additional X chromosome who come from families with similar educational (and likely cognitive ability) levels. Thus, while performance of the XXY group fell within the average range on most tasks, including EF tasks, scores were approximately a half to a whole standard deviation below controls. While these may be considered statistically medium-large effects, these effects are not of the magnitude reported for disorders such as autism or intellectual disability where the clinical impact of EF deficits is likely to be far greater. Thus, evaluating the extent of EF deficits in XXY in the larger context of research on individuals with a range of developmental disabilities may be helpful. Furthermore, it is important to note that while overall EF weaknesses are somewhat mild in the XXY group, there is great variability in EF performance within the XXY group, such that some individuals are more significantly impacted than others. Furthermore, youth with comorbid XXY + ADHD appear to have the greatest EF deficits, and thus, may be most in need of support and intervention at school and in the community.

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

Lastly, while the XXY group in the current study performed less well than both control groups on tasks commonly referred to as EF tasks, it is essential to acknowledge that these tasks, particularly Verbal Fluency and Trail Making, place multiple cognitive demands on participants, demands that are likely to fall outside of what is traditionally considered the EF domain. For instance, Verbal Fluency and Trail Making have significant speed and motor/oral–motor components. Thus, a general processing speed or motor deficit could be contributing to lower task performance. Furthermore, performance on Verbal Fluency is likely to be influenced by dyslexia and other language-based learning difficulties that were not evaluated in this study. Thus, phonological processing or other linguistic processes could be contributing to performance deficits.

FUTURE DIRECTIONS AND CONCLUSIONS

Further research is needed to isolate the underlying processes related to EF task performance in XXY. This will likely best be accomplished using experimental manipulation of language, motor, and speed demands in EF tasks. Furthermore, research linking EF skills to their corresponding neural substrates in XXY is desired. Structural (DeLisi et al., 2005; Giedd et al., 2007; Warwick et al., 1999) and functional (Steinman, Ross, Lai, Reiss, & Hoeft, 2009; van Rijn et al., 2008) neuroimaging studies are under way. However, additional research may refine our understanding of the neuroanatomical/physiological underpinnings of deficits on EF tasks possessed by youth with XXY and those with attentional, linguistic, and processing speed deficits more generally. Finer-grained descriptions of the morphometric and physiological underpinnings of EF weaknesses and other neurocognitive dysfunction in XXY may be used as endophenotypes in genetic expression studies of candidate genes on the X chromosome thought to be associated with the cognitive deficits characteristic of the disorder. Such studies have implications not only for individuals with XXY, but also for individuals with other causes of learning, attentional, and intellectual disability deficits.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research Program of the NIH, National Institute of Mental Health. We thank the children and families who volunteered their time for this research.

REFERENCES

Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65–94.

Bedard, A.C., Martinussen, R., Ickowicz, A., & Tannock, R. (2004). Methylphenidate improves visual-spatial memory in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*, 260–268.

Bender, B.G., Linden, M.G., & Harmon, R.J. (2001). Neuropsychological and functional cognitive skills of 35 unselected adults with sex chromosome abnormalities. *American Journal of Medical Genetics*, *102*, 309–313.

Bender, B.G., Linden, M., & Robinson, A. (1991). Cognitive and academic skills in children with sex chromosome abnormalities. *Reading and Writing*, *3*, 315–327.

Bender, B.G., Linden, M.G., & Robinson, A. (1993). Neuropsychological impairment in 42 adolescents with sex chromosome abnormalities. *American Journal of Medical Genetics*, *48*, 169–173.

Bender, B.G., Puck, M.H., Salbenblatt, J.A., & Robinson, A. (1986). Dyslexia in 47, XXY boys identified at birth. *Behavior Genetics*, *16*, 343–354.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289–300.

Blair, C., & Razza, R.P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663.

Boada, R., Janusz, J., Hutaff-Lee, C., & Tartaglia, N. (2009). The cognitive phenotype in Klinefelter syndrome: A review of the literature including genetic and hormonal factors. *Developmental Disabilities Research Reviews*, *15*, 284–294.

Bojesen, A., Juul, S., & Gravholt, C.H. (2003). Prenatal and postnatal prevalence of Klinefelter syndrome: A national registry study. *Journal of Clinical Endocrinology and Metabolism*, *88*, 622–626.

Boone, K.B., Swerdloff, R.S., Miller, B.L., Geschwind, D.H., Razani, J., Lee, A., … Paul, L. (2001). Neuropsychological profiles of adults with Klinefelter syndrome. *Journal of the International Neuropsychological Society*, *7*, 446–456.

Bruining, H., Swaab, H., Kas, M., & van Engeland, H. (2009). Psychiatric characteristics in a self-selected sample of boys with Klinefelter syndrome. *Pediatrics*, *123*, e865–e870.

Cambridge Cognition Limited. (1999). *Cambridge neuropsychological test automated battery*. Cambridge: Cambridge Cognition.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deary, I.J., Pattie, A., Wilson, V., & Whalley, L.J. (2005). The cognitive cost of being a twin: Two whole-population surveys. *Twin Research and Human Genetics*, *8*, 376–383.

DeLisi, L.E., Maurizio, A.M., Svetina, C., Ardekani, B., Szulc, K., Nierenberg, J., … Harvey, P.D. (2005). Klinefelter's syndrome (XXY) as a genetic model for psychotic disorders. *American Journal of Medical Genetics Part B*, *135*, 15–23.

Dennis, M., Francis, D.J., Cirino, P.T., Schachar, R., Barnes, M.A., & Fletcher, J.M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*, 331–343.

Diamond, A., Barnett, W.S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*, 1387–1388.

Fales, C.L., Knowlton, B.I., Holyoak, K.J., Geschwind, D.H., Swerdloff, R.S., & Gonzalo, I.G. (2003). Working memory and relational reasoning in Klinefelter syndrome. *Journal of the International Neuropsychological Society*, *9*, 839–846.

Gaddes, W.H., & Crockett, D.J. (1975). The Spreen-Benton aphasia tests, normative data as a measure of normal language development. *Brain and Language*, *2*, 257–280.

Geschwind, D.H., Boone, K.B., Miller, B.L., & Swerdloff, R.S. (2000). Neurobehavioral phenotype of Klinefelter syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*, 107–116.

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press

Giedd, J.N., Clasen, L.S., Wallace, G.L., Lenroot, R.K., Lerch, J.P., Wells, E.M., … Samango-Sprouse, C. (2007). XXY (Klinefelter syndrome): A pediatric quantitative brain magnetic resonance imaging case-control study. *Pediatrics*, *119*, e232–e240.

Giedd, J.N., Lalonde, F.M., Celano, M.J., White, S.L., Wallace, G.L., Lee, N.R., & Lenroot, R.K. (2009). Anatomical brain magnetic resonance imaging of typically developing children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*, 465–470.

Graham, J.M. Jr., Bashir, A.S., Stark, R.E., Silbert, A., & Walzer, S. (1988). Oral and written language abilities of XXY boys: Implications for anticipatory guidance. *Pediatrics*, *81*, 795–806.

Halperin, J.M., Healey, J.M., Zeitchik, E., Ludman, W.L., & Weinstein, L. (1989). Developmental aspects of linguistic and mnestic abilities in normal children. *Journal of Clinical and Experimental Neuropsychology*, *11*, 518–528.

Hollingshead, A.B., & Redlich, F.C. (1958). *Social class and mental illness*. New York: Wiley.

Hyde, J.S. (1990). Meta-analysis and the psychology of gender differences. *Journal of Women in Culture and Society*, *16*, 55–73.

Kaufman, J.B., Brent, D., Rao, U., & Ryan, N. (1996). *Kiddie-Sads-Present and Lifetime Version* (K-SADS-PL).

Lezak, M.D., Howieson, D.N., & Loring, D.W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Netley, C., & Rovet, J. (1984). Hemispheric lateralization in 47, XXY Klinefelter's syndrome boys. *Brain and Cognition*, *3*, 10–18.

Nielsen, J., & Wohlert, M. (1990). Sex chromosome abnormalities found among 34,910 newborn children: Results from a 13-year incidence study in Arhus, Denmark. *Birth Defects Original Articles Series*, *26*, 209–223.

Pennington, B.F., Bender, B.G., Puck, M., Salbenblatt, J., & Robinson, A. (1982). Learning disabilities in children with sex chromosome anomalies. *Child Development*, *53*, 1182–1192.

Pennington, B.F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*, 51–87.

Ratcliffe, S.G., Butler, G.E., & Jones, M. (1990). Edinburgh study of growth and development of children with sex chromosome abnormalities. IV. *Birth Defects Original Articles Series*, *26*, 1–44.

Ratcliffe, S.G., Masera, N., Pan, H., & McKie, M. (1994). Head circumference and IQ of children with sex chromosome abnormalities. *Developmental Medicine and Child Neurology*, *36*, 533–544.

Ross, J.L., Roeltgen, D.P., Stefanatos, G., Benecke, R., Zeger, M.P., Kushner, H., … Zinn, A.R. (2008). Cognitive and motor development during childhood in boys with Klinefelter syndrome. *American Journal of Medical Genetics A*, *146*, 708–719.

Ross, J.L., Zeger, M.P., Kushner, H., Zinn, A.R., & Roeltgen, D.P. (2009). An extra X or Y chromosome: Contrasting the cognitive and motor phenotypes in childhood in boys with 47, XYY syndrome or 47. XXY Klinefelter syndrome. *Developmental Disabilities Research Reviews*, *15*, 309–317.

Rovet, J., Netley, C., Bailey, J., Keenan, M., & Stewart, D. (1995). Intelligence and achievement in children with extra X aneuploidy: A longitudinal perspective. *American Journal of Medical Genetics*, *60*, 356–363.

Rovet, J., Netley, C., Keenan, M., Bailey, J., & Stewart, D. (1996). The psychoeducational profile of boys with Klinefelter syndrome. *Journal of Learning Disabilities*, *29*, 180–196.

Spreen, O., & Gaddes, W.H. (1969). Developmental norms for 15 neuropsychological tests age 6 to 15. *Cortex*, *5*, 170–191.

Steinman, K., Ross, J., Lai, S., Reiss, A., & Hoeft, F. (2009). Structural and functional neuroimaging in Klinefelter (47,XXY) syndrome: A review of the literature and preliminary results from a functional magnetic resonance imaging study of language. *Developmental Disabilities Research Reviews*, *15*, 295–308.

Stewart, D.A., Bailey, J., Netley, C.T., Rovet, J.F., & Park, M.E. (1986). Growth and development of midadolescent children with X and Y chromosome aneuploidy: The Toronto study. In S.P. Ratcliffe (Ed.), *Sex chromosome aneuploidy: Follow-up studies* (Vol. 22, pp. 119–182). New York: Alan R. Liss.

Temple, C.M., & Sanfilippo, P.M. (2003). Executive skills in Klinefelter's syndrome. *Neuropsychologia*, *41*, 1547–1559.

Theilgaard, A. (1984). A psychological study of the personalities of XYY- and XXY-men. *Acta Psychiatrica Scandinavica*, *69*, 1–133.

Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*, 167–177.

Tombaugh, T.N., Rees, L., & McIntyre, N. (1998). Normative data for the Trail Making Test. Personal communication cited in Spreen and Strauss. *A compendium of neuropsychological tests: Administration, norms and commentary* (2nd ed.). New York: Oxford University Press.

van Rijn, S., Aleman, A., Swaab, H., Vink, M., Sommer, I., & Kahn, R.S. (2008). Effects of an extra X chromosome on language lateralization: An fMRI study with Klinefelter men (47,XXY). *Schizophrenia Reseach*, *101*, 17–25.

van Rijn, S., Swaab, H., Aleman, A., & Kahn, R.S. (2006). X Chromosomal effects on social cognitive processing and emotion regulation: A study with Klinefelter men (47,XXY). *Schizophrenia Research*, *84*, 194–203.

Warwick, M.M., Doody, G.A., Lawrie, S.M., Kestelman, J.N., Best, J.J., & Johnstone, E.C. (1999). Volumetric magnetic resonance imaging study of the brain in subjects with sex chromosome aneuploidies. *Journal of Neurology, Neurosurgery, and Psychiatry*, *66*, 628–632.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychological Corporation/Harcourt Brace Jovanovich.

Webbink, D., Posthuma, D., Boomsma, D.I., de Geus, E.J.C., & Visscher, P.M. (2008). Do twins have lower cognitive ability than singletons? *Intelligence*, *36*, 539–547.

Weiss, E.M., Kemmler, G., Deisenhammer, E.A., Fleischhacker, W.W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, *35*, 863–875.

https://doi.org/10.1017/S1355617711000312 Published online by Cambridge University Press