PAPER

Deficits in volitional oculomotor control align with language status in autism spectrum disorders

David J. Kelly, Robin Walker and Courtenay Frazier Norbury

Department of Psychology, Royal Holloway, University of London, UK

Abstract

Eye-tracking paradigms are increasingly used to investigate higher-level social and cognitive processing in autism spectrum disorder (ASD). However, the integrity of the oculomotor system within ASD is unclear, with contradictory reports of aberrant eye-movements on basic oculomotor tasks. The purpose of the current study was to determine whether reducing population heterogeneity and distinguishing neurocognitive phenotypes can clarify discrepancies in oculomotor behaviour evident in previous reports. Reflexive and volitional eye-movement control was assessed in 73 children aged 8–14 years from four distinct groups: Autism Language Normal (ALN), Autism Language Impaired (ALI), non-autistic Language Impaired (LI) and Typically Developing (TD). Eye-movement control was measured using pro- and antisaccade tasks and a novel search distracter task to measure distractibility. Reflexive eye-movements were equivalent across groups, but deficits in volitional eye-movement control were found that aligned with language status, and were not specific to ASD. More than 80% of ALI and LI children presented error rates at least 1.5 SDs below the TD mean in an antisaccade task. In the search distracter task, 35.29% of ALI children and 43.75% of LI children had error rates greater than 1.5 SDs compared with 17.64% of ALN children. A significant proportion of children with neurodevelopmental disorders involving language function have pronounced difficulties suppressing reflexive saccades and maintaining fixations in the presence of competing stimuli. We extend the putative link between ALI and LI populations to non-language tasks, and highlight the need to account for co-morbidity in understanding the ontogenesis of ASD.

Introduction

Autism spectrum disorders (ASD) are neurodevelopmental disorders defined by socio-communicative impairments, and restricted and repetitive interests and behaviours. Furthermore, ASD is characterized by atypicalities in joint attention and mutual eye gaze (Nation & Penny, 2008), and reduced visual attention for social stimuli, such as human faces (Falck-Ytter & von Hofsten, 2011). Language abilities in ASD are variable and may range from essentially non-verbal to scores on standard measures of structural language within the normal range, with a substantial proportion of cognitively able individuals experiencing co-morbid language deficits that are similar to those observed in non-autistic children with language impairment (LI: Lindgren, Folstein, Tomblin & Tager-Flusberg, 2009; Loucas, Charman, Pickles, Simonoff, Chandler, Meldrum & Baird, 2008). Against this backdrop of phenotypic heterogeneity, eye-movements provide a convenient measure of online cognitive processing (Rayner, 1998) and may be particularly advantageous when working with ASD populations as eye-tracking is non-invasive and does not typically require a verbal or other overt response. Consequently, eye-tracking is becoming an increasingly popular method to probe cognition and social understanding in ASD (Benson & Fletcher-Watson, 2011; Boraston & Blakemore, 2007). Despite this surge in research using eye-tracking methods to probe a range of cognitive processes, relatively few studies have explored the ability of ASD individuals to control volitional eye-movements, which are crucial to interpreting performance in most cognitive processing tasks. Therefore, conclusions about social preferences and higher-order cognitive skills are being made without full reference to the functionality of the underlying oculomotor system, or the influence of co-morbid deficits such as language impairment on volitional eye-movement control. Thus, the aim of the current study was to measure reflexive and volitional eye-movement capabilities in sub-groups of children with ASD and a matched LI comparison group in order to better understand the specificity of oculomotor anomalies within ASD and the potential sources of individual differences in eye-movement control.

Previous eye-tracking studies have revealed atypical eye-movement patterns in ASD in tasks related to social...
processing such as face scanning and emotion recognition (Kirchner, Hatris, Heekeren & Dziobek, 2011; Klin, Jones, Schultz, Volkmar & Cohen, 2002; Pelphrey, Sasse, Reznick, Paul, Goldman & Piven, 2002; Riby & Hancock, 2008), theory of mind (Senju, Southgate, White & Frith, 2009) and complex scene processing (Benson, Piper & Fletcher-Watson, 2009; Freeth, Foulsham & Chapman, 2011; OHearn, Lakusta, Schroer, Minshew & Luna, 2011). However, there remains considerable debate over the universality of scanning atypicalities across ASD individuals, with some authors reporting typical viewing patterns to social stimuli, including faces (Fletcher-Watson, Leekam, Benson, Frank & Findlay, 2009) and others reporting that atypicalities in social viewing preferences may only characterize a sub-group of the ASD population (Norbury, Brock, Cragg, Einav & Nation, 2009; Pierce, Conant, Hazin, Desmond & Stoner, 2011). A possible explanation for the disparity between empirical reports concerns the integrity of the underlying oculomotor system. If meaningful conclusions are to be drawn from eye-movement studies that attempt to tap higher-level socio-cognitive processing, it is essential to know whether the oculomotor system is functioning normally within ASD populations. However, here too inconsistencies exist between reported findings.

Initial investigations found reduced saccade velocity and landing accuracy (Rosenhall, Johansson & Gillberg, 1998) in individuals with ASD relative to TD peers, in addition to increased saccade activity to a blank screen between stimulus presentations (Kemner, Verbaten, Cuperus, Camfferman & Van Engeland, 1998). Measurement of prosaccades has been used to test the ability of participants with ASD to generate reflexive, visually triggered saccades from a central fixation point (FP) to a peripheral target. In general, individuals with ASD do not have reduced saccadic reaction time (SRT) to peripheral targets (Luna, Doll, Hegedus, Minshew & Sweeney, 2007; Minshew, Luna & Sweeney, 1999; Sceif, Karmiloff-Smith, Campos, Elsabbagh, Driver & Cornish, 2005; Van Der Geest, Kemner, Camfferman, Verbaten & Van Engeland, 2001; see Goldberg, Lasker, Zee, Garth, Tien & Landa, 2002, for a report of longer SRTs in ASD) and display cueing effects to both arrows and eye-gaze cues (Kuhn, Benson, Fletcher-Watson, Kovshoff, McCormick, Kirkby & Leekham, 2010). In a variation of the prosaccade paradigm, removing the central FP prior to appearance of the peripheral target (gap task) reduces SRT compared to the situation with a continuously displayed fixation stimulus (Sceif et al., 2005; Saslow, 1967). The difference in SRT between the gap-overlap conditions (gap effect) has been attributed to the disengagement of attention (Fischer & Weber, 1993), the release of low-level fixation mechanisms (Munoz & Wurtz, 1992) and higher-level warning-signal effects (Reuter-Lorenz, Oonk, Barnes & Hughes, 1995). Only two studies have investigated the gap effect in ASD: one found no differences between individuals with ASD and the typical developing comparison group (Goldberg et al., 2002), while the other found a smaller gap effect in their ASD sample (Van Der Geest et al., 2001). Studies of infant siblings at genetic risk of ASD have found impairments in both SRT and a reduction in the gap effect, suggesting that reduced attentional control may be an early marker of ASD that subsequently derails social-communicative development (Elsabbagh, Volein, Holmboe, Tucker, Csibra, Baron-Cohen, Bolton, Charman, Baird & Johnson, 2009). In contrast to the prosaccade task, the antisaccade task involves higher-level cognitive processes as participants are required to inhibit a response to the peripheral target and instead look to the contra-lateral location. The antisaccade task is often used to probe top-down executive control of eye-movements (Munoz & Everling, 2004). Across studies, individuals with ASD make more directional, prosaccade errors (i.e. looks to the peripheral stimulus) relative to TD comparison groups (Minshew et al., 1999; Van Der Geest et al., 2001) and error rates do not appear to improve across developmental time (Minshew et al., 1999). However, deficits on the antisaccade task are evident in a number of other neurodevelopmental disorders, notably Fragile X syndrome (Sceif et al., 2005) and ADHD (Munoz, Armstrong, Hampton & Moore, 2003). No previous study of oculomotor control in ASD has included a non-ASD comparison group with developmental disorder; thus it is unclear whether reported deficits are specific to ASD or are reflective of co-morbid pathology.

Individuals with ASD are at increased risk of co-morbid diagnoses. In population studies, approximately 28% of children with ASD also meet criteria for ADHD (Simonoff, Pickles, Charman, Chandler, Loucas & Baird, 2008) and 48% meet criteria for LI (Loucas et al., 2008). The role of additional language impairment (LI) on behavioural presentation is particularly relevant as individuals with autism and additional language impairments (ALI) are thought to represent a distinct neurocognitive phenotype which shares etiological and neurobiological risk factors with LI (Tager-Flusberg & Joseph, 2003; Tomblin, 2011; Vernes, Newbury, Abrahams, Winchester, Nicod, Groszer, Alarcon, Oliver, Davies, Geschwind, Monaco & Fisher, 2008). With regard to eye-movements, direct comparisons of individuals with ASD who do and do not have LI reveal striking group differences. For example, Takarae, Minshew, Luna and Sweeney (2004) divided ASD participants into two groups according to whether they had exhibited delayed language development. Interestingly, although no differences were found for saccade peak velocity or SRT, children with language delays showed increased variance in saccade accuracy, leading the authors to conclude that a motor deficit underpinned oculomotor difficulties in this group. Using a task requiring higher-level processing, Norbury et al. (2009) reported fewer fixations to the eye regions of scene protagonists in dynamic social stimuli, but only for participants with ASD who had
normal range language abilities (ALN); peers with ALI did not differ from TD children. Although these findings might appear contradictory, Norbury and colleagues report further that their ALN population presented elevated levels of restricted and repetitive interests and behaviours relative to the ALI group, which may have influenced the viewing goals of the ALN group. At present, the causal connection between language and oculomotor behaviour is unclear; it has been suggested that early deficits in volitional control of eye-movements interfere with the establishment of joint attention in ASD, with cascading effects on language acquisition (Brenner, Turner & Muller, 2008). Alternatively, aberrant development of the neural circuits that support language, attention and motor capacities may confer vulnerability to both language and oculomotor development. Finally, impaired language development may disrupt executive control of eye-movements, resulting in differences in volitional eye-movements or increased distractibility when there are multiple distractions in the visual field. Whatever the causal pathway, in order to make sense of eye-tracking studies tapping higher-order social and cognitive processes, it is essential to establish basic oculomotor control and how this may be differentially affected within ASD by distinct neurocognitive phenotypes.

For instance, eye-movement studies that demonstrate reduced fixation time to eyes or faces in ASD suggest that these differences arise because of reduced interest in social stimuli (see Rice, Moriuchi, Jones & Klin, 2012). An alternative explanation may be that in the presence of competing visual stimuli, individuals with ASD are unable to maintain fixation on key aspects of the scene. Similarly, it has been reported that individuals with ASD do not modulate visual scanning patterns in response to task demands (Benson et al., 2009). It may be that individuals with ASD have the appropriate goal in mind, but that scanning is compromised by inefficient control of eye-movements. Most previous studies have included participants with ASD who score within the average range on a measure of verbal reasoning (usually vocabulary) and, as a group, are broadly matched to a typically developing comparison group. However, vocabulary is a recognized peak of ability within ASD and may overestimate general language abilities (Mottron, 2004). It is also not unusual to find that the ASD group is more variable, both in verbal reasoning scores and on the experimental measure. It is therefore not clear how well group means reflect individual differences within the group. Therefore it is currently unknown to what extent variation in language ability could affect visual scanning patterns in higher-order cognitive tasks (though see Norbury et al., 2009).

In the current study, we investigated oculomotor control of participants with autism and language impairment (ALI), autism and language scores within the normal range (ALN), non-autistic participants with LI and typically developing peers (TD). We used standard prosaccade and antisaccade tasks that require the participant to disengage from a central fixation point and direct a saccadic movement towards a specified location. Differences in SRT to disengage from a target have been found between ASD and TD controls, largely for social stimuli (i.e. human faces; Chawarska, Volkmar & Klin, 2010). We chose non-social stimuli in order to explore oculomotor behaviour in order to minimize confounds in performance associated with social stimuli. The gap effect is known to be highly reproducible, making it an appropriate task to assess baseline oculomotor function. In addition to saccade tasks, the participants ability to locate a target object and subsequently maintain fixation on this target in the presence of competing distracters was assessed using a search distracter task across two different conditions. Conditions were identical except for the number of distracters (two or four) displayed on the screen. This task can provide insight into an individuals ability to orient gaze appropriately to a clearly defined target and the ability to maintain fixation in the presence of competing visual stimuli. This provides a measure of voluntary control and distractibility. To our knowledge, these oculomotor tasks have never been used with a developmental LI population. To the extent that ALI and LI represent overlapping phenotypes, similarities in oculomotor performance are predicted across these groups. If, on the other hand, deficits in oculomotor control are more pronounced in those with co-morbidity, children with ALI can be expected to have the most severe deficits on oculomotor tasks, particularly those requiring voluntary inhibition of responses. The performance of the ALN group was more challenging to predict; to the extent that language serves as a mechanism for facilitating executive control (Marcovitch & Zelazo, 2009), we anticipated that the ALN group would demonstrate fewer task difficulties relative to language impaired groups. On the other hand, if aberrant gaze behaviour is an early developmental marker of ASD (Elsabbagh et al., 2009), deficits in oculomotor control would be evident across the ASD spectrum, including those with ALN.

**Methods**

**Participants**

Ninety-eight children aged 8–14 years were recruited to the study from the South East of England; 25 children were excluded for obtaining standard scores on measures of non-verbal reasoning of < 60 ($n = 10$), noncompliance ($n = 5$), age < 7 years ($n = 4$) or additional diagnoses such as hearing impairment or chromosomal anomaly ($n = 6$). Informed, written consent was obtained from all parents, verbal assent was obtained from all children, and the protocol was approved by the Research Ethics Committee at Royal Holloway, University of London.
Children with ASD (ALI, n = 18 and ALN, n = 17) all held an existing diagnosis of ASD based on DSM-IV/ICD-10 criteria derived from a multi-disciplinary team assessment external to the research group. ASD was the primary diagnosis cited on the Statement of Special Educational Need (SEN), a legal document in the UK that specifies entitlement to special educational provision; all were receiving specialist support for ASD in mainstream schools or units serving children with ASD. In addition, all children obtained scores of 7 or greater on Module 3 of the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore & Risi, 1999). Children with ALI also obtained standard scores of less than 80 on the Total Language Composite of the Clinical Evaluation of Language Fundamentals (CELF-4UK: Semel, Wiig & Secord, 2003) and were receiving language-based interventions from a speech-language therapist. Children with LI (n = 16) all held an existing diagnosis of Language Impairment on the Statement of SEN and were receiving full-time special educational support for language impairment and intervention from a speech-language therapist. In addition, they obtained standard scores of less than 80 on the CELF-4UK and scores of 6 or below on the ADOS. None of the children were receiving medication at the time of testing. TD children (n = 22) were recruited from local schools in the community and did not have any reported special educational needs, or a history of ASD or language delay. Verbal (VIQ) and non-verbal (NVIQ) abilities were assessed using the Matrix Reasoning and Definitions sub-tests of the Wechsler Abbreviated Scales of Intelligence (Wechsler, 1999). Receptive vocabulary was measured using the Receptive One Word Picture Vocabulary Test (Gardner, 1990). All groups were matched for age; the ALN and TD groups were additionally matched on all cognitive and language measures. As is typical for older school-aged children with language impairment (see Botting, 2005; Kjelgaard & Tager-Flusberg, 2001), the ALI and LI children presented with low-average receptive vocabulary, verbal and non-verbal reasoning scores and significantly impaired structural language skills (as measured by the CELF), in which Total Language scores were more than 2 SDs below both chronological age expectations and non-verbal reasoning abilities. Both the ALI and LI groups had significantly poorer scores on language and cognitive measures relative to ALN and TD peers, but did not differ from one another (see Table 1).

Eye-tracking acquisition and analysis

Eye-movements were recorded binocularly at a sampling rate of 60 Hz using a Tobii T120 eye tracker, which has an average gaze position error of 0.5° and a spatial resolution of 0.2°. Attempts were made to sample eye-movements at 120 Hz, but data were lost as a consequence of extreme head movements from some participants. The initial calibration was conducted at the beginning of each experimental task using a 5-point fixation procedure in Tobii Studio software and repeated throughout the testing session as required. All experiments were implemented using E-Prime software (Psychology Software Tools Inc., PA), with a 640 x 480 screen resolution. Children were seated in a comfortable position directly in front of the computer monitor at a viewing distance of 60 cm. Instructions were provided verbally and also displayed on the screen subsequently.

Raw data were extracted and analysed using custom written Matlab (The Mathworks, MA) code. In the pro- and anti-saccade tasks, trials were considered valid if the participant was fixating the central fixation point at the moment of target onset and successfully moved their eyes horizontally towards (prosaccade) or away from (antisaccade) the stimulus. In the search distracter tasks, trials were considered valid when the participant was fixating the central fixation point at the moment of

<table>
<thead>
<tr>
<th>Table 1 Descriptive statistics for age, non-verbal ability, verbal ability, vocabulary, language and symptom scores</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>Measure</td>
</tr>
<tr>
<td>Chronological age (years)</td>
</tr>
<tr>
<td>WASI Matric Reasoning</td>
</tr>
<tr>
<td>ALN &amp; LI</td>
</tr>
<tr>
<td>Vocabulary</td>
</tr>
<tr>
<td>Receptive OWVPY</td>
</tr>
<tr>
<td>CELF-4UK</td>
</tr>
<tr>
<td>60–136</td>
</tr>
<tr>
<td>67–118</td>
</tr>
<tr>
<td>63.70 (10.06)</td>
</tr>
<tr>
<td>46–79</td>
</tr>
<tr>
<td>ADOS</td>
</tr>
<tr>
<td>7–20</td>
</tr>
</tbody>
</table>

Data are presented as Mean (SD) Range. Note: WASI T-scores transformed to standard scores with x 100 and SD 15 for ease of comparison.
target onset and successfully located the specified target by fixating it. In all tasks, eye-movements were considered to be anticipatory if they occurred < 100 ms from target onset. Fixations were defined as stable looking (±0.5°) for a minimum of 100 ms.

**Prosaccade task**

Seventy-two children successfully completed the prosaccade task; one child (LI) was excluded from the task due to missing data caused by excessive movement. In Gap trials, a central fixation point in the form of a schematic smiley face was displayed for 500–800 ms before disappearing from the screen. Following a 200 ms delay, a target appeared on the horizontal meridian 10° to the left or right of the centre and was displayed for 800 ms (see Figure 1a). The target was a cartoon monster or alien picture measuring 25 × 25 pixels. When displayed on the monitor with a 640 x 480 resolution, from 60 cm the size of the images measured 1.25° × 1.34° visual angle. In overlap trials, the central fixation point was displayed for 500–800 ms before the target appeared on the screen. The central fixation point and target were displayed simultaneously for 200 ms before the central fixation point disappeared from the screen leaving the target displayed for a further 800 ms (see Figure 1b). Participants were instructed to look at the central fixation point at the start of each trial and then move their eyes to the target as soon as it appeared on the screen, returning their gaze to the centre of the screen when the target disappeared. Each child completed 40 trials, which comprised 20 gap and 20 overlap trials. The number of left/right trials was counterbalanced and the order of trials was fully randomized so the target location could not be predicted. A practice session contained eight trials in which the Tobii Gaze Replay extension for E-Prime was used, to allow the experimenter to observe online whether the child was performing the task as instructed. Children progressed to the main task when 50% accuracy on practice trials was attained.

**Antisaccade task**

In total 64 children successfully participated in the standard antisaccade task; nine children were excluded from the task due to missing data caused by excessive movement or inattentiveness (ALI, n = 2; ALN, n = 4; TD, n = 3). The procedure for the antisaccade task was identical to the prosaccade task, but the instructions differed. The child was instructed not to look towards the stimulus (i.e. monster or alien), but instead to look to the opposite side of the screen to the approximate location where the target would be if displayed on that side. As is typical in anti-saccade tasks, there was no requirement for accuracy of fixation landing in the mirror location; instead we were interested in the number of trials on which they initiated a movement to the opposite side of the screen versus the number of times they made a prosaccade error (fixating the stimulus). Prior to experimental testing, a short practice session with feedback in the form of Gaze Replay was provided and children progressed to the experimental task when they achieved accuracy of at least 50%.

**Search distracter task**

In total 62 children participated in a two distracter condition and 59 children participated in a four distracter condition. One child (ALI) was excluded from the two distracter condition due to missing data caused by excessive movement and 10 TD children did not participate due to time restrictions. Four children were excluded from the four fixation distracter task due to missing data caused by excessive movements (ALI, n = 2; ALN, n = 1; TD, n = 1) and 10 TD children did not participate due to time restrictions.

A central fixation schematic smiley face was displayed in the centre of the screen for 800–1000 ms. The central fixation point was then removed and three dragons measuring 25 × 29 pixels (1.25° × 1.55° visual angle)
appeared on the screen for 1500 ms or 3000 ms in a circular configuration (see Figure 2). The participants were instructed to find the red dragon on the screen. Furthermore, participants were told explicitly that they should not look away from the target once it had been located. A single red dragon (the target) was present in every trial accompanied by two further dragons (distracters). The colour of the distracters varied across trials, but they were always the same colour as each other (e.g. blue, green). In total, the target could appear in one of eight different locations, which were analogous to compass points. The locations of the target and distracters were counterbalanced across trials. In the second task condition, five dragons were displayed on each trial (one target and four distracters). All timings and other details were identical to the two distracter condition. In both conditions, participants completed 32 trials (1500 ms, \( n = 16 \); 3000 ms, \( n = 16 \)). Practice trials and criterion to proceed to the main testing session were equivalent to the pro- and anti-saccade tasks.

Results

A full breakdown of participant characteristics for each task, error types and non-significant main effects and interaction terms is provided in the Supporting Information. For each task we analyse latency, or the time taken to fixate the target, and the accuracy of that fixation. Trials were considered accurate when the participants saccade landed within a region covering the target and surrounding area (2.5° × 2.68° visual angle for saccade tasks; 2.5° × 3.1° visual angle for the search distracter tasks).

Prosaccade task

Figure 3 illustrates the mean latency and accuracy rates for each group from both saccade tasks. Two 4 (Group) × 2 (Condition: Gap or Overlap) ANOVAs were conducted on saccade latency and target fixation accuracy. A main effect of condition was found with all groups exhibiting the gap effect by displaying significantly shorter saccade latencies in the gap condition, \( F(1, 67) = 221.36, p < .001, \eta_p^2 = .77 \). There were no main effects of group, nor were there any significant group × condition interactions (Fs < 1). Directional error rates (i.e. looks away from the stimulus) were low across all groups (Table S3), and there were no significant main effects or interactions, Fs < 1 (see Table S10 for detailed results). Thus, despite group differences in verbal and non-verbal IQ, basic reflexive oculomotor abilities were not disrupted in these populations.

Antisaccade task

Two 4 (Group) × 2 (Condition: Gap or Overlap) ANOVAs were conducted on saccade latency and directional errors (i.e. looks to the stimulus rather than the opposing horizontal location). Again, all groups exhibited the gap effect by displaying significantly shorter saccade latencies to target in the gap condition relative to the overlap condition, \( F(1, 60) = 80.160, p < .001, \eta_p^2 = .572 \) (Figure 2 and Table S5). Neither the main effect of group, nor the group × condition interaction was significant, \( F < 1.1 \). Directional errors were made by the ALI and LI groups on 20.35%, and 18.76% of trials. In contrast, such errors occurred less frequently in the ALN and TD groups (13.76% and 9.62% of trials, respectively). These group differences were significant, \( F(3, 60) = 3.320, p < .025, \eta_p^2 = .126 \). Planned
comparisons revealed that the ALI and LI groups made significantly more directional errors than TD peers (TD vs. ALI, \( p < .006 \); TD vs. LI, \( p < .022 \)). The ALN group did not differ significantly from TD peers (\( p = .283 \)). Neither the main effect of Condition nor the group \( \times \) condition interaction was significant (Table S10). In real terms, ALI and LI children made directional errors on average every one in five trials, whereas ALN and TD children only made errors every one in 10 trials.

Children with ALI/LI clearly found this task challenging and, in the absence of a visible goal for fixation, it is possible that these groups misunderstood task instructions or didn’t know what to do. To address this, we assessed corrective shifts in gaze following the initial directional error. A high percentage of corrective shifts were made by all groups: ALI = 72.52%, ALN = 85.72%, SLI = 80.15%, TD = 84.96%. A one-way ANOVA revealed significant between-groups differences, \( F(3, 60) = 4.646, p < .005, \eta^2 = .189 \), with TD children making more corrective shifts relative to the ALI group (\( p < .004 \)). Given the relatively low overall percentage of directional errors and the high proportion of corrective gaze shifts observed in these trials, it appears that the differences in error rates observed between groups do not reflect difficulty in following instructions.

**Search distracter task**

Initial analyses revealed no group differences in latency to fixate the target stimulus in either condition or in the number of fixations made prior to locating the target, \( F_S < 1 \). The ability to maintain fixation on the target was assessed using a 4 (Group) \( \times \) 2 (Time: Short versus Long) \( \times \) 2 (Condition: Two versus Four distracters) ANOVA. Here, there was a significant effect of Time on post-target fixations, \( F(1, 58) = 125.067, p < .001, \eta^2 = .695 \), with all groups displaying more post-target fixations at longer trial lengths. A main effect of Group, \( F(3, 58) = 4.226, p < .009, \eta^2 = .187 \), was also evident. Children in the ALI and LI groups made a significantly greater number of post-target fixations than TD peers (TD vs. ALI, \( p < .002 \); TD vs. LI, \( p < .009 \)), while the ALN group did not (\( p = .115 \)). None of the interaction terms involving Group were significant, indicating that the LI/ALI groups were distracted even when the task demands were minimal (Tables S8–S10 for full details).

Inspection of the error data across tasks revealed greater variability within the clinical populations. To investigate individual differences in volitional errors, we normalized directional error rates from the antisaccade task and number of post-target fixations from the search distracter task for each participant using TD means and standard deviations. A threshold of \(-1.5\) SD was selected as a cut-off to determine which individuals possessed a deficit in volitional eye-movement control as this was the extreme lower boundary for TD performance (only one TD child fell below this boundary on either volitional task). For the antisaccade task, 93.75% of ALI children and 81.25% of LI children scored below the threshold compared with just 23.07% ALN children and 5.26% TD children, \( \chi^2(3) = 37.480, p < .001 \). For the search distracter task, 35.29% of ALI children and 43.75% of LI children scored below the threshold compared with 17.64% and 0% for ALN and TD, respectively, \( \chi^2(3) = 8.255, p < .04 \). This is depicted in Figures 4a and 4b, where each individual participant is represented as a single data point. Comparison of children in the ALI and LI groups who scored below threshold in at least one task with those that did not revealed no significant difference in non-verbal IQ, \( t(30) = .62; p = .95 \). However, those scoring below threshold did have significantly lower verbal reasoning \( t \)-scores than peers (\( M = 38.8 \) versus \( M = 46.2; t(30) = 2.495; p = .028 \); Cohens \( d = .77 \)). There was no

---

**Figure 4** (a) Individual participant z-scored directional error rates in antisaccade task (the –1.5 SDs threshold is marked by the red horizontal line). Each participant is represented by a single data point; (b) Individual participant z-scored post-target fixation error rates in fixation distracter task (the –1.5 SDs threshold is marked by the red horizontal line). Each participant is represented by a single data point.
difference between these two groups on CELF-IV scores ($r < 1$), perhaps because these two groups were selected to have low scores on this measure.

**Discussion**

In this study, children with neurodevelopmental disorders demonstrated normal control and speed of reflexive eye-movements, suggesting that at the most basic level, the underlying oculomotor system is intact (see Rommelse, van der Stigchel & Sergeant, 2008). Eye-movement deficits were only apparent in tasks that involved volitional control but were not limited to individuals with ASD diagnoses and were not explained by differences in non-verbal cognitive ability. Instead, deficits aligned with language status. Specifically, children with autism spectrum disorders and additional language impairment, and non-autistic children with language impairments, had greater difficulty suppressing reflexive shifts of gaze and maintaining fixation on a target in the presence of competing distracters.

Previous research using oculomotor tasks in ASD populations has yielded inconsistent findings, with some investigators reporting increased overall saccade latencies but no differences in the magnitude of the gap effect (Goldberg et al., 2002), while others have found the opposite pattern (Van Der Geest et al., 2001), and still others have suggested impairments in both (Elsabbagh et al., 2009). Takarae et al. (2004) reported that oculomotor control deficits in ASD, identified using a visually guided saccade task, were associated with language delay. The extension of oculomotor deficits to non-autistic children with LI provides further support for the notion that ALI represents a distinct neurocognitive phenotype that shares overlapping genetic and neurobiological risk factors with LI (Tager-Flusberg & Joseph, 2003). Hodge, Makris, Kennedy, Caviness, Howard, McGrath, Steele, Frazier, Tager-Flusberg and Harris (2009) demonstrated that individuals with ALI and LI phenotypes have similar neurodevelopmental anomalies in fronto-cortico-cerebellar circuits, which underpin language, motor control and attention, and are contiguous with neural circuits implicated in volitional oculomotor control (Munoz & Everling, 2004; Pierrot-Deseilligny, Rivaud, Gaymard & Agid, 1991). Antisaccade errors have been linked to working-memory processes that may be related to the ability to maintain an instruction and apply it at the appropriate time (Roberts, Hager & Heron, 1994; Walker, Husain, Hodgson & Kennard, 1998). The inability to adequately suppress a voluntary response, particularly in the presence of numerous competing stimuli, could negatively impact the developmental trajectories of skill acquisition across numerous domains, including language and social understanding (see Karmiloff-Smith, 1998; Norbury et al., 2009).

In addition to reduced language ability, both the ALI and LI exhibited lower non-verbal IQ relative to TD and ALN groups. To isolate the affects of language on oculomotor performance, it may have been preferable to exclude those children with language impairments who had non-verbal IQ scores more than $-1\,\text{SD}$ below the mean. We elected not to do this for a number of reasons. First, non-verbal IQ has not generally been used as an exclusion criterion when identifying the LI phenotype in ASD (see Tager-Flusberg & Joseph, 2003) and is not part of the proposed diagnostic criteria for Language Impairment in the revised DSM-5 (www.dsm5.org). Second, longitudinal studies of non-autistic children with LI have consistently demonstrated a decrease in standardized non-verbal reasoning scores over time (Botting, 2005), even when the same assessments have been used (see Bishop & Adams, 1992, and Stothard, Snowling, Bishop, Chipchase & Kaplan, 1998). Thus, those that may meet strict discrepancy criteria for specific language impairment in the early school years may no longer do so in later childhood. To exclude those in the ALI and LI groups with low non-verbal IQ would have resulted in unacceptably small groups that were non-representative of the wider population. Finally, although the ALI and LI groups had lower non-verbal reasoning scores relative to ALN and TD peers, they did not differ from each other. To the extent that we are interested in the specificity of oculomotor deficits as characterizing ASD performance, this is an appropriate comparison.

Where group differences in NVIQ exist, it is often suggested that NVIQ be controlled in statistical analysis. Dennis, Francis, Cirino, Schachar, Barnes and Fletcher (2009) explain why this is both theoretically and statistically inappropriate to do. Essentially, the differences in NVIQ seen here are not the result of poor sampling, but rather reflect non-random, pre-existing differences that are associated with diagnosis. To control for NVIQ would in effect control for the variable we are most interested in, impaired language development. For that reason, we did not use analysis of covariance in our statistical analyses.

It is therefore possible that the differences we observe have little to do with language and are attributable to general cognitive delays that affect task performance. If so, we might anticipate that the ALI and LI populations would have also been impaired on the prosaccade task, but this was not the case. In fact, all children included in our analyses were capable of completing all the tasks they were set and all produced valid trials. In addition, the ALI and LI groups made corrective errors on the antisaccade task, demonstrating an understanding of task instructions. Instead, high rates of error were specific to volitional tasks and were more likely in those with the lowest verbal reasoning abilities.

Furthermore, to our knowledge, there is no evidence to suggest that non-verbal reasoning is associated with volitional eye-movement control, and no mechanistic explanation for why such an association might be expected. In contrast, previous studies (Norbury et al., 2009; Takarae et al., 2004) have shown an association
between eye-movements and language status. An important consideration is whether eye-movement patterns and language development are causally related. In ASD, it is clear that at least for a proportion of individuals, anomalies in eye-movements are apparent before the onset of spoken language as they are evident in infant siblings of autistic children (Elsabbagh et al., 2009) and infants and toddlers with ASD (Chawarska et al., 2010; Pierce et al., 2011). In addition, deficits in oculomotor control are also characteristic of unaffected, first-degree relatives of autistic individuals (Mosconi, Kay, DCruz, Guter, Kapur, Macmillan, Stanford & Sweeney, 2010) and implicate deficits in left frontotemporal cortical circuits that overlay neural pathways crucial for language development. Taken together, this body of research indicates that aberrant eye-movements may serve as a precursor to developmental delays in joint attention and imitation, leading to lifelong disruptions of language acquisition and social processing (Brenner et al., 2008).

Longitudinal studies will elucidate whether these early anomalies in oculomotor control are indeed predictive of the ALI phenotype. However, in this study, inefficiencies in oculomotor behaviour were also seen in non-autistic children with LI. At the present time, there is a paucity of research investigating the earliest behavioural markers of LI and therefore we do not know whether the patterns of eye-movement control we see here would be evident in infants and toddlers at risk for LI. The lack of social deficit in LI suggests that for this group, deficits in volitional eye-movement control may be a consequence of impaired language function rather than a cause of it. Again, longitudinal studies, comparing children with different neurodevelopmental disorders, will be needed to establish causal relationships.

The pattern of deficit observed in the ALI and LI groups is also consistent with findings from other neurodevelopmental disorders such as Fragile X (Scerif et al., 2005) and particularly ADHD (Munoz et al., 2003), and may therefore reflect further co-morbidities. At least one-third of children with ASD have co-morbid ADHD (Simonoff et al., 2008), and a similar proportion of non-autistic children with LI have clinically significant deficits on verbal and non-verbal measures of executive control (Henry, Messer & Nash, 2011). Our findings may therefore point to an additional co-morbidity that we did not explicitly measure, though none of our participants were currently being medicated for ADHD. There is no indication at present that co-morbid ADHD is more common in individuals with ASD who have concomitant language impairment; however, children with ALN may be less affected because language is an important mediator of executive control (Marcovitch & Zelazo, 2009). Language may help children to reflect on the goals of the task at hand, and to internalize arbitrary rules (e.g. look to the other side of the screen or keep looking at the red dragon), contributing to task success. It is notable that even within the language impairment groups, those with the most severely impaired verbal reasoning abilities were the most likely to have extremely high error rates on volitional eye-movement tasks.

Population heterogeneity is rarely taken into account in autism research though there is increasing evidence that children with different neurocognitive phenotypes involving language may show different visual scanning patterns (Norbury et al., 2009) and that similar scan patterns may reflect different underlying processes in those with discrepant verbal–non-verbal abilities (Rice et al., 2012). Our findings suggest that measurement of oculomotor control could further enhance interpretation of eye-tracking studies tapping higher-order cognitive processes. For instance, reduced fixation time to eyes and faces is often taken as evidence of reduced interest in social stimuli. Our findings suggest that individuals with lower language levels are more likely to be distracted by competing visual stimuli, potentially affecting their visual sampling of complex images or dynamic scenes. Rice et al. (2012) report that those with discrepant verbal and non-verbal abilities demonstrated more off-screen looks, consistent with this suggestion. Variations in language skill may also be important in understanding top-down control of scanning, for instance, in modulating fixation patterns according to task instructions (e.g. Benson et al., 2009). Finally, similarities between the ALI and LI groups in this study further suggest that cross-disorder comparisons are essential for identifying the specificity and developmental consequences of aberrant oculomotor behaviour.

In summary, there is a current explosion of eye-movement research in ASD exploring higher-level social and cognitive processes. The present findings suggest that a proportion of individuals with ASD, particularly those with concomitant language impairment, have deficits in volitional oculomotor control that may render such research difficult to interpret. Our findings also demonstrate for the first time phenotypic overlap between ALI and LI populations on ostensibly non-verbal tasks, though the role of verbal mediation in non-verbal executive control warrants further investigation. It is clear that the pattern of findings observed in this study is not limited to language impairment, but rather that volitional control of eye-movements may serve as a marker of neurodevelopmental anomaly, in which language acquisition is especially vulnerable.

Acknowledgements

We are grateful to the children and their families who participated in this study, as well as the schools who hosted our research. We also thank Rebecca Lucas and Anna Powell for their assistance with data collection. This work was supported by an Economic and Social Research Council grant (RES-061-25-0409) awarded to CFN.
References


Supporting Information

Additional supporting information may be found in the online version of this article:

Table S1. Participant characteristics: Prosaccade task
Table S2. Prosaccade task: Mean (SD), saccade latency range, percentage of errors, post-hoc t-test results for gap-overlap latencies.
Table S3. Prosaccade Task: Breakdown of errors
Table S4. Participant characteristics: Antisaccade task
Table S5. Antisaccade task: Mean (SD), saccade latency range, percentage of errors, post-hoc t-test results for gap-overlap latencies.
Table S6. Antisaccade Task: Breakdown of errors
Table S7. Participant characteristics: Fixation Distractor Tasks
Table S8. Additional variable analysis: 2 distractor condition. SDs in parenthesis.
Table S9. Additional variable analysis: 4 distractor condition. SDs in parenthesis.
Table S10. Summary of non-significant ANOVAs

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
Eye-tracking paradigms are increasingly used to investigate higher-level social and cognitive processing in autism spectrum disorder (ASD). However, the integrity of the oculomotor system within ASD is unclear, with contradictory reports of aberrant eye-movements on basic oculomotor tasks. The purpose of the current study was to determine whether reducing population heterogeneity and distinguishing neurocognitive phenotypes can clarify discrepancies in oculomotor behaviour evident in previous reports.
Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper’s edge. Please remember that illegible mark-ups may delay publication. Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query reference</th>
<th>Query</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AUTHOR: please check the graphical abstract.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>AUTHOR: Figure 2 has been saved at a low resolution of 288 dpi. Please resupply at 300 dpi. Check required artwork specifications at <a href="http://authorservices.wiley.com/submit_illust.asp?site=1">http://authorservices.wiley.com/submit_illust.asp?site=1</a></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>AUTHOR: Figure 3 is of poor quality. Please check required artwork specifications at <a href="http://authorservices.wiley.com/submit_illust.asp?site=1">http://authorservices.wiley.com/submit_illust.asp?site=1</a></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>AUTHOR: Figure 4 is of poor quality. Please check required artwork specifications at <a href="http://authorservices.wiley.com/submit_illust.asp?site=1">http://authorservices.wiley.com/submit_illust.asp?site=1</a></td>
<td></td>
</tr>
</tbody>
</table>
Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

<table>
<thead>
<tr>
<th><strong>Instruction to printer</strong></th>
<th><strong>Textual mark</strong></th>
<th><strong>Marginal mark</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave unchanged</td>
<td>· · · under matter to remain</td>
<td>New matter followed by  [ or [ \circlearrowleft  ]  [ or [ \circlearrowright  ]</td>
</tr>
<tr>
<td>Insert in text the matter indicated in the margin</td>
<td>/ through single character, rule or underline or [ [ through all characters to be deleted [ [</td>
<td>new character / or new characters /</td>
</tr>
<tr>
<td>Delete</td>
<td>/ through letter or [ [ through characters</td>
<td>[ [</td>
</tr>
<tr>
<td>Substitute character or substitute part of one or more word(s)</td>
<td>/ through letter [ [ through characters</td>
<td>[ [</td>
</tr>
<tr>
<td>Change to italics</td>
<td>under matter to be changed</td>
<td>[ [</td>
</tr>
<tr>
<td>Change to capitals</td>
<td>under matter to be changed</td>
<td>[ [</td>
</tr>
<tr>
<td>Change to small capitals</td>
<td>under matter to be changed</td>
<td>[ [</td>
</tr>
<tr>
<td>Change to bold type</td>
<td>under matter to be changed</td>
<td>[ [</td>
</tr>
<tr>
<td>Change to bold italic</td>
<td>Encircle matter to be changed</td>
<td>[ [</td>
</tr>
<tr>
<td>Change italic to upright type</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Change bold to non-bold type</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert ‘superior’ character</td>
<td>/ through character or [ [ where required</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert ‘inferior’ character</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert full stop</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert comma</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert single quotation marks</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert double quotation marks</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert hyphen</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>Start new paragraph</td>
<td>(As above)</td>
<td>[ [</td>
</tr>
<tr>
<td>No new paragraph</td>
<td>[ [</td>
<td>[ [</td>
</tr>
<tr>
<td>Transpose</td>
<td>[ [ linking [ [ characters</td>
<td>[ [</td>
</tr>
<tr>
<td>Close up</td>
<td>/ through character or [ [ where required</td>
<td>[ [</td>
</tr>
<tr>
<td>Insert or substitute space</td>
<td>/ through character or [ [ where required</td>
<td>[ [</td>
</tr>
<tr>
<td>between characters or words</td>
<td>between characters or words affected</td>
<td>[ [</td>
</tr>
<tr>
<td>Reduce space between characters or words</td>
<td>between characters or words affected</td>
<td>[ [</td>
</tr>
</tbody>
</table>

(As above)