This is the peer reviewed version of the following article: [Gonzalez-Barrero, A.M., and Nadig, A.S. (2017). Can Bilingualism Mitigate Set-Shifting Difficulties in Children With Autism Spectrum Disorders? Child Development 90, 1043–1060.]

Can Bilingualism Mitigate Set-Shifting Difficulties in Children with Autism Spectrum Disorders?

Ana Maria Gonzalez-Barrero[1,2] and Aparna Nadig[1,2]

[1]*School of Communication Sciences and Disorders, McGill University, Montreal, Canada*

[2]*Centre for Research on Brain, Language and Music (CRBLM), Montreal, Canada*

Corresponding Author Contact Information
Ana Maria Gonzalez-Barrero
School of Communication Sciences and Disorders
McGill University
Montreal, Canada
ana.gonzalez@mail.mcgill.ca

Funding
This project was funded by doctoral fellowships from the Fonds de Recherche du Québec - Société et Culture and the Social Sciences and Humanities Research Council of Canada to Gonzalez-Barrero.

Abstract

This study investigated the effects of bilingualism on set-shifting and working memory in children with Autism Spectrum Disorders (ASD). Bilinguals with ASD were predicted to display a specific bilingual advantage in set-shifting, but not working memory, relative to monolinguals with ASD. Forty 6- to 9-year-old children participated (20 ASD, 20 typically-developing). Set-shifting was measured using a computerized Dimensional Change Card Sort task, and by parent-report of executive functioning in daily life. Results showed an advantage for bilingual relative to monolingual children with ASD on the DCCS task but not for set-shifting in daily life. Working memory was similar for bilinguals and monolinguals with ASD. These findings suggest that bilingualism may mitigate some set-shifting difficulties in children with ASD.

The impact of bilingualism on cognition is an ongoing topic of debate. Of particular interest is whether the use of two languages confers an advantage to executive functions (EF). A number of researchers have reported enhanced performance on EF tasks for typically-developing bilingual children (e.g., Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008), whereas others have failed to find significant differences between monolingual and bilingual groups (e.g., Morton & Harper, 2007; Namazi & Thordardottir, 2010). EF abilities have also been investigated in monolingual children with Autism Spectrum Disorders (ASD), generally showing poorer performance relative to typically-developing peers. However, it is not known if bilingualism may hold consequences for the EF of children with ASD. In the present study, we test for the first time whether the bilingual advantage hypothesis extends to children with ASD. Specifically, we compare four groups of children: bilingual and monolingual children, with ASD or with typical development. We examine whether cognitive flexibility, or set-shifting skills, may be enhanced in bilinguals relative to monolinguals, and include working memory skills as a control EF not expected to be affected by bilingualism.

Executive Functions in Autism Spectrum Disorders

Autism Spectrum Disorders (ASD) are neurodevelopmental disorders characterized by impairments in social communication and social interaction, accompanied by restricted and repetitive behaviors and interests (American Psychiatric Association, 2013). Executive functioning has been a major area of interest in autism research (e.g., Robinson, Goddard, Dritschel, Wisley, & Howlin, 2009), with *the theory of executive dysfunction in autism* being proposed as an explanatory account for the condition (Ozonoff, Pennington, & Rogers, 1991, see Hill, 2004 for a critical review of this account). Executive functions comprise an array of cognitive abilities necessary for goal oriented problem-solving in everyday life, including planning, inhibition, working memory, and set-shifting (Eigsti, 2011; Gioia, Isquith, Guy, & Kenworthy, 2002). These abilities have neural bases in the frontal cortex and frontalstriatal and frontalparietal pathways (Schmitz et al., 2006). The executive function of set-shifting (also referred to as cognitive flexibility) is particularly relevant to ASD. Set-shifting is the EF domain exhibiting the most pronounced impairment in daily life for children with ASD relative to both children with typical development (Granader et al., 2014, in a sample with over 400 participants in each group) and those with other developmental disorders (Gioia et al., 2002). Though not included in the diagnostic criteria for ASD, set-shifting difficulties are closely related to the second defining domain of ASD symptomology: repetitive or restricted behavior and interests (APA, 2013). Symptoms falling in this domain that are commonly (but not universally) experienced by people with ASD include difficulty with transitions, narrowness of focus, perseveration in interests and activities, desire for sameness in the environment, and inflexible adherence to routines (Leekam, Prior, & Uljarevic, 2011).

Set-shifting has been defined as the ability to accurately switch back and forth between tasks given specific demands (Miyake et al., 2000). A quantitative review of set-shifting in individuals with ASD by Leung and Zakzanis (2014) included 72 studies, most of which tested adults but also including the few studies that have tested children and adolescents, yielding a combined sample of 2,137 individuals with ASD and 2,185 controls. This review covered 19 lab-based experimental (also referred to as performance) tasks for which multiple specific measures of set-shifting were evaluated. Across these studies, neurotypical participants outperformed participants with ASD with mean effect sizes from 0.30 – 1.88, as measured by Cohen's d (with benchmarks of .2 for small, .5 for medium, and .8 for large effects, Cohen, 1992). Interestingly, the one informant-based measure of set-shifting in daily life that was included in the review

(BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000a), bore much larger effects, with a mean effect size of 2.14 for children based on parent-report and 5.04 from one study using self-report with adults. This review highlights three key points: (a) across a broad range of samples, tasks, and measurement constructs (but not in each study included) individuals with ASD exhibited an impairment in set-shifting, (b) this effect was highly variable in magnitude, which is not surprising giving the scope of the review, but as such did not meet the authors' stringent criteria for use of set-shifting as a clinical marker for ASD (d >= 3.0, and < 5% of overlap in data points between distributions), (c) effect size is consistently and substantially larger for set-shifting difficulties in daily life relative to that observed on experimental tasks. This divergence can be attributed to the fact that experimental tasks isolate set-shifting skills from real-life demands, imposing a standard structure and simplifying the problem space relative to set-shifting situations encountered in daily life.

A number of factors have been highlighted in the literature to explain variable findings on executive function tasks in individuals with ASD. One just mentioned is that of task selection (e.g., report of EF in daily life vs. on performance tasks, and across performance tasks with more or less explicit rules or structure, Leung & Zakzanis, 2014; Van Eylen et al., 2011). A second is with respect to testing modality (e.g., live experimenter administration vs. computer administration). Specifically, multiple studies have reported that individuals with ASD perform better on a given task when presented by computer rather than live administration by an experimenter, whereas typically-developing comparison groups perform similarly across modalities (reviewed by Kenworthy, Yerys, Anthony, & Wallace, 2008). This has been reported for set-shifting (Ozonoff, 1995; Pascualvaca, Fantie, Papageorgiou, & Mirsky, 1998), as well as planning tasks (Kenworthy et al., 2008), and has been attributed to (a) a decrease in the social demands of completing the task and/or (b) the motivation typically-developing children derive from live interaction that their counterparts with ASD may not display. A third is the massive heterogeneity amongst participants with ASD in terms of nonverbal IQ (NVIQ) and language abilities, necessitating proper characterization of samples and the use of appropriate matching procedures with comparison groups (Russo et al., 2007).

A commonly employed experimental task of set-shifting in children is the Dimensional Change Card Sort (DCCS) task (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006). In a first “pre-switch” phase children are asked to sort a series of simple images depicted on cards (e.g., boats and rabbits) according to one dimension (e.g., color). Afterwards, in a “post-switch” phase, they are asked to sort the same images according to a different dimension (e.g., shape). Therefore, the task requires that children disengage from the previously used rule and switch to a new dimension that is explicitly stated (Prior & Macwhinney, 2010). Typically-developing 5-year-olds are able to perform the switch correctly, while younger children tend to continue sorting by the first dimension (Zelazo, 2006). Children who perform correctly during the post-switch phase are then administered a more advanced phase in which mixed rather than blocked trials are presented (i.e., border version; Zelazo, 2006). In this advanced phase, the task is to sort images by one of the previously used dimensions depending on a visual cue (e.g., sort by color if the image has a border but by shape if the image has no border). As reported by Zelazo (2006), different measures (e.g., reaction time) from the border version have been used to assess set-shifting skills in both school-age children and adults (e.g., Bialystok & Martin, 2004; Diamond & Kirkham, 2005).

The DCCS, in different testing modalities, has been employed to examine set-shifting abilities in monolingual children with ASD (e.g., Dichter et al., 2010; Faja & Dawson, 2014; Yi et al., 2012). Faja and Dawson (2014) studied the performance of 20 typically-developing children and 23 children with ASD matched on age (6 to 8 years), gender, and IQ using the standard card presentation of the DCCS task with a live examiner. Results showed that accuracy on the post-switch phase for the children with ASD was comparable to that of typically-developing children. However, the typically-developing children outperformed their peers with ASD on the border version. Yi and colleagues (2012) reported that 18 3- to 9-year-olds with ASD, matched on verbal mental age to a group of 31 typically-developing children (3 to 6 years old) exhibited deficits on both the post-switch phase and border versions of a computerized DCCS task. Group differences on the post-switch phase were likely found because it was still developmentally sensitive to this younger sample. Dichter and colleagues (2010) employed a computerized card sort task (with a modified administration procedure, and using written words to explicitly provide the rule for each block) with 6- to 17-year-olds with and without ASD who were matched on chronological age but who differed on NVIQ. These authors reported that participants with ASD performed similarly to the typically-developing comparison group on key measures of switching, but their results are difficult to interpret given lack of matching on NVIQ, a very broad age range, and differences in task design.

While impaired performance on set-shifting tasks has been widely documented, other domains of EF, such as verbal working memory, have been shown to be spared in children with ASD (Williams, Goldstein, & Minshew, 2006; however see Schuh & Eigsti, 2012 for an alternative view). Working memory is defined as the ability to temporarily maintain and manipulate information to accomplish a specific task (Baddeley, 1992). Experimental paradigms have been used to investigate working memory using visual as well as auditory stimuli (e.g., Ozonoff & Strayer, 2001). The available evidence suggests that individuals with ASD show difficulties with complex working memory tasks that encompass high processing load (Minshew & Goldstein, 2001), especially in the visual domain (e.g., Williams, Goldstein, Carpenter, & Minshew, 2005). In contrast, performance on simple verbal working memory tasks, such as word recall (e.g., Russell, Jarrold, & Henry, 1996) and digit recall (Bennetto, Pennington, & Rogers, 1996; Faja & Dawson, 2014; Williams et al., 2005; Williams et al., 2006), has been shown to be unimpaired.

Overall, children with ASD exhibit an uneven profile in their EF abilities, with consistent impairments on set-shifting skills alongside relatively preserved verbal working memory for simple tasks (as demonstrated by Faja & Dawson, 2014 in 6- to 8-year-olds). So far, studies examining EF in ASD have only included monolingual children. We used Faja and Dawson's (2014) methods as a model to examine whether bilingualism may confer a specific advantage in characteristically impaired set-shifting skills, but not verbal working memory, in children with ASD.

The Bilingual Advantage Hypothesis

Some researchers have reported that typically-developing bilinguals exhibit advantages in particular executive functions relative to monolinguals (e.g., Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009). The bilingual advantage hypothesis posits that living as a bilingual and using two linguistic systems involves increased cognitive demands. In particular, set-shifting skills are involved when (active and proficient) bilinguals exert control over their two linguistic systems to continuously switch between two competing languages to successfully meet the demands of their communication context (e.g., Bialystok & Viswanathan, 2009). The ongoing practice of switching between languages is thought to yield global advantages in set-shifting that can be observed in non-linguistic domains (Bialystok et al., 2009; Green, 1998;

Prior & MacWhinney, 2010). In contrast, it has been suggested that a bilingual advantage is not observed in other EF domains such as working memory (e.g., Bialystok, 2009; Engel de Abreu, 2011; however see Delcenserie & Genesee's, 2016 findings with adults). A potential cause for the bilingual advantage effect being more easily observed on set-shifting tasks is that these paradigms involve greater cognitive demands (e.g., switching between sets, updating, and inhibiting a previously established rule) than simple working memory tasks (although the latter also requires cognitive control to some degree, Engel the Abreu, 2011).

Of greatest relevance to the current study, a bilingual advantage in set-shifting has been reported for typically-developing children in studies using the DCCS task previously discussed or adaptations thereof (Barac & Bialystok, 2012; Bialystok, 1999; Carlson & Meltzoff, 2008). Barac and Bialystok (2012) administered a computerized color-shape switch task to three groups of bilingual 5- to 7-year-olds (Chinese-English, French-English, and Spanish-English bilinguals) and to a group of English monolingual children. Findings revealed that whereas monolingual and bilingual groups had similar reaction times (RT) for blocks in which switching was not required, for mixed blocks (e.g., border version of the DCCS task) the three bilingual groups outperformed the monolingual group by exhibiting shorter RTs.

It should be noted that the bilingual advantage hypothesis has been challenged by other researchers who have found comparable performance on EF tasks for bilinguals and monolinguals (e.g., Paap & Greenberg, 2013; Morton & Harper, 2007; Namazi & Thordardottir, 2010). These authors ascribed the enhanced performance reported elsewhere for bilingual participants to higher socio-economic status (Morton & Harper, 2007), better memory skills (Namazi & Thordardottir, 2010), and shortcomings of the statistical analyses conducted such as the presence of Type I errors (Paap & Greenberg, 2013), among others. Consequently, whether bilingualism confers and advantage to EF abilities remains a debated topic which warrants further research.

Though our focus in the present study is on executive functions, it is important to address language ability in relation to ASD, as well as in relation to bilingualism. First, when all children with ASD are grouped, they are significantly delayed in language development (e.g., Luyster et al., 2008; Weismer, Lord, & Esler, 2010). Importantly, this group-level delay co-exists with massive individual variability in language skills among children with ASD. One study found that by school age approximately half of children with ASD were language impaired as assessed by standardized tests, a quarter had borderline skills within two standard deviations of the normal range, and the last quarter scored in the normal range or above (Kjelgaard & Tager-Flusberg, 2001). Therefore, we carefully characterize the language abilities of our sample and include children with ASD who have a range of language abilities (see Participants section of Methods). Research questions focusing on the language skills of the current sample are addressed in a separate manuscript (Gonzalez-Barrero & Nadig, submitted). Second, in children who are bilingually exposed language skills are also heterogeneous. Bilinguals' language skills vary along a continuum related to the amount of language exposure received in their second language, with optimal mastery of two languages occurring with balanced exposure, that is, between 40% to 60% exposure to each language (Elin Thordardottir, 2011). Given that the starting point for the bilingual advantage hypothesis is living as a proficient bilingual, we ensured this in our sample by employing a very stringent definition of bilingualism as described in Methods under "Bilingual Status".

A Bilingual Advantage in Autism Spectrum Disorder?

As reviewed above, set-shifting is a domain of executive functioning that has been consistently reported to be impaired in individuals with ASD (e.g., Gioia et al., 2002; Granader et al., 2014; Hill, 2004; Leung & Zackzanis, 2014). Does bilingualism hold benefits for specific executive functions in children with ASD, as has been reported for typically-developing children (e.g., Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008; Poulin-Dubois et al., 2011)? That is, can being bilingual mitigate the set-shifting impairment observed in children with ASD? This was our novel research question, for which we posed the following hypotheses: (a) all children with ASD will exhibit impaired performance relative to typically-developing children on measures of set-shifting, (b) however, bilingual children with ASD will perform better on set-shifting tasks relative to monolingual children with ASD, who do not need to manage two linguistic systems; (c) bilinguals and monolinguals with ASD will not differ on tasks that assess verbal working memory, included as a control EF not expected to be affected by bilingualism. We evaluated these hypotheses using rigorous group matching, a stringent, multi-source definition of bilingualism, and comprehensive evaluation of a potential bilingual advantage. Specifically, we measured set-shifting with an experimental task alongside parent report of children's executive functions in daily life to provide a more ecologically valid view of EF skills, as has been called for in the literature (Kenworthy et al., 2008).

Method

Participants

Forty children with a chronological age range of 6 to 9 years participated in the present study (chronological age *M* = 8.0 years, *SD* = 9 months). Data was collected between February 2014 and April 2016. There were 20 typically-developing children (henceforth TYP, 10 monolinguals, 10 bilinguals) and 20 children with ASD (10 monolinguals, 10 bilinguals).

Exclusion criteria for the TYP group were parental report of history of language, learning or developmental difficulties; physical, visual or hearing limitations; or any family members who had been diagnosed with ASD. All children in the TYP group attended regular schools.

Participants with ASD, who did not have intellectual disability, were recruited from autism organizations, schools, therapy programs, and a database from previous studies. They all had a formal clinical diagnosis obtained from licensed clinicians or multidisciplinary groups from health care institutions and parents were asked to provide a copy of the diagnostic report to confirm the participant's status. In addition, the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003) was used to confirm ASD symptoms. This questionnaire consists of 40 questions which provide information about the child's social functioning and early communication abilities. A score of 15 or higher on the SCQ is consistent with a diagnosis of ASD. A one-way analysis of variance (ANOVA) revealed a statistically significant difference on SCQ scores between groups, $F(3,35) = 45.50$, $p < .001$. Post-hoc tests showed that as expected, participants with ASD exhibited higher scores than TYP children ($p < .001$). However, there was no statistically significant difference between the monolingual and bilingual TYP groups ($p = .11$) or between the monolingual and bilingual ASD groups ($p = .98$).

The study was conducted in Montreal, Canada, a multicultural city where French is the official language though many people also speak English. A quarter of participants had at least one immigrant parent (TYP: 0 monolinguals, 3 bilinguals; ASD: 3 monolinguals, 4 bilinguals). Children were speakers of English, French, or Spanish or a combination of any of these languages. Bilingual speakers of other languages were not included in order to control for language typology and because of the availability of the same standardized measures in these languages. Some participants in the bilingual groups (2 in the TYP bilingual and 3 in the ASD bilingual group) had a minimal exposure (i.e., less than 10%) to a third language over their lifetime but were not proficient in this language. The participants' language of instruction was either French or English. However, given that French is the official language in the province of Quebec, it is common that children attending English schools receive some amount of French language instruction.

Details of participant characteristics and demographic information is presented in Table 1. To ensure matching on key variables, a series of one-way ANOVAs were conducted confirming no significant difference between groups on the following five characteristics: NVIQ (NVIQ > 80), chronological age (6 to 9 years), maternal education as a proxy for socio-economic status (mothers from all groups had attained at least a high-school degree), and gender (the majority of children were boys, given the higher prevalence of ASD in males (Baio, 2012)). Nonverbal IQ was assessed using the Leiter-R (Roid & Miller, 1997), which is a completely nonverbal test appropriate for the assessment of children from different language backgrounds as well as for children with ASD.

Insert Table 1 about here

Two standardized tests were used to assess participants' language skills. Receptive vocabulary was measured with the Peabody Picture Vocabulary Test in English (PPVT-4; Dunn & Dunn, 2007), French (Échelle de vocabulaire en images Peabody-EVIP-R; Dunn, Theriault-Whelan, & Dunn, 1993) and Spanish (Test de Vocabulario en Imagenes Peabody-TVIP; Dunn, Padilla, Lugo, & Dunn, 1986). Vocabulary measures were included given the influence of vocabulary on EF (e.g., Bialystok, Craik, & Luk, 2008; Buac, Gross, & Kaushanskaya, 2016). The recalling sentences subtest from the Clinical Evaluation of Language Fundamentals (CELF-

4; Semel, Wiig, & Secord, 2003), along with its French (Évaluation clinique des notions langagières fondamentales; Wiig, Secord, Semel, Boulianne, & Labelle, 2009), and Spanish versions (CELF 4 Spanish Edition; Semel, Wiig, & Secord, 2006) were used to determine if participants had language impairment. Sentence repetition tasks such as this one are considered a useful clinical marker for language impairment and have been used in studies examining structural language disorder in ASD (Riches, Loucas, Baird, Charman, & Simonoff, 2010). Four of 10 participants in the monolingual ASD group and 3 of 10 participants in the bilingual ASD group met criteria for language impairment as indicated by scores 1 SD below the mean on the recalling sentences subtest. This cut-off provides adequate sensitivity and specificity for the identification of language impairment (Conti-Ramsden et al., 2001; Elin Thordardottir et al., 2011). Participants with ASD and concomitant language impairment were included to reflect the heterogeneity of language abilities in this population (Kjelgaard & Tager-Flusberg, 2001). All children in the TYP groups had scores within the average range on the recalling sentences subtest.

Bilingual status. A comprehensive language background questionnaire (Gonzalez-Barrero & Nadig, submitted) including participants' history of language exposure over their lifetime, current amount of language exposure, and proficiency in all the languages of exposure was administered to parents to aid in the determination of participants' bilingual status. However, percent of lifetime exposure to two languages on its own did not accurately reflect fluent bilingualism, especially for the ASD participants. Given this, determination of bilingual status was based on a combination of several indices: (a) greater than 20 percent of lifetime exposure to each language (i.e., dominant and non-dominant language) according to parent-report, (b) the ability to complete at least 5 out of 8 language-based tasks of the testing protocol in both languages, (c) a score of 3 or 4 on a 4-point scale of language proficiency in each of the bilingual children languages as rated by parents, and (d) mean ratings of 2 or above on a 4-point scale of language proficiency according to the assessment of three external raters. These raters were blind to bilingual status and based their ratings on observation of video records of the testing sessions. The 20% cut-off to determine bilingual status was based on evidence from Pearson, Fernandez, Lewedeg, and Oller (1997), who found that it may be difficult to elicit utterances in a language for children with language exposure lower than 20%. Concerning age of first exposure to each language, there were 7 simultaneous bilingual children in the group with ASD (i.e., children who have been exposed to both languages before three years of age; Paradis, Genesee, & Crago, 2011) and 3 early sequential bilingual children (mean age of first exposure to L2 = 4.3 years), whereas in the TYP bilingual group there were 8 simultaneous bilinguals and 2 early sequential bilinguals (mean age of first exposure to L2 = 3 years).

Monolingual children were defined as those who: (a) had not been exposed to a language other than English (or French for the French-L1 participants) more than 20 percent of their lifetime, (b) if exposed to an L2, had scores of 1 or 2 on a 4-point scale of language proficiency completed by parents, or (c) could not complete the testing protocol in both languages, even if they did not meet criteria 1 and 2. The vast majority of monolingual children who had some language exposure to an L2 over their lifetime (9 in the TYP group and 8 in the group with ASD) did not have an amount of exposure greater than 20%. One child in the monolingual ASD group had a history of exposure to a second language of 22%, however his parent proficiency rating in the L2 was low (1 out of 4) and this child could not maintain a simple conversation with a research assistant in the second language. Therefore, he was kept in the monolingual group.

Procedure

Ethics approval was obtained from a university institutional review board. Parental consent and children's assent was also obtained. This project was part of a larger study examining the effects of bilingualism on cognition and language in children with ASD. Participants were tested individually by trained research assistants in a quiet room at a university lab or at the participant's home. Testing consisted of one 2-hour session for monolingual participants, and two sessions for bilingual participants (because language tests were administered in separate sessions for each language). The protocol of language tests, described in detail in Gonzalez-Barrero and Nadig (submitted), included subtests from the CELF-4, a non-word repetition task, a picture-description task, and a free-play language sample. Sessions for the bilingual participants took place approximately 2 weeks apart, and one language (i.e., English, French, or Spanish) was used exclusively during each session. Breaks and snacks were provided as needed and participants received a small gift upon completion of each session.

EF measures (BRIEF, DCCS, and verbal working memory) were administered during the first testing session. Experimental EF tasks (DCCS task and verbal working memory) were administered in the participant's dominant language. Also during the first testing session, parents completed the BRIEF in their preferred language (English, French, or Spanish), as well as the SCQ and the child language background questionnaire.

Executive function assessments.

Parent ratings of executive function in daily life. The Behavior Rating Inventory of Executive Functioning (BRIEF, parent form; Gioia, Isquith, Guy, & Kenworthy, 2000a) was administered to parents. The BRIEF can be used to assess EF in daily life for TYP children as well as children with neurodevelopmental disorders, aged 5 to 18 years (Gioia et al., 2000a). Its reliability (e.g., internal consistency = .80 to .98) and validity are considerate adequate (Gioia et al., 2000b). This questionnaire has been used in previous research examining EF in children with ASD (e.g., Leung, Vogan, Powell, Anagnostou, & Taylor, 2016) and is considered an ecologically valid measure of executive functioning (Kenworthy et al., 2008). The test consists of 86 questions that assess 8 clinical dimensions: inhibit, shift, emotional control, initiate, working memory, plan/organize, organization of materials, and monitor. These subscales can be grouped into two indexes (i.e., Behavioral Regulation Index and Metacognition Index), which, when combined, comprise an overall EF score (i.e., Global Executive Composite, GEC). T-scores are calculated and provide information about whether the child exhibits clinical significant executive dysfunction in daily life (i.e., t-score > 65). Special attention was given to the shift and working memory subscales given our research questions.

Performance measure of set-shifting. Set-shifting was assessed using a computerized version of the Dimensional Change Card Sort task (DCCS; Zelazo, 2006) developed with E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The task was presented on a Toshiba Satellite A350 laptop with a 16-in. screen. The computerized testing modality was used to decrease the social demands of face-to-face interactions for children with ASD, as suggested by prior research on set-shifting in this population (e.g., Ozonoff, 1995; Pascualvaca et al., 1998). Furthermore, the use of a computerized task allowed us to gather precise reaction time.

Our setup of the computerized task followed Bialystok and Martin (2004) and Barac and Bialystok (2012). A sticker with a red boat was placed on the right side of the laptop keyboard (“p” key) and a sticker with a blue rabbit was placed on the left side (“w” key). A black cover on the keyboard was used to hide all other keys from view. Stimuli presented on the screen consisted of red rabbits and blue boats, thus the stimuli to be sorted was never identical to the indicators on the keyboard. Participants sat in front of the laptop, the experimenter sat next to it.

Our computerized DCCS task included seven phases, adhering to the administration guidelines of Zelazo (2006). In the first phase (i.e., demonstration phase), the experimenter introduced the task and explained the rules to the participant. In the second phase (i.e., practice phase, 3 trials), the child was asked to perform the task: the stimuli appeared on the center of the screen and participants were instructed to press the appropriate key to categorize the image (e.g., "*We are going to play the color game, in the color game all the blue ones go here* -pointing to the key with the blue rabbit- *and all the red ones go here* -pointing to the key with the red boat. *Here is a blue one, where does it go?*). The image remained on the screen until the participant pressed a key to respond. An inter-trial interval of 2000 ms preceded the next trial. During the practice phase, visual feedback on the screen (e.g., a cartoon of a checkmark signaling "good work") and verbal feedback from the experimenter were provided after each trial. To continue to the pre-switch phase, the child was required to have at least 2 out of 3 correct practice trials. In the third phase (i.e., pre-switch phase, 6 trials), the experimenter reminded the child about the rule (e.g., sort by color) to reduce memory demands, and the participant was again asked to sort images following this rule. No feedback was provided during the experimental trials. In the fourth phase (i.e., post-switch phase, 6 trials), the child was asked to change the sorting rule and to now classify the images by another dimension (e.g., "*okay, now we are not going to play the color game anymore; now we are going to play the shape game*…" Zelazo, 2006, p. 297). The order of presentation of sorting dimension was counterbalanced, with half of children receiving color first and the other half receiving shape first. Following Zelazo (2006), participants passed the post-switch phase if they were able to correctly sort on 5 out of 6 trials.

A more advanced phase was administered to participants who passed the post-switch phase. This "border version" (Zelazo, 2006) was demonstrated in a fifth phase where children were asked to sort by color if the image displayed on the screen had a border, or by shape if the image had no border, consequently referred to as "mixed condition" in the literature. A practice border version with feedback was presented (i.e., sixth phase, 2 trials). If the child obtained at least one correct response, he or she was asked to complete the border version (seventh phase, 12 trials). Children were considered to pass the border version if they correctly sorted on at least 9 out of 12 trials.

Accuracy (i.e., passing or failing) on the pre-switch, post-switch, and border version phases of the DCCS task, along with reaction time (RT) in milliseconds were analyzed. Following Diamond and Kirkham (2005), for RT analyses only correct trials were considered and trials that were less than 200 ms or 2.5 SD above the mean for each group were not included as these are considered to be either too quick or too slow to reflect processing of the stimuli. For TYP participants, after removing incorrect trials (17%) and trials that did not meet the response time criteria just noted (3%), 80% of possible trials remained for inclusion in RT analyses. For participants with ASD, removal of incorrect trials = 15%, and trials excluded given response time criteria = 3%, 82% were kept for RT analyses. The mean RT of the last two trials of the pre-switch phase was subtracted from the mean RT of the first two trials of the post-switch phase to obtain a *RT switch cost* difference score. This approach to RT analysis follows that used by Dichter and colleagues (2010) with a group of monolingual children with ASD. In addition, mean RT for the border version (i.e., mixed condition) was examined to investigate if the bilingual advantage reported by Barac and Bialystok (2012) for TYP bilingual children could be replicated in the present study.

Performance measure of verbal working memory. To assess short-term and working memory, the number repetition subtest of the Clinical Evaluation of Language Fundamentals

(CELF-4; Semel, Wiig, & Secord, 2003), along with its French (Évaluation clinique des notions langagières fondamentales; Wiig, Secord, Semel, Boulianne, & Labelle, 2009), or Spanish versions (CELF 4 Spanish Edition; Semel, Wiig, & Secord, 2006) were used. Administration and scoring followed the guidelines provided in the CELF-4 manual. In this task there are 8 number lengths (ranging from two to nine digits, with 2 items at each length), that the child was asked to repeat immediately after the examiner (i.e., number repetition forward). In addition, there are 7 number lengths (ranging from two to eight digits, with 2 items at each length) that the child was asked to repeat in reverse order (i.e., number repetition backwards). Number repetition forward is considered a measure of short-term memory, whereas number repetition backwards is a more complex measure that taps into both short-term and working memory (e.g., Eigsti, 2011; Engel de Abreu, 2011). Although we were interested in the working memory measure of number repetition backwards, both subtests were given to maintain standard administration involving a gradual increase in complexity. Only the scaled score for number repetition backwards was considered for analysis.

Our research question was whether bilingualism can mitigate set-shifting difficulties in children with ASD, via a bilingual advantage. To recap, we hypothesized that both monolingual and bilingual children with ASD will exhibit impaired performance relative to typically-developing children on measures of set-shifting. Given prior findings (Leung & Zakzanis, 2014), we expected set-shifting impairments to be larger on a rating measure of EF in daily life (BRIEF) than in an experimental task (DCCS task). The critical hypothesis was that, despite this background of a general set-shifting impairment, bilinguals with ASD will perform better on set-shifting than monolinguals with ASD. Finally, we hypothesized that bilinguals and monolinguals with ASD will not differ with respect to the control EF domain of verbal working memory. We did not have specific hypotheses as to how ratings of EF in daily life and performance or experimental measures of EF would compare, as they have not previously been compared directly when investigating a bilingual advantage.

Results

Data Analysis

To examine the impact of bilingualism on EF, two-way (2 Diagnostic group [TYP, ASD] x 2 Language status [monolingual, bilingual]) analyses of variance (ANOVA) were conducted on continuous variables (i.e., shifting and working memory scores from the BRIEF, DCCS RTs, and working memory scaled score from the CELF). For categorical variables (i.e., DCCS task accuracy/passing), Fisher's exact tests were performed. In addition, correlation analyses were conducted to explore the relation between scores based on parent-report and direct testing of EF measures. The assumptions of statistical tests were checked and parametric tests where used where appropriate. In cases where the assumption of normality was not met, Kendall's tau was used for correlations. Following a conservative approach, Bonferroni corrections were used for measures where multiple tests were performed (e.g., DCCS task and BRIEF) to avoid Type I errors. Therefore, the applicable alpha level for significance is presented for each measure.

Means and standard deviations on all EF measures are provided in Table 2. To assess the validity of the responses provided by parents on measures of EF in daily life we used the inconsistency scale from the BRIEF. This scale compares similar items to which parents are expected to provide consistent answers. A score derived from this scale indicates whether the questionnaire can be considered acceptable, questionable, or inconsistent. Results showed that for the majority of children, scores were consistent. However, two participants in the monolingual ASD group showed questionable scores and were therefore excluded from analyses involving the BRIEF. Two statistical analyses were conducted using the BRIEF scores (i.e., shift subscale and working memory subscale); therefore the alpha level for significance was set at p <.025.

Insert Table 2 about here

Set-Shifting (BRIEF)

A 2 (Diagnostic group) x 2 (Language status) ANOVA on shift subscale scores, obtained via parent report on the BRIEF, revealed only a significant main effect of diagnostic group, $F(1, 34) = 43.12$, p = .<.001, partial $\eta^2 = .56$, where TYP participants obtained lower scores than participants with ASD. As previously described, higher scores on the BRIEF (i.e., t-scores > 65) reflect a greater degree of difficulty in EF (Gioia et al., 2000a). Conversely, there was no significant main effect of language, $F(1, 34) = .08$, $p = .78$, partial $\eta^2 = .00$, nor was the interaction between diagnostic group and language significant, $F(1, 34) = .88$, $p = .36$, partial $\eta^2 = .03$.

Set-Shifting (DCCS task accuracy/passing)

Results showed that children in all groups passed the pre-switch phase, and the majority of participants in each group passed the post-switch phase (passing criteria: 5 correct trials out of 6). Consequently, following Diamond and Kirkham (2005), statistical analyses were not conducted on accuracy data in these phases (available in Table 2) given the minimal variation observed. The percentage of participants passing the pre-switch, post-switch and border versions of the DCCS task using the criteria of Zelazo (2006) are presented in Figure 1. For the border version (passing criteria: 9 correct trials out of 12), Fisher's exact test showed a significant difference across the four groups (p = .024). To further examine this difference, pairwise Fisher's exact tests were performed. Results revealed that, in line with our second hypothesis, the number of bilinguals with ASD who passed the border version was significantly higher than the number of monolinguals with ASD who passed (p = .026). In fact, the bilingual ASD group also outperformed the TYP bilingual group on the border version of the DCCS task (p = .009). On the other hand, the TYP groups did not differ significantly by language status (p = .347). The monolingual TYP and monolingual ASD groups performed similarly (p =.619).

Insert Figure 1 about here

Set-Shifting (DCCS RTs)

The significance level for analyses concerning RTs was set at p <.025, given two statistical analyses conducted, one on switch cost between pre- and post-switch phases and the other on mean RT of the border phase. A 2-way ANOVA was performed on RT switch cost (i.e., mean RT of first two trials from post-switch phase minus mean RT of last two trials from pre-switch phase). The main effect of diagnostic group was not significant, $F(1, 33) = 2.97$, p = .09, partial η^2 = .08, although a trend was observed in which the TYP groups were faster than the groups with ASD. The main effect of language status, $F(1, 33) = .46$, p = .50, partial η^2 = .01, was not statistically significant, contrary to previous studies with TYP children (e.g., Barac & Bialystok, 2012). Finally, the interaction between diagnostic group and language group was not significant ($F(1, 33) = .46$, p = .50, partial η^2 = .01).

For the border version mean RTs were analyzed using a 2-way ANOVA. The main effect of diagnostic group was not significant, $F(1, 28) = 2.82$, p = .10, partial η^2 = .09, and the main effect of language revealed that, contrary to previous findings, the monolingual TYP and ASD groups responded faster than the bilingual TYP and ASD participants in the border version , $F(1, 28) = 6.32$, $p = .018$, partial $\eta^2 = .18$. The interaction between diagnostic group and language was not statistically significant, $F(1, 28) = .02$, $p = .89$, partial $\eta^2 = .00$ (see Figure 2).

Insert Figure 2 about here

Verbal Working Memory (BRIEF)

A 2-way ANOVA conducted to examine working memory in everyday life, as assessed by parent-report on the corresponding BRIEF subscale, revealed a significant main effect only for diagnostic group, $F(1, 34) = 12.63$, $p = .00$, partial $\eta^2 = .27$, with the ASD group having higher scores, reflecting poorer performance, relative to the TYP groups. The main effect of language, $F(1, 34) = .005$, $p = .95$, partial $\eta^2 = .00$, and the interaction, $F(1, 34) = 1.39$, $p = .25$, partial $\eta^2 = .04$, were not statistically significant.

Verbal Working Memory (CELF-4)

Participants' performance on the experimental verbal working memory measure (i.e., scaled score from the number repetition backwards subtest of the CELF-4) was examined. In line with our third hypothesis, results revealed that there was no significant main effect for diagnostic group, $F(1, 36) = .40$, $p = .53$, partial $\eta^2 = .01$, or for language, $F(1, 36) = .96$, $p = .34$, partial $\eta^2 = .03$, nor a significant interaction between diagnostic group and language for this variable, $F(1, 36) = .003$, $p = .95$, partial $\eta^2 = .00$.

Correlations between BRIEF Scores and Experimental Tasks

Finally, Kendall's tau-b correlations were run to examine the relation between parent report of EF behaviors in daily life as measured by the BRIEF and experimental tasks of EF. For the set-shifting variables (i.e., BRIEF shifting subscale and DCCS task switch cost) there was not a statistically significant correlation for the TYP group, $\tau_b = .122$, $p = .505$, or for the group with ASD, $\tau_b = -.226$, $p = .252$. Similarly, results for the working memory measures (i.e., BRIEF working memory subscale and number repetition subtest from the CELF-4) revealed no significant correlations for the TYP group, $\tau_b = .108$, $p = .530$, or for the group with ASD, $\tau_b = -.226$, $p = .206$.

Discussion

In this study we tested for the first time whether being proficiently bilingual can mitigate the set-shifting impairment observed in children with ASD. Our first hypothesis was that both monolingual and bilingual children with ASD would exhibit impaired performance relative to typically-developing children on measures of set-shifting. This was the case for a parent-report measure of set-shifting (BRIEF), which confirmed clinical levels of difficulty. However, we did not observe a decrement at the level of diagnostic group in an experimental task of set-shifting (the DCCS task), given the high rate of passing by bilingual children with ASD. Our second hypothesis proposed that if the bilingual advantage extends to ASD, bilingual children with ASD should perform better on set-shifting tasks relative to monolingual children with ASD, who do not need to coordinate two linguistic systems. Our results from the DCCS task support this hypothesis: a significantly higher percentage of bilingual children than monolingual children with ASD passed the border version of the DCCS task. However, no bilingual advantage was observed in parent report of set-shifting difficulties in daily life on the BRIEF. Our third hypothesis was that bilingual children with ASD would not differ from their monolingual peers with ASD on tasks that assess verbal working memory, included as a control EF not expected to be affected by bilingualism. This hypothesis was supported by our results on both a performance task (number repetition backwards) and parent report of working memory in daily life. We now consider each of these findings in turn.

Set-Shifting in Daily Life

Regardless of language status, and as anticipated, children with ASD exhibited set-shifting difficulties in everyday life relative to their same-age TYP peers, as shown by results from the BRIEF. Both groups with ASD scored within the dysfunction range on this subscale, whereas the TYP participants' scores were in the average (typical) range. This confirmation of set-shifting difficulties in daily life on a rating measure is consistent with a substantial body of work indicating that set-shifting is a "peak" area of EF impairment in ASD (Gioia et al., 2002; Granader et al., 2014), and that rating measures are the most sensitive to detect them (Leung & Zakzanis, 2014; Toplak, West, & Stanovich, 2013). By establishing set-shifting impairment, this first finding sets the backdrop for testing whether bilingualism may mitigate these difficulties in ASD. At the same time, similar performance of bilingual and monolingual children with ASD on the BRIEF was contrary to our second hypothesis of a bilingual advantage in set-shifting skills. We did however observe a bilingual advantage for children with ASD on the DCCS task discussed next. We return to the (lack of) relationship between rating and performance or experimental task measures later in the Discussion.

Set-Shifting in an Experimental Task

The key finding of this study was a bilingual advantage on the advanced border version of the DCCS task, where bilingual children with ASD passed at a significantly higher rate than monolingual children with ASD, upholding our second hypothesis. The fact that we detected this difference on ability to complete the border, rather than simpler post-switch phase is likely due to the age of our participants. Notably, this is also the phase of the DCCS task where significant effects were reported by Faja and Dawson (2014), who tested monolingual children with ASD of similar ages. Although novel and revealing, this finding should be taken as preliminary evidence given our small sample size. Replication with a larger but similarly well-characterized sample to our own (i.e., stringent confirmation of bilingual status, matching on NVIQ and maternal education as well as age, description of language abilities, etc.) is needed to confirm a bilingual advantage in children with ASD.

Two other aspects of our DCCS task results, however, ran counter to our predictions. Following our first hypothesis, we expected typically-developing children to outperform both groups of children with ASD. Yet, in our sample, the DCCS task did not capture the impairment in set-shifting reported by previous studies using the same task with monolingual children with ASD (Faja & Dawson, 2014; Yi et al., 2012). This is likely because the DCCS task is not a highly sensitive measure of set-shifting impairment, with a 95% confidence interval of [-1.82, 2.85] pooled over the two studies just mentioned (Leung & Zakzanis, 2014). That is, the DCCS task does not reliably show better performance for typically-developing children relative to children with ASD. Other studies have also reported spared set-shifting abilities in children and adults with ASD on some experimental tasks (e.g., Edgin & Pennington, 2005; Hill & Bird, 2006). Second, while we detected a bilingual advantage in children with ASD, this was not the case for typically-developing children.

Why might the bilingual advantage be observable only in children with ASD and not in typically-developing children in our study? One possibility is related to testing modality (e.g., the computer presentation employed here, vs. live administration by an experimenter). As discussed in the introduction, children with ASD have been found to exhibit better performance on EF tasks that employ computerized versus live administration by an examiner (Kenworthy et al., 2008; Ozonoff, 1995; Pascualvaca et al., 1998), which has been attributed to the reduced social demands required of them. Typically-developing children, on the other hand, do not perform better on computerized tasks and may actually benefit from social motivation when interacting with an experimenter (Kenworthy et al., 2008). In fact, during the administration of the computerized DCCS task in our study, it was observed that some TYP children (5 monolinguals, 2 bilinguals) looked at the experimenter as if awaiting confirmation on their performance during the border version (although no feedback was provided on experimental trials). Although this was occasionally observed, this behavior was less prominent among the ASD groups (1 monolingual, 2 bilinguals). In addition, children in both ASD groups seemed more motivated and engaged with the computerized task than did their TYP peers; some of them verbally expressed high levels of enjoyment and even wanted to continue "*playing the computer game*".

A second possible reason for the enhanced performance of the bilingual ASD group relative to the bilingual typically-developing group is the degree of explicitness of the task. It has been reported that individuals with ASD perform better on tasks that provide clear and explicit rules (e.g., DCCS task) relative to tasks where rules are implicit and need to be inferred (e.g., Wisconsin Card Sorting Task, WCST), (Van Eylen et al., 2011). Since sorting criteria for the DCCS task were explicitly stated, this may explain the relatively high performance of the ASD groups in the pre-switch and post-switch phases. Yet, it was in the border version where the presence or absence of the visual cue of a border indicated the rule to use, that bilinguals with ASD outperformed monolinguals with ASD as well as typically-developing bilinguals. It is possible that this visual cue was perceived as "explicit" and easier for bilingual (but not monolingual) children with ASD to use given their bilingual advantage, but was not perceived as an explicit cue by typically-developing children.

A third contributor to the lack of a bilingual advantage for typically-developing children in the current study may be related to the participants included in the monolingual typically-developing group. All of these children spoke French, whereas the other three groups of participants included both French and English (dominant) speakers. This difference was unavoidable because the study was conducted in Montreal, where it is possible to recruit monolingual French-speaking children (as defined in the methods). It was not, however, possible to find typically-developing school-age children who were English monolinguals without a significant amount of exposure to French, as children in regular English language schools receive French instruction and hear French in multiple social contexts. English and French school boards are separate in Quebec and may provide different background experiences. Consequently, the comparison between monolingual and bilingual typically-developing children was not as well matched in this respect as it was for children with ASD.

Finally, a fourth reason why a bilingual advantage may not have been observed for typically-developing 6- to 9-year-olds in our study is that the task may no longer have been developmentally sensitive for them. Bialystok (1999) used the DCCS task (live card administration, original procedure) and found a bilingual advantage for typically-developing 4- and 5-year-olds. Barac and Bialystok (2012) used a computerized color-shape switching task with typically-developing 5- to 7-year-olds, also documenting a bilingual advantage in switch cost. As our primary research questions focused on children with ASD our sample was older, in line with prior studies on children with ASD, and we employed a very simple task of set-shifting as a starting point to observe potential differences related to bilingualism.

DCCS Response Time Measures

In research on typically-developing children and adults, the bilingual advantage in set-shifting is most commonly found in RT measures (e.g., Prior & MacWhinney, 2009; but see Morton & Harper, 2007). The general finding is that bilinguals exhibit smaller switching costs (delays when changing from one rule to another, or on switch blocks compared to same rule blocks) relative to monolinguals during set-shifting paradigms (Prior & MacWhinney, 2009). It is important to note that our examination of RT measures was exploratory given that we employed the original procedure of the DCCS task although it was computer-administered, and RT measures are not involved in the conventional scoring of this task. We presented a maximum of 24 trials compared to 200 trials over 5 blocks (with either the same rule or a switch in rule) that are used in many card-sort tasks focusing on RT measures. Keeping that in mind, we did not find a bilingual advantage for either the ASD or TYP groups with respect to switch cost between pre- and post-switch phases of the task, which has been reported previously for typically-developing children in other types of card sort tasks (e.g., Bialystok & Viswanathan, 2009; Prior & Macwhinney, 2009). There was a trend for children with ASD to present increased switch cost times relative to children with typical development, which is in line with previous studies reporting overall slower RTs in monolingual children with ASD relative to typically-developing controls on computerized card sort tasks (e.g., Dichter et al., 2010).

Similarly, for the border version (a phase where the rule to be used was mixed), no bilingual advantage was found, contrary to prior reports for typically-developing children using color-shape card switch tasks (Barac & Bialystok, 2012). In fact, the reverse was found: In our study, both groups of monolinguals tended to respond more quickly on border version trials than did bilinguals with typical development or ASD. Once again, this finding should be interpreted with caution since the task procedure was not comparable to other studies analyzing RTs.

However, at least for the bilingual ASD group, who were significantly better at passing this phase, this finding appears to represent a speed/accuracy trade-off.

Working Memory in Daily Life and on an Experimental Task

Results from the BRIEF working memory subscale revealed that children with ASD exhibited difficulties in this domain in their everyday life, as reflected in somewhat elevated scores. However, while set-shifting scores of children with ASD were in the executive dysfunction range, this was not the case for working memory scores, replicating the finding that set-shifting is a peak area of difficulty in ASD (Gioia et al., 2002, Granader et al., 2014). Finally, our third hypothesis that working memory should not be affected by being proficiently bilingual was upheld, both on the BRIEF and on our performance measure of verbal working memory. Scores on the number repetition backwards subtest of the CELF-4 showed comparable performance between all four groups of participants. This also in line with previous work showing verbal working memory to be a spared aspect of EF in children with ASD (Faja & Dawson, 2014).

Relation between Experimental Tasks and Skills in Daily Life

The dissociation between impairments in EF skills in everyday life and intact performance on the experimental EF tasks has been discussed by autism researchers (e.g., Kenworthy et al., 2008). These researchers have advocated for more comprehensive and ecologically valid approaches in the study of cognitive flexibility in ASD. Our study was novel in employing both experimental or performance measures and a rating measure of EF skills in daily life to investigate a bilingual advantage. We found weak, non-significant correlations between switch cost on the DCCS task and the shift subscale of the BRIEF (TYP, $\tau_b = .122$, ASD, $\tau_b = -.226$). In populations other than ASD, the dissociation between performance-based and rating measures of EF is widely documented (Isquith, Roth, & Gioia, 2013) and was the subject of a practitioner review by Toplak, West, and Stanovich (2013). These authors reviewed 20 studies that compared both performance-based and rating measures of EF, including child and adult, nonclinical and clinical samples, particularly attention deficit hyperactivity disorder (ADHD) and traumatic brain injury. Across these studies, they report a median correlation of only .19 between performance-based and rating measures of EF, with only 24% of the correlations reported in these studies reaching significance (Toplak, West, & Stanovich, 2013). The lack of a strong relationship between these measures across a range of populations indicates that they in fact measure different constructs. Performance-based measures provide optimal conditions and structure the problem-space with a well-defined task and often an explicit rule to follow. In contrast, the use of analogous EF abilities in daily life requires applying them as the need arises to meet ones' goals. Toplak and colleagues (2013, p. 140) deem these complementary measures given that each provide important information about EF abilities, with performance-based measures serving as an index of efficiency of processing, and rating measures reflecting success in rational goal pursuit in natural environments.

How can we bridge the divide between the significant impairment in EF abilities reported in daily life, and the good or even enhanced (in the case of bilingual children with ASD) performance on experimental tasks? It may be possible to use tasks where children with ASD present spared performance to improve EF abilities in daily life. For instance, if children with ASD are able to switch between sets using a computer-based task, this ability could be built on through intervention. One could gradually shift from computer-based cognitive flexibility paradigms to the integration of EF skills in daily-life, for example by adding one additional dimension of complexity at a time.

Contributions and Limitations of the Current Study

The present study is the first to examine the effects of bilingualism on the set-shifting and working memory skills of proficient bilingual children with ASD. To ascertain children's bilingual status, this study employed a rigorous 4-step procedure including parent-report of amount of language exposure, parent ratings of language proficiency, direct testing in both languages of the bilingual participants, and ratings of language proficiency by external judges. In addition, all four groups of participants were matched on NVIQ, age, and level of maternal education, and to the extent possible, dominant language. Participants were from the same city and for the most part shared similar cultural and educational backgrounds. Importantly, both experimental and rating measures of set-shifting and working memory in daily life were gathered.

Taken together, these methodological strengths allow us to interpret results as due to bilingual exposure rather than confounding factors. However, this methodological rigor, as well as the difficulties inherent in recruiting special populations, limited the sample size in each group. It is possible that with only 10 children per subgroup we did not have enough power to detect significant differences between monolingual and bilingual typically-developing children, yet as discussed earlier, we did have enough power to detect differences amongst children with ASD. Potential differences in the strength of a bilingual advantage across typically-developing children and those with developmental disorders should be explored in future research.

Generalizability of Bilingual Advantage in Executive Function for Children with ASD

We present the first evidence that a bilingual advantage (Bialystok & Martin, 2004; Bialystok & Viswanathan, 2009) can be observed in children with ASD. It is important to note the limitations on to whom and where this effect may generalize. Our participants with ASD were school-age and did not have intellectual delay, with 7/20 meeting research criteria for language impairment. Crucially, to have bilingual status in our study participants had to be proficiently bilingual as described earlier, and they lived in a society where the two languages in question were official languages used in multiple contexts.

The effect of bilingualism on EF in ASD was selective: there was an advantage on an experimental task of set-shifting but not for verbal working memory. The selective influence of bilingualism on specific cognitive skills had been discussed and this same pattern has been found in typically-developing individuals (e.g., Bialystok, 2009).

We detected a bilingual advantage in set-shifting in children with ASD using a relatively brief and simple computerized task that reduced social demands and provided explicit rules. This task was also untimed, that is, participants could respond when ready. There was a trend for children with ASD to demonstrate longer switch costs. The bilingual children with ASD who were the most accurate even on the advanced border phase of the DCCS task also tended to be slower on all response time measures. This longer response time raises the possibility that an ASD bilingual advantage may not be observed in tasks or situations involving time pressure.

We also presented evidence that the bilingual advantage in set-shifting for children with ASD does not extend to ratings of set-shifting in daily life. As discussed above this is a more general phenomena where performance on experimental EF tasks are minimally related to the application of EF in daily life (Toplak et al., 2013). It is possible that the cognitive processes engaged and enhanced by the proficient use of two languages are not related to the broader EF skills tapped by rating measures. Alternatively, any potential bilingual advantage effects in ratings of EF in daily life may only be observed at a much later age, with accumulated bilingual experience. This is another important question for future research.

In a separate study with a largely overlapping sample (Gonzalez-Barrero & Nadig, 2017b) we report a similar bilingual advantage on a verbal fluency task (e.g., "*name all of the animals you can think of*"), which assesses both lexical skills and multiple EF skills. Thus, it appears this bilingual advantage in ASD is not limited to the nonverbal domain or to the particular task we employed in this study. On the verbal fluency task bilinguals outperformed monolinguals with ASD in the total number of words produced, but not on switching between semantic clusters of words. Consequently we suggested that the bilingual advantage observed on the verbal fluency task reflects generativity and initiation, rather than set shifting. Further work is clearly warranted to explore the extent and reliability of bilingual advantages for children with ASD.

The present findings extend a bilingual advantage in set-shifting skills, as measured by an experimental task, to children with ASD. We demonstrate that, contrary to common belief, bilingualism is not harmful for children with ASD (e.g., with respect to their language abilities, Kay-Raining Bird, Genesee, & Verhoeven, 2016), and in fact, may provide some advantages, such as mitigating prominent set-shifting difficulties. There is growing evidence that bilingualism acts as a protective factor that delays the onset of dementia in normal aging (for a review see Bialystok, Abutalebi, Bak, Bruke, & Kroll, 2016). Likewise, bilingualism, under the right conditions, may act as a protective factor for certain EF difficulties in populations with neurodevelopmental disorders. If replicated, this finding could provide critical evidence to inform educational decisions taken by the increasing number of families with children with ASD for whom the use of two or more languages is a valued practice (Kay-Raining Bird et al., 2016).

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556-559.

Baio, J. (2012). *Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, 14 Sites, United States, 2008*. Morbidity and Mortality Weekly Report. Surveillance Summaries. *61*(3), 1-19. Atlanta, GA: Centers for Disease Control and Prevention.

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child Development, 83*, 413-422. doi: 10.1111/j.1467-8624.2011.01707.x

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child development, 67*, 1816-1835. doi: 10.1111/j.1467-8624.1996.tb01830.x

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*, 636-644. doi: 10.1111/1467-8624.00046

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition. 12*, 3-11. doi: 10.1017/S1366728908003477

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 7, 325-339. doi: 10.1111/j.1467-7687.2004.00351.x

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition*, *112*, 494-500. doi: 10.1016/j.cognition.2009.06.014

Buac, M., Gross, M., & Kaushanskaya, M. (2016). Predictors of processing-based task performance in bilingual and monolingual children. *Journal of Communication Disorders*, *62*, 12-29.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*, 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*, 741-748. doi: 10.1111/1469-7610.00770

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*, 59-86. doi: 10.1016/j.cognition.2006.12.013

Delcenserie, A., & Genesee, F. (2016). The effects of age of acquisition on verbal memory in bilinguals. *International Journal of Bilingualism*. doi: 10.1177/1367006916639158

Diamond, A., & Kirkham, N. (2005). Not quite as grown-up as we like to think parallels between cognition in childhood and adulthood. *Psychological Science, 16*, 291-297. doi: 10.1111/j.0956-7976.2005.01530.x

Dichter, G. S., Radonovich, K. J., Turner-Brown, L. M., Lam, K. S. L., Holtzclaw, T. N., & Bodfish, J. W. (2010). Performance of children with autism spectrum disorders on the Dimension-Change Card Sort task. *Journal of Autism and Developmental Disorders, 40*, 448-456. doi: 10.1007/s10803-009-0886-1

Edgin, J. O., & Pennington, B. F. (2005). Spatial cognition in autism spectrum disorders: Superior, impaired, or just intact? *Journal of Autism and Developmental Disorders*, *35*, 729-745. doi: 10.1007/s10803-005-0020-y

Eigsti, I. M. (2011). Executive functions in ASD. In D. A. Fein (Ed.). *The neuropsychology of autism*, (pp. 185-203). New York: Oxford University Press.

Elin Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*, 426-445. doi: 10.1177/136700691140320

Elin Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., ... Chilingaryan, G. (2011). Sensitivity and specificity of French language and processing measures for the identification of primary language impairment at age 5. *Journal of Speech, Language, and Hearing research, 54*, 580-597. doi: 10.1044/1092-4388(2010/09-0196)

Engel de Abreu, P. M. (2011). Working memory in multilingual children: Is there a bilingual effect?. *Memory*, *19*, 529-537. doi: 10.1080/09658211.2011.590504

Faja, S., & Dawson, G. (2014). Performance on the dimensional change card sort and backward digit span by young children with autism without intellectual disability. *Child Neuropsychology*, *20*, 692-699. doi: 10.1080/09297049.2013.856395

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*, 483-527.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000a). *Behavior rating inventory of executive function (BRIEF)*. Lutz, FL: PAR.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000b). Test review Behavior Rating Inventory of Executive Function. *Child Neuropsychology, 6*, 235-238.

Gioia, G. A., Isquith, P. K., Kenworthy, L., & Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*, 121-137. doi: 10.1076/chin.8.2.121.8727

Gonzalez-Barrero, A. M. & Nadig, A. (2017a). *Bilingual children with Autism Spectrum Disorders: The impact of language exposure in lexical and morphological skills at school age.* Manuscript submitted for publication.

Gonzalez-Barrero, A. M., & Nadig, A. (2017b). Verbal fluency in bilingual children with Autism Spectrum Disorders. *Linguistic Approaches to Bilingualism, 7,* 460-475.

Granader, Y., Wallace, G. L., Hardy, K. K., Yerys, B. E., Lawson, R. A., Rosenthal, M., ... & Schultz, R. T. (2014). Characterizing the factor structure of parent reported executive function in autism spectrum disorders: The impact of cognitive inflexibility. *Journal of Autism and Developmental Disorders*, *44*, 3056-3062.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67-81. doi: 10.1017/S1366728998000133

Hill, E. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, *24*, 189-233. doi:10.1016/j.dr.2004.01.001

Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia*, *44*, 2822-2835. doi: 10.1016/j.neuropsychologia.2006.06.007

Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child*, *2*, 125-132. doi:10.1080/21622965.2013.748389

Kay-Raining Bird, E., Genesee, F., & Verhoeven, L. (2016). Bilingualism in children with developmental disorders: A narrative review. *Journal of Communication Disorders*, *63*, 1-14. doi: 10.1016/j.jcomdis.2016.07.003

Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding executive control in autism spectrum disorders in the lab and in the real world. *Neuro*, 18, 320-338. doi: 10.1007/s11065-008-9077-7

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16*, 287-308. doi: 10.1080/01690960042000058

Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders: A review of research in the last decade. *Psychological Bulletin*, *137*, 562-593.

Leung, R. C., Vogan, V. M., Powell, T. L., Anagnostou, E., and Taylor, M. J. (2016). The role of executive functions in social impairment in autism spectrum disorder. *Child Neuropsychology*, *22*, 336-344. doi: 10.1080/09297049.2015.1005066

Leung, R. C., & Zakzanis, K. K. (2014). Brief report: cognitive flexibility in autism spectrum disorders: a quantitative review. *Journal of Autism and Developmental Disorders*, *44*, 2628-2645. doi: 10.1007/s10803-014-2136-4

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 1426-1438. doi: 10.1007/s10803-007-0510-1

Minshew, N. J., & Goldstein, G. (2001). The pattern of intact and impaired memory functions in autism. *Journal of Child Psychology and Psychiatry, 42*, 1095-1101. doi: 10.1111/1469-7610.00808

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. doi: 10.1006/cogp.1999.0734

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*, 719-726. doi: 10.1111/j.1467-7687.2007.00623.x

Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingualism Education and Bilingualism, 13*, 597-616. doi: 10.1080/13670050.2010.488288

Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, *9*, 491-500. doi: 10.1037/0894-4105.9.4.491

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, *32*, 1081-1105. doi: 10.1111/j.1469-7610.1991.tb00351.x

Ozonoff, S., & Strayer, D. L. (2001). Further evidence of intact working memory in autism. *Journal of Autism and Developmental Disorders*, *31*, 257- 263. doi: 10.1023/A:1010794902139

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*, 232-258. doi: 10.1016/j.cogpsych.2012.12.002

Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual language development & disorders: A handbook on bilingualism and second language learning* (2nd ed.). Baltimore, MD: Paul H. Brookes Pub. Co.

Pascualvaca, D. M., Fantie, B. D., Papageorgiou, M., & Mirsky, A. F. (1998). Attentional capacities in children with autism: Is there a general deficit in shifting focus? *Journal of Autism and Developmental Disorders, 28*, 467-478. doi: 10.1023/A:1026091809650

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*, 41-58. doi: 10.1017/S0142716400009863

Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology*, *108*, 567-579. doi: 10.1016/j.jecp.2010.10.009

Prior, A., & Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*, 253-262. doi: 10.1017/S1366728909990526

Psychology Software Tools, Inc. (2012). [E-Prime 2.0]. Pittsburgh, PA. Retrieved from http://www.pstnet.com.

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language & Communication Disorders, 45*, 47-60. doi: 10.3109/13682820802647676

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition, 71*, 362-368. doi: 10.1016/j.bandc.2009.06.007

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 6*, 673-686. doi: 10.1111/j.1469-7610.1996.tb01459.x

Russo, N., Flanagan, T., Iarocci, G., Berringer, D., Zelazo, P. D., & Burack, J. A. (2007). Deconstructing executive deficits among persons with autism: Implications for cognitive neuroscience. *Brain and Cognition, 65*, 77-86. doi: 10.1016/j.bandc.2006.04.007

Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., & Murphy, D. G. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biological Psychiatry*, *59*, 7-16. doi: 10.1016/j.biopsych.2005.06.007

Schuh, J., & Eigsti, I. M. (2012). Working memory, language skills, and autism symptomatology. *Behavioral Sciences, 2*, 207-218. doi: 10.3390/bs2040207

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*, 131-143.

Van Eylen, L., Boets, B., Steyaert, J., Evers, K., Wagemans, J., & Noens, I. (2011). Cognitive flexibility in autism spectrum disorder: Explaining the inconsistencies? *Research in Autism Spectrum Disorders*, *5*, 1390-1401. doi: 10.1016/j.rasd.2011.01.025

Weismer, S. E., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders, 40*, 1259-1273. doi: 10.1007/s10803-010-0983-1

Williams, D. L., Goldstein, G., Carpenter, P. A., & Minshew, N. J. (2005). Verbal and spatial working memory in autism. *Journal of Autism and Developmental Disorders*, *35*, 747-756. doi: 10.1007/s10803-005-0021-x

Williams, D. L., Goldstein, G., & Minshew, N. L. (2006). The profile of memory function in children with autism. *Neuropsychology, 20,* 21-29. doi: 10.1037/0894-4105.20.1.21

Yi, L., Liu, Y., Li, Y., Fan, Y., Huang, D., & Gao, D. (2012). Visual scanning patterns during the Dimensional Change Card Sorting Task in children with autism spectrum disorder. *Autism Research and Treatment, 2012*, 1-11. doi: 10.1155/2012/123053

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297-301. doi: 10.1038/nprot.2006.46

Tables

Table 1

Participants Characteristics and Demographic Information

	Monolingual TYP *(n = 10)* *M (SD)*	Bilingual TYP *(n = 10)* *M (SD)*	Monolingual ASD *(n = 10)* *M (SD)*	Bilingual ASD *(n = 10)* *M (SD)*	*P*
NVIQ	111 *(7.44)*	108 *(7.61)*	109 *(11.64)*	111 *(9.63)*	.86
Age in months	93 *(7.56)*	95 *(8.93)*	100 *(11.94)*	97 *(7.23)*	.36
Range (Years)	*6;10 – 8;9*	*6;11 – 8;11*	*6;1 – 9;3*	*7;3 - 9;2*	
Maternal Education (Years)	15 *(1.73)*	16 *(1.49)*	16 *(2.06)*	15 *(3.03)*	.29
Gender	7 males	7 Males	10 Males	8 Males	.35
Dominant Language	10 French	6 French 4 English	6 French 4 English	5 French 5 English	.08
Receptive Vocabulary [abcd]	136 *(13.37)*	118 *(6.20)*	111 *(22.89)*	98 *(15.75)*	.00
Sentence Repetition[ac]	13 *(2.10)*	10 *(1.17)*	9 *(4.24)*	8 *(4.14)*	.01
SCQ[bcde]	2 *(1.27)*	5 *(3.44)*	18 *(6.67)*	19 *(3.63)*	.00

Note. TYP = typically-developing; ASD = autism spectrum disorders; NVIQ = nonverbal IQ composite score from the Leiter-R; Receptive vocabulary = standard score from the Peabody Picture Vocabulary-4 or its French version; SCQ = Social Communication Questionnaire total score. [a]: Significant difference between monolingual TYP and bilingual TYP, [b]: Significant difference between monolingual TYP and monolingual ASD, [c]: Significant difference between monolingual TYP and bilingual ASD, [d]: Significant difference between bilingual TYP and monolingual ASD, [e]: significant difference between bilingual TYP and bilingual ASD.

Table 2

Results from Executive Function Measures by Group

	Monolingual TYP *M* (*SD*)	Bilingual TYP *M* (*SD*)	Monolingual ASD *M* (*SD*)	Bilingual ASD *M* (*SD*)
BRIEF T-scores	(*n* = 10)	(*n* = 10)	(*n* = 8)	(*n* = 10)
Set-shifting [a]	45 (*5.60*)	49 (*9.07*)	69 (*11.41*)	67 (*11.50*)
Working memory [a]	51 (*10.23*)	46 (*9.78*)	59 (*12.52*)	63 (*10.96*)
DCCS task pre and post switch phases	(*n* = 10)	(*n* = 10)	(*n* = 10)	(*n* = 10)
Task accuracy				
Pre-switch (6 trials)	5.8 (*.42*)	5.9 (*.32*)	5.7 (*.48*)	6.0 (*.00*)
Post-switch (6 trials)	5.5 (*.85*)	5.7 (*.68*)	4.2 (*2.3*)	4.9 (*2.13*)
RT in milliseconds				
Pre-Switch	1326 (*284*)	1558 (*567*)	1431 (*278*)	1972 (*570*)
Post-Switch	1503 (*299*)	1451 (*303*)	1732 (*298*)	1918 (*267*)
Switch cost in milliseconds	(*n* = 8)	(*n* = 9)	(*n* = 7)	(*n* = 8)
	881 (*547*)	751 (*640*)	1156 (*1076*)	1368 (*854*)
DCCS task border phase	(*n* = 8)	(*n* = 9)	(*n* = 7)	(*n* = 8)
Task accuracy				
Border version [c d] (12 trials)	9.5 (*2.39*)	7.0 (*2.24*)	9.0 (*2.24*)	10.88 (*1.13*)
RT in milliseconds				
Border version [b]	2908 (*835*)	3786 (*1664*)	2233 (*502*)	3217 (*522*)
Working memory	(*n* = 10)	(*n* = 10)	(*n* = 10)	(*n* = 10)

Scaled score number repetition	12 (*2.35*)	11 (*2.11*)	11 (*3.06*)	10 (*3.31*)

Note. TYP = typically-developing children; ASD = autism spectrum disorders; BRIEF = Behavior Rating Inventory of Executive Functioning; DCCS = Dimensional Change Card Sort task. [a]: Main effect of diagnostic group, [b]: main effect of language, [c]: significant difference between monolingual TYP and bilingual TYP, [d]: significant difference between bilingual TYP and bilingual ASD.

Figures

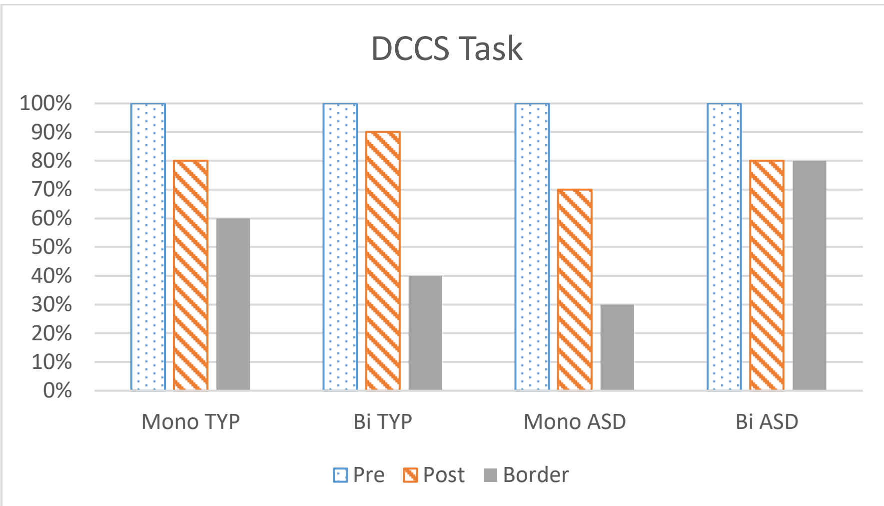

Figure 1. Percentage of participants passing each phase of the Dimensional Change Card Sort task by Group.

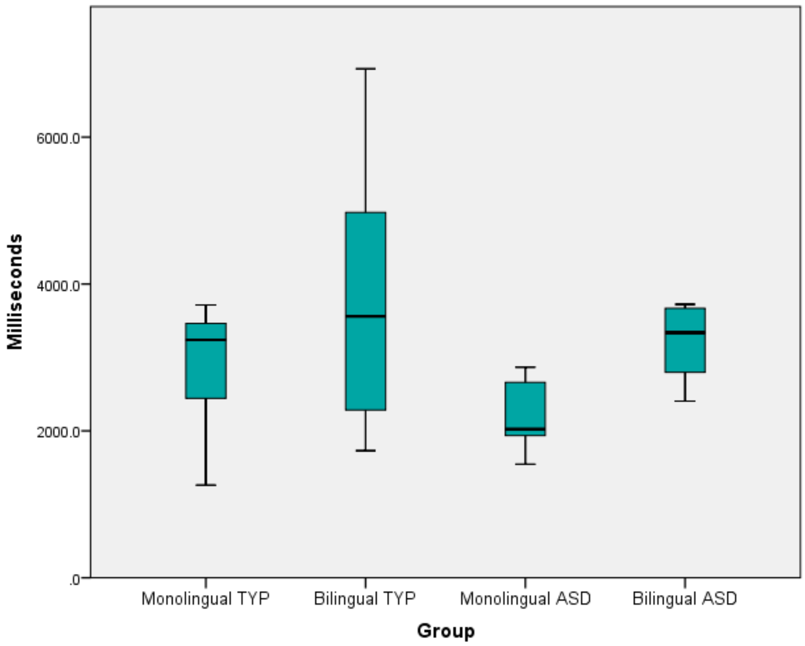

Figure 2. Mean reaction time for Border Version of the Dimensional Change Card Sort Task by Group.