Independent effects of bilingualism and socioeconomic status on language ability and executive functioning

Alejandra Calvo, Ellen Bialystok *

York University, Canada

ARTICLE INFO

Article history:
Received 6 March 2013
Revised 17 November 2013
Accepted 19 November 2013

Keywords:
Socioeconomic status
Bilingualism
Language ability
Executive functioning

ABSTRACT

One hundred and seventy-five children who were 6-years old were assigned to one of four groups that differed in socioeconomic status (SES; working class or middle class) and language background (monolingual or bilingual). The children completed tests of nonverbal intelligence, language tests assessing receptive vocabulary and attention based on picture naming, and two tests of executive functioning. All children performed equivalently on the basic intelligence tests, but performance on the language and executive functioning tasks was influenced by both SES and bilingualism. Middle-class children outperformed working-class children on all measures, and bilingual children obtained lower scores than monolingual children on language tests but higher scores than monolingual children on the executive functioning tasks. There were no interactions with either group factors or task factors. Thus, each of SES and bilingualism contribute significantly and independently to children's development irrespective of the child's level on the other factor.

1. Introduction

There is great interest in understanding the environmental conditions that affect cognitive ability and the mechanisms behind their influence as a complement to more biologically-based approaches to intelligence and performance. Two experiences that have been extensively investigated in this regard are bilingualism and socioeconomic status (SES), both of which have been shown to correlate with measures of cognitive performance and language ability throughout development. Robust effects of SES have been found across cognitive skills, including language, memory, and intelligence (Bradley & Corwyn, 2002; Hoff-Ginsberg, 1991; McCall, 1981) showing a relation between higher SES and better outcomes. In contrast, the effects of bilingualism on cognitive functioning vary in their direction, with positive outcomes for cognitive measures but negative outcomes for verbal measures (Akhtar & Menjivar, 2012). However, it is possible that these experiences interact and their effect depends on a specific level of the other. Thus, it may be that bilingualism only leads to cognitive advantages for certain levels of SES, such as middle-class children, or that SES only compromises ability for certain levels of language experience, such as monolingual children. Empirically studying this question is complicated by the fact that SES and bilingualism themselves are often correlated, making it difficult isolate the effect of each on performance. However, precisely because these two experiences frequently intersect it is particularly important to distinguish between the influence of each, both practically in terms of children's development and theoretically in terms of the possible mechanism underlying each. The purpose of the present study is to examine the effects of SES and bilingualism independently to determine the role each plays on cognitive and language outcomes, the extent to which their influence on development is similar or not, and whether their combined effects are interactive or independent.

The role of SES on intellectual functioning and academic performance is well established: children growing up in
families with more financial resources and parents who are more educated obtain higher scores on cognitive measures than do children without these advantages (Bradley & Corwyn, 2002; Brooks-Gunn & Duncan, 1997; McLoyd, 1998; Sirin, 2005). The difference in IQ scores between high and low SES groups is reported to be about one standard deviation (Bradley & Corwyn, 2002; Seifer, 2001).

Studies of SES have focused as well on specific cognitive systems, particularly language acquisition. Typically, children from low SES backgrounds have lower levels of both receptive and expressive language skills than more affluent children (Arriaga, Fenson, Cronan, & Pethick, 1998; Hart & Risley, 1995; Locke, Ginsborg & Peers, 2002; Qi, Kaiser, Milan, & Hancock, 2006). Using the Peabody Picture Vocabulary Test (PPVT), the difference in score for children in medium vs. low SES groups is 0.75 to 1 standard deviation (Noble, Norman, & Farah, 2005; Qi et al., 2006). These differences are more evident in complex language tasks, such as Clinical Evaluation of Language Fundamentals Assessment—Preschool (CELF-P; Semel, Wiig, & Secord, 2003). In a sample of low SES preschoolers, Locke, Ginsborg, and Peers (2002) reported that more than half the children met criteria for diagnosis of at least moderate language impairment by scoring 1 standard deviation or more below the population mean. These findings are consistent with research showing differences not only in the level but also the trajectory and rate of vocabulary growth as a function of SES (Arriaga et al., 1998; Dollaghan et al., 1999; Hart & Risley, 1995; rescorla & Alley, 2001).

SES has also been shown to influence the development of executive functioning (EF), a relation that Noble et al. (2005) argued could be central in explaining SES effects on IQ and achievement. Supporting this view, components of EF (planning, monitoring, switching) have been theoretically and empirically linked to general intelligence (Gray & Thompson, 2004; Kail, 2000; Kyllonen, 2002). For example, Carpenter, Just, and Shell (1990) derived a model for general intelligence through simulation studies heavily dependent on the working memory component of EF. In contrast, highly automatized tasks that are not part of overall assessments of intelligence do not require EF because they are driven by environmental cues and are ultimately run by ‘plans and programs’ already in long term memory (Miller & Cohen, 2001; Rabbitt, 1997; Shiffrin & Schneider, 1977). Thus, effects of SES are particularly apparent for tasks that engage the EF system.

Disparities in EF for children at different SES levels have been reported even in infancy. Lipina, Martelli, Vuelta, and Colombo (2005) evaluated the performance of 280 infants (6- to 14-months old) in the A not B task (Diamond, 1985) and found that infants from families with the lowest SES made more errors than their higher SES peers. Ardila, Rosselli, Matute, and Guajardo (2005) administered a battery of EF tests to 622 children ranging in age from 5- to 14-years old who differed in SES as measured by parents’ education. Higher SES children performed better overall, and parental education was a significant predictor of performance for the majority of EF measures in the battery. Finally, Noble and colleagues (Noble et al., 2005) showed that language and EF were the two neurocognitive systems most affected by SES. In two follow-up studies aimed at uncovering more precise relations (Farah et al., 2006; Noble, McCandliss, & Farah, 2007), they examined the relation between SES and working memory, cognitive control, and reward processing. Significant SES differences were found for working memory and cognitive control measures, both aspects of EF, with no difference in reward processing.

A parallel body of research has examined the effects of bilingualism on children’s language and EF abilities. Two aspects of language ability in children that have been extensively studied in bilinguals and monolinguals are vocabulary and metalinguistic skills. Bilingual children typically obtain lower scores than monolinguals on measures of both receptive (Bialystok, Luk, Peets, & Yang, 2010) and productive (Oller & Eilers, 2002) vocabulary. In the studies by Oller and colleagues of Spanish–English bilingual children, lower vocabulary scores were found in both English and Spanish assessments (Fernández, Pearson, Umbel, Oller, & Molinet-Molina, 1992; Oller, Pearson, & Cobo-Lewis, 2007), were independent of the level of usage of each language (Oller & Eilers, 2002), and persisted after accounting for SES (Cobo-Lewis, Pearson, Eilers, & Umbel, 2002). Using the PPVT with a heterogeneous sample of more than 1,700 bilingual children between the ages of 3- and 10-years old, Bialystok, Luk, et al. (2010) reported significantly higher scores for monolinguals at every age examined. Importantly, studies that match samples on SES replicate the vocabulary discrepancy in which monolinguals obtain higher English vocabulary scores than bilinguals (Hoff et al., 2012; Vagh, Pan, & Mancilla-Martinez, 2009).

The results for metalinguistic awareness are different: bilinguals typically show more advanced metalinguistic development than monolingual children in tasks examining the understanding of arbitrariness of linguistic labels (Bialystok, 1988; Cummins, 1978; Cummins & Mcalchay, 1978; Feldman & Shen, 1971; Ianco-Worrall, 1972) or requiring selective attention to information from form or meaning (Bialystok, 1986, 1988; Cromdal, 1999). These paradigms require EF to direct attention to the relevant feature (usually form) and ignore salient distracting information (usually meaning), implicating EF into language processing (Bialystok, 2001).

Evidence that bilingual children outperformed their monolingual peers on metalinguistic tasks that required EF led to the hypothesis that there might be a general EF advantage from bilingualism in nonverbal processing as well. Numerous studies have now supported this idea (see Akhtar & Menjivar, 2012 for review). Beginning again with infants, Kovács and Mehler (2009) compared 7-month-old infants being raised in homes that were monolingual or bilingual on an A-not-B type task in which they had to learn a new response to obtain the reward. Infants from bilingual homes were significantly more successful in learning the new response than were those exposed to only one language, suggesting that the basis for EF differences is established in the first few months of life. Research with preschool and early school-aged children has shown better performance by bilinguals on a Simon task (Martin-Rhee & Bialystok, 2008), flanker task (Yang, Yang, & Lust, 2011), Stroop task (Poulin-Dubois, Blaye, Coutya, & Bialystok, 2011), and the dimensional change card sort
(Bialystok, 1999; Bialystok & Martin, 2004). In all these cases, children must respond to a perceptual display that includes both target information indicating the correct response and a misleading cue leading to the incorrect response. For example, the flanker task requires that children indicate the direction that a central arrow is pointing, but on the more difficult incongruent trials, the flanking arrows are pointing in the opposite direction and must be ignored. Tasks that involve other aspects of EF have also been performed better by bilingual children than by monolinguals, such as multi-tasking (Bialystok, 2011), creativity (Adi-Japha, Berberich-Artzi, & Libnawi, 2010), and spatial perspective-taking (Greenberg, Bellana, & Bialystok, 2013). Carlson and Meltzoff (2008) administered a large battery of EF tasks and found advantages for Spanish–English bilingual children over monolingual children on conflict tasks, although there was no group difference in tasks that required withholding or delaying a response. A meta-analysis by Adesope, Lavin, Thompson, & Bialystok, 2012) in which bilingualism was isolated demonstrated these bilingual advantages in EF for low SES groups. Carlson and Meltzoff (2008) compared low SES Spanish–English bilingual children to middle-class monolingual children and found advantages in conflict tasks for the bilinguals once differences in vocabulary scores between the groups were statistically controlled. A study by Mezzacappa (2004) administered the ANT flanker task to a large group of 6-year-olds who differed in ethnicity and SES. There were significant advantages for children with high SES, but the low SES Hispanic children outperformed all other children on the incongruent trials of the flanker task, the condition requiring the greatest involvement of EF. Retrospectively, Mezzacappa noted that about two-thirds of the Hispanic children spoke Spanish at home, making them bilingual at least to some degree, so although bilingualism was not formally assessed, he attributed the Hispanic advantage in this task to bilingualism.

In both those studies, the language groups were different on other dimensions as well. In the study by Carlson and Meltzoff, the low SES bilingual children were living below the poverty line and facing challenges from their social circumstances that were not measured but undoubtedly differentiated them from the more affluent monolingual children to whom they were compared. In the study by Mezzacappa, ethnicity was also confounded with language group (see also Morton & Harper, 2007 regarding SES and language group). This problem was avoided in a recent study by Engel de Abreu, Cruz-Santos, Tourinho, Martin, and Bialystok (2012) in which bilingualism was isolated within a tightly controlled band of very low SES. They compared 8-year-old monolingual children in Portugal with comparable children whose parents had emigrated from that region to Luxembourg and were being raised as Portuguese-Luxembourgish bilinguals. Thus, all the children came from identical backgrounds in a poor region of Portugal. A large battery of measures assessing SES produced equivalent outcomes for families in the two language groups, and all children performed equivalently on background tasks measuring cognitive level, with the exception of vocabulary size where monolinguals outperformed bilinguals (cf., Bialystok, Luk et al., 2010). The relevant outcome measure was performance on a battery of tasks that either involved EF or not, a distinction confirmed by a factor analysis. The results showed that tasks that did not involve EF were performed similarly by children in both language groups, but tasks that included EF demands were performed significantly better by the bilingual children. The only difference between the two groups in addition to bilingualism was immigration, but recent studies have ruled out immigration as a contributor to these effects (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Bialystok & Viswanathan, 2009; Schweizer, Craik, & Bialystok, 2013).

To summarize, both SES and bilingualism have been shown to influence language and cognitive functioning in children, but because they have been studied in isolation, it is not known what the relation between them might be. There are three main possibilities. The first is that each of these experiences affects children’s development independently of the other. In this case, tasks that recruit EF will be solved better by bilingual children than monolingual children regardless of SES and by middle SES children than lower SES children regardless of bilingualism. Similarly, tasks based on language knowledge will be performed better by higher SES children and monolingual children than by their respective counterparts. Such independent and non-interacting outcomes is consistent with the view that each experience operates through a different mechanism. The second option is that these experiences interact and the effect of each is limited by specific conditions set by the other factor. Turkheimer, Haley, Waldron, D’Onofrio, and Gottesman (2003) argued that life experiences have a larger effect on cognitive performance than genetics for low SES children, but the reverse was true for higher SES children. If this is the case, then there would be larger effects of bilingualism in lower SES children than in higher SES comparison groups. However, the alternative prediction is that the more supportive middle-class environments would allow greater effects of bilingualism to emerge. For this second option, therefore, there are enabling conditions that must be met for experiential factors to impact children’s cognitive development, so claims about the effect of one of the experiences without reference to the status of the other experience would be misleading. Finally, it may be that the effects of these experiences are additive, so children with two positive experiences will experience relatively greater gains than do children with just one. Unlike the previous possibility in which there are conditions under which one of the experiences has no effect, in this case, both experiences modify cognitive performance but two sources of experience would lead to larger gains. Thus, there would be larger effects for bilingualism in higher SES children than in lower SES children on executive function tasks, and higher vocabulary scores in monolingual children who are also higher SES.

The present study used a factorial design to investigate monolingual and bilingual children from working class (WC) and middle class (MC) families on measures of
selective attention, inhibition, WM, and language ability. Unlike previous research, the study used a subtle distinction between WC and MC levels to avoid potential influence from factors associated with poverty that may exacerbate the influence of SES. This is the first study to factorially examine bilingualism and SES, so it has the potential to uncover the unique influence of each as well as the possible interaction between them.

2. Method

2.1. Participants

One hundred and ninety-three 6- to 7-year old children were recruited from 8 different public schools in Toronto. Exclusion criteria included being an immigrant with less than two years in an English-speaking country ($n = 2$) or not meeting criteria for either bilingual or monolingual status because children had been exposed to a non-English language but parents reported that the children’s proficiency in that language was limited ($n = 12$). Four children were excluded because they did not complete both testing sessions. The final sample consisted of 175 participants (86 boys and 89 girls). Background data are reported in Table 1.

2.2. Measures

Language and social background questionnaire (LSBQ). The LSBQ was completed by parents when they signed the permission forms. The questions included demographic information, including date of birth of the child, places of birth for both parents and the child, and date of arrival in the country if they were foreign born. It also contained questions about the languages that children spoke and the frequency with which each was used for both the child and the family in different contexts, the languages spoken by family members, and the child’s proficiency in those languages as well as information about how each of the languages was learned. Estimates of home language environment were quantified as the mean rating of 10 questions in which parents indicated the language used for specific activities (e.g., watching television, reading books) and interactions (e.g., parents speaking to children, siblings speaking to each other) on a 5-point scale in which 1 signified “always in English”, 3 signified “equal balance between English and other language” and 5 signified “always in non-English language”. These scores were reported in Table 1. Parents were also asked to provide their highest level of education and current occupation. The reported maternal and paternal occupational statuses were coded following the National Occupational Classification (2001) from Statistics Canada and transformed into an occupational earnings score based on Boyd’s Income Index in the Socioeconomic Scale of Canadian Occupations (Boyd, 2008). The scores represent the ranking of the average income (as a percentile) of each occupation title. In families with two working parents, the highest occupational status reported was used to represent the family’s occupational status. Education and income scores are also reported in Table 1.

Information from the LSBQ was used to classify children into four groups. SES was determined initially by maternal education and then confirmed by examining father’s education and occupation and income level. Parental education is the most commonly used indicator of SES in research with children and adolescents (Ensminger & Fothergill, 2003), is highly predictive of other variables such as income and occupation (Bornstein, Hahn, Suwalsky, & Haynes, 2003; Entwistle & Astone, 1994), and has been shown to have the most predictive power for cognitive performance compared to other SES indicators (Mistry, Biesanz, Chien, Howes, & Benner, 2008; Noble, Wolmetz, Ochs, Farah, & McCandliss, 2006; Scarr & Weinberg, 1978). Following Stevens, Lauinger, and Neville (2009), high school education was used as the criterion difference to distinguish between the two SES groups such that children whose mothers had completed at least some post-secondary education were classified as middle class (MC, $n = 111$) and those with high school or less education as working class (WC, $n = 64$). There was a significant correlation between maternal and paternal education, $r(169) = .55$, $p < .0001$, and the mean of these was calculated as an overall Education score. A 2-way ANOVA indicated a significant difference between SES groups, $F(1,168) = 187.27$, $p < .0001$, but no effect of language

Table 1
Mean score (and standard deviation) for demographic variables and standardized background measures by group.

<table>
<thead>
<tr>
<th></th>
<th>WC Monolingual</th>
<th>WC Bilingual</th>
<th>MC Monolingual</th>
<th>MC Bilingual</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>20</td>
<td>44</td>
<td>46</td>
<td>65</td>
</tr>
<tr>
<td>Mean age (in months)</td>
<td>80.0 (4.6)</td>
<td>81.6 (3.8)</td>
<td>80.4 (4.0)</td>
<td>79.9 (4.2)</td>
</tr>
<tr>
<td>Gender (% females)</td>
<td>55</td>
<td>57</td>
<td>46</td>
<td>49</td>
</tr>
<tr>
<td>Maternal education</td>
<td>1.9 (0.3)</td>
<td>1.7 (0.5)</td>
<td>3.5 (0.7)</td>
<td>3.7 (0.7)</td>
</tr>
<tr>
<td>Maternal education</td>
<td>1.9 (0.9)</td>
<td>2.4 (1.2)</td>
<td>3.2 (0.9)</td>
<td>3.6 (1.1)</td>
</tr>
<tr>
<td>Income index</td>
<td>44.1 (15.8)</td>
<td>44.0 (20.2)</td>
<td>68.2 (23.4)</td>
<td>65.6 (22.6)</td>
</tr>
<tr>
<td>Home language</td>
<td>1.2 (0.2)</td>
<td>2.6 (0.6)</td>
<td>1.2 (0.2)</td>
<td>2.6 (0.6)</td>
</tr>
<tr>
<td>K-BIT standard score</td>
<td>101.0 (13.6)</td>
<td>101.0 (16.6)</td>
<td>101.8 (13.1)</td>
<td>106.0 (14.7)</td>
</tr>
<tr>
<td>PPVT standard score</td>
<td>101.6 (8.9)</td>
<td>94.8 (11.6)</td>
<td>104.0 (7.1)</td>
<td>99.3 (11.7)</td>
</tr>
<tr>
<td>Pair cancelation WJ-III standard score</td>
<td>102.8 (7.8)</td>
<td>103.6 (6.7)</td>
<td>104.8 (8.1)</td>
<td>105.2 (8.3)</td>
</tr>
<tr>
<td>Animal cancelation WISC IV scale score</td>
<td>9.7 (2.2)</td>
<td>8.7 (3.1)</td>
<td>10.7 (2.2)</td>
<td>9.6 (3.4)</td>
</tr>
</tbody>
</table>

Maternal education range = 1 (less than high school) to 5 (graduate/professional degree).
Income index = percentile ranging from 1 to 99 based on Boyd’s (2008) socioeconomic scale.
Home language = 1 (entirely English), 3 (equally English and other language) , 5 (no English).
group and no interaction, $F_s < 1$. The mean Education score for parents whose children were classified as WC monolinguals was 1.9 (range 1–3), for WC bilinguals was 2.1 (range 1–3.5), for MC monolinguals was 3.4 (range 2–5), and for MC bilinguals was 3.6 (range 2–5). Education was also significantly correlated with occupation and income level, $r(162) = .52$, $p < .0001$. There was no correlation between home language scores and either Education ($r = .00$) or occupation ($r = .04$) further indicating no overlap in the primary classifications for SES and bilingualism.

All children were being educated in English and all children were proficient in English (all English PPVT scores were within $\pm 2$ SD of 100), so the primary determinant of bilingualism was the home language. Children were initially considered to be bilingual if parents reported that children’s comprehension and production in another language was good, very good or excellent and parents spoke to them in the other language most of the time. This classification designated 109 children as bilingual and 66 children as monolingual. This distribution in which 62% of the children are bilingual is consistent with the local population demographic where approximately 60% of families do not use English as the home language (Statistics Canada, 2007). A 2-way ANOVA for scores on the home language usage scale showed a significant difference between language groups, $F(1, 145) = 280.30$, $p < .0001$, with no effect of SES and no interaction, $F_s < 1$. Bilingual children spoke 26 different languages, with Chinese/Mandarin ($n = 18$), Tamil ($n = 11$) and Urdu ($n = 12$) as the only languages spoken by more than 10% of bilingual children.

Of the 8 schools, one contributed children only to the MC group, with the other 7 schools contributing children to both groups. Similarly, two schools contributed only monolingual children with the other 6 schools providing both monolingual and bilingual children. Thus, the school environment and home neighborhoods were very similar for all children in the study.

Matrices subtest of the Kaufmann brief intelligence test, 2nd edition (K-BIT 2; Kaufman & Kaufman, 2004). This test assessed non-verbal reasoning. The child selects the one item from five choices that is similar to or completes the visual analogy of the target stimulus. For this and all standardized measures, standard procedures for administration and scoring were followed.

Peabody picture vocabulary task, 3rd edition (PPVT-III). The PPVT is a standardized measure of receptive vocabulary in English (Dunn & Dunn, 1997). Words of increasing complexity were read and the child was required to pick the picture that best illustrated the meaning of the read word.

Nonverbal visual attention. The Pair Cancelation subtest of the Woodcock Johnson Tests of Cognitive Abilities-III (Woodcock, McGrew, & Mather, 2001) required children to find target stimuli among distracters. Children circled pairs of pictures where a ball was followed by a dog in a page containing rows of black and white pictures of dogs, balls, and cups.

Verbal visual attention. The Cancelation subtest of Wechsler Intelligence Scale for Children-Fourth Edition (WISC–IV; Wechsler, 2003) required children to find targets that were defined as pictures of animals within distracter stimuli that were pictures of common inanimate objects. The stimuli were presented in either a random arrangement or structured in columns and rows. Children needed to identify, and likely name the stimuli to carry out the classification.

Flanker task. This task was programmed in E-prime (Psychology Software Tools, 2002) and presented on a tablet computer. The goal was to indicate the direction in which a target red chevron was pointing (< or >). The target chevron was presented either in isolation or surrounded by four other black stimuli. The trial started with a central fixation cross for 250 ms and was followed by the stimulus display which stayed on the screen until a response was made or after 2000 ms had elapsed.

In the control condition, a single chevron was shown in the center of the screen. In the visual search condition, the red chevron was flanked by four black diamonds. The target chevron was randomly presented in one of 3 positions in the row, the center or one space to the right or left. In the conflict condition, the flankers were black chevrons facing in the same (congruent) or opposite (incongruent) direction as the target chevron that was in one of the 3 central positions.

Five blocks were administered in a fixed order: 1. control (24 trials), 2. visual search (24 trials), 3. conflict (48 trials, half congruent), 4. visual search (24 trials), and 5. control (24 trials), for a total of 144 trials. Within each block trials were presented in randomized order for both direction (pointing right or left) and position (in visual search and conflict blocks). Children responded by pressing the touch screen to show the side the red chevron was pointing. Reaction time and accuracy were recorded by the program.

Frog Matrices Task (FMT). The Frog Matrices Task (Morales, Calvo, & Bialystok, 2013) is a measure of children’s ability to maintain information in mind under different degrees of cognitive control, such as ignoring interference and manipulating the to-be-remembered information. Children were shown a $3 \times 3$ matrix on a touch-screen tablet computer and were told that each of the nine cells represented a pond in which a frog could be resting and that they needed to remember all the ponds the frog had visited. In all conditions, an increasing number of frogs (from 2 to 6) were presented by increasing the stimulus length by one pond every second trial, creating a total of 10 trials per condition. Each trial had three components: first, the to-be-remembered items were shown, then a blank screen was presented for 2000 ms, followed by a blank matrix which stayed on the screen until the child’s responses were recorded. The task was programmed in E-prime and children responded by touching the ponds on the screen.

There were four conditions. In the first, simple span (see Fig. 1), the frogs were presented all at once for 2000 ms, and following the blank screen, children could enter their responses in any order. In the second condition, masked, the blank screen was replaced by a distracter matrix showing 3 frogs in irrelevant locations. Children were told to remember the frogs shown on the first matrix and ignore the second one. In the third condition, sequential, the frogs were shown one at a time for one second each and children were instructed to remember the frogs in the order of their
appearance, testing memory for both the position of each frog and temporal order of its presentation. In the fourth condition, operational, frogs were also presented sequentially but instead of reporting the locations in the order in which they were presented, children were instructed to reorder the response. The cells in the matrix were connected by lines that created a path through the ponds. The path was visible in all conditions but was not mentioned until the operational condition, when it became relevant. Children were told that the frogs had left some treasures for them on the ponds, and to collect the treasures, they must follow the path by walking on it with their finger and pressing only on the ponds where a frog had been. The rules were that they could not go back and had to continue moving in the direction of the path. Therefore, children were required to reorganize the order in which the frogs had been presented but recall all the locations.

For each condition, span was calculated as the longest string length for which the child remembered all the frogs on at least one of the two trials. For example, if a child was correct on both items of length two (2 + 2) and was fully correct on one item of length three and incorrect on the other, and incorrect on all items of length four, then the span would be 3. Total scores were calculated by adding the number of frogs correctly recalled up to the child’s span. For the sequential and operational conditions, children received separate scores for position (i.e., whether the selected pond was in the trial) and order (i.e., whether the pond was placed in the correct response sequence). An overall total was calculated as the sum of the scores for position and order. Following this procedure, the maximum total for the sequential and operational conditions was twice as much (80) as that for the simple and masked conditions (40), so total scores were converted to proportions to compare performance across conditions.

Fig. 1. Illustration of the frog matrices task for each of the four conditions.

2.3. Procedure

Parental permission forms were distributed to all first grade students in the selected schools. Interested parents
returned the signed permission form and the LSSEQ to school with their child. Children were tested individually in two sessions, in a quiet room in the school. Tasks were presented in a fixed order. The first session consisted of the Woodcock Johnson Pair Cancelation, Frogs Matrices Task and Flanker Task; the second session, the WISC Cancelation, K-BIT and PPVT were administered. At the end of each session children were given stickers and thanked for their participation.

### 3. Results

Table 1 reports the mean scores and standard deviations for the demographic variables and background measures. Univariate outliers were identified (z-scores where \( p < .001 \)) and removed from the pair cancelation task and the FMT. In the pair cancelation task, two MC bilinguals had extremely high scores and for the FMT, two WC monolinguals had extremely low scores. There were no group differences in age for language group, \( F < 1 \), SES, \( F(1,174) = 1.14, \) n.s., or their interaction, \( F(1,174) = 1.74, \) n.s., and gender distribution was equivalent for both SES and language group, \( X^2 < 1 \).

**Background tasks.** A 2-way ANOVA for SES and language group was performed on scores from the K-BIT as a measure of abstract reasoning (see Table 1). There were no significant main effects of language, \( F < 1 \), SES, \( F(1,171) = 1.41, \) n.s., or their interaction, \( F < 1, \) confirming no difference between groups in this measure.

Vocabulary was assessed by the PPVT test of receptive vocabulary (Table 1). A 2-way ANOVA indicated significant main effects for both language group, \( F(1,170) = 11.30, \) \( p < .01 \), \( \eta^2 = .06, \) and SES, \( F(1,170) = 4.03, \) \( p < .05, \) \( \eta^2 = .02, \) and no interaction, \( F < 1 \). These effects showed that monolinguals (\( M = 103.26, \) \( SD = 7.68 \)) had higher vocabulary scores than bilinguals (\( M = 97.46, \) \( SD = 11.82 \)) and that children from MC families (\( M = 101.26, \) \( SD = 10.26 \)) had higher scores than those from WC families (\( M = 96.91, \) \( SD = 11.22 \)).

The pair cancelation task is a test of nonverbal attention and perceptual matching. A 2-way ANOVA showed no significant effects of SES, \( F(1,169) = 1.99, \) n.s., language group, \( F < 1, \) or their interaction, \( F < 1 \).

In the animal cancelation task, pictures need to be assigned to semantic categories to determine which ones to cross out. Thus, the task includes a linguistic component, which did not differ from each other. The operational condition (\( M = 0.32, \) \( SD = .20 \)) was the most difficult. The data are plotted in Fig. 2.

**Language and executive functioning composites.** To evaluate the overall effect of SES and bilingualism on language ability and executive functioning, a composite score was calculated as the average of the z-score of tasks reflecting each of these constructs. For language ability, the calculations included the z-score of the standard PPVT score and bilingual children, \( r = .27, \) \( p = .01 \). A follow-up ANCOVA on animal cancelation scores with PPVT scores as the covariate erased all effects and showed no significant difference for language group, \( F(1,166) = 1.16, \) n.s., SES, \( F(1,166) = 2.11, \) n.s., or their interaction, \( F < 1, \) once vocabulary level was considered. Thus, the adjusted scores indicate equivalent attention ability between the two groups, but the unadjusted standard scores reflect verbal ability.

**Chevron flanker task:** accuracy and reaction time were analyzed separately and are shown in Table 2. RTs longer than 1850 or shorter than 150 ms were excluded (2.2% of data). Seven children with accuracy levels lower than chance were excluded from the RT analyses. The possibility of speed-accuracy trade-off was evaluated by correlating accuracy with RT for each condition. The only significant correlation was for the conflict block, \( r = -.22, \) with better performance associated with faster responses, that is, the opposite of a speed accuracy tradeoff.

A 3-way mixed factor ANOVA on accuracy scores for language group, SES, and block condition showed a significant effect of condition, \( F(2,161) = 52.72, \) \( p < .0001, \) \( \eta^2 = .35 \). There was also a significant effect of language group, \( F(1,162) = 9.77, \) \( p < .002, \) \( \eta^2 = .08, \) and a marginally significant effect of SES, \( F(1,162) = 3.44, \) \( p = .06, \) \( \eta^2 = .02, \) and no interactions, \( Fs < 1 \). The main effects indicated that bilinguals (\( M = 0.87, \) \( SD = .12 \)) were more accurate than monolinguals (\( M = 0.83, \) \( SD = .10 \)) and that children from MC families (\( M = 0.86, \) \( SD = .12 \)) had marginally higher scores than those from WC families (\( M = 0.84, \) \( SD = .10 \)). For condition, pairwise contrasts showed that differences between all conditions were significant \( p < .05, \) with control trials (\( M = 0.89, \) \( SD = .09 \)) being easier than search trials (\( M = 0.86, \) \( SD = .09 \)), which in turn were easier than conflict trials (\( M = 0.80, \) \( SD = .12 \)).

The same analyses were applied to the RT data with the only significant effect being condition, \( F(2,161) = 520.03, \) \( p < .001, \) \( \eta^2 = .87 \). RT was faster for control trials (\( M = .866, \) \( SD = 117 \)) than neutral (\( M = 1030, \) \( SD = 136 \)) or conflict trials (\( M = 1126, \) \( SD = 159 \)). There were no other significant main or interaction effects.

**FMT working memory.** A 3-way mixed factor ANOVA on proportion scores indicated significant main effects for condition, \( F(3,165) = 166.03, \) \( p < .001, \) \( \eta^2 = .75, \) language group, \( F(1,167) = 6.24, \) \( p = .01, \) \( \eta^2 = .04, \) and SES, \( F(1,167) = 4.66, \) \( p < .05, \) \( \eta^2 = .03, \) with no interactions, \( Fs < 1 \). Bilinguals (\( M = 0.56, \) \( SD = .20 \)) were more accurate than monolinguals (\( M = 0.50, \) \( SD = .27 \)) and children from MC families (\( M = 0.55, \) \( SD = .20 \)) were more accurate than those from WC families (\( M = 0.50, \) \( SD = .27 \)). For condition, the simple condition was the easiest (\( M = 0.80, \) \( SD = .27 \)), with higher scores than either masked (\( M = 0.49, \) \( SD = .29 \)) or sequential conditions (\( M = 0.50, \) \( SD = .24 \)), which did not differ from each other. The operational condition (\( M = 0.32, \) \( SD = .20 \)) was the most difficult. The data are plotted in Fig. 2.
the standardized score of the animal cancelation task. For executive functions, the average included the \(z\)-score of the overall mean proportional score of the FMT, and the standardized score of overall accuracy on the flanker task. Overall mean scores were used for tasks with multiple components to ensure that all tasks had the same weighting on the composite score. This procedure was also justified by the absence of any interaction between condition and group for any of the tasks.

Table 2
Mean score (and standard deviation) for accuracy and reaction time on the three conditions of the chevron flanker task by group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>WC Monolingual</th>
<th>WC Bilingual</th>
<th>MC Monolingual</th>
<th>MC Bilingual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (proportion correct)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>0.85 (.10)</td>
<td>0.91 (.07)</td>
<td>0.89 (.08)</td>
<td>0.91 (.08)</td>
</tr>
<tr>
<td>Search</td>
<td>0.82 (.13)</td>
<td>0.88 (.10)</td>
<td>0.85 (.08)</td>
<td>0.90 (.07)</td>
</tr>
<tr>
<td>Conflict</td>
<td>0.76 (.15)</td>
<td>0.81 (.13)</td>
<td>0.81 (.09)</td>
<td>0.83 (.12)</td>
</tr>
<tr>
<td>Reaction time (in ms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>863 (139)</td>
<td>848 (90)</td>
<td>877 (133)</td>
<td>871 (113)</td>
</tr>
<tr>
<td>Search</td>
<td>1000 (165)</td>
<td>1023 (114)</td>
<td>1047 (163)</td>
<td>1032 (119)</td>
</tr>
<tr>
<td>Conflict</td>
<td>1105 (182)</td>
<td>1106 (123)</td>
<td>1140 (173)</td>
<td>1137 (163)</td>
</tr>
</tbody>
</table>

Similarly, for language scores, there was a significant effect of language group, \(F(1,171) = 14.10, p < .001, \eta^2 = .08\), a significant effect of SES, \(F(1,171) = 6.45, p < .01, \eta^2 = .04\), and no interaction, \(F < 1\). For language composites, Cohen’s \(d\) values were .45 for SES and .60 for bilingualism; for executive function composites, Cohen’s \(d\) values were .25 for SES and .51 for bilingualism. The direction of these effects, however, was different for the two composites. Higher SES and monolingualism were associated with better language outcomes, whereas higher SES and bilingualism were associated with better executive functioning outcomes. Effect sizes were also calculated for the interaction term for each composite score, revealing a very small effect size for language, \(f = .04\), and a small effect size for executive function, \(f = .13\). Correlations between the indicators of SES and composite scores calculated separately for each of the two SES groups showed no relation, all \(r\)s close to 0, indicating little variance within each SES group. Finally, the independence of each of the factors is demonstrated in the finding that WC monolingual children obtained language scores equivalent to MC bilingual children, \(F < 1\), and WC bilingual children obtained EF scores equivalent to MC monolingual children, \(F = 1.35, p = .25\), indicating no additional effect from the level of the other factor on performance.

4. Discussion

Previous research has reported that both SES and bilingualism influence children’s language and cognitive development, but the natural co-occurrence of these factors, especially in the United States, has made it difficult to determine if they have independent contributions to development or if their effect on development depends on specific values of the other factor. Such an outcome would diminish the general effect of each by confining its influence to certain circumstances. The present study is the first to factorially vary levels of each of these experiences to determine not only their individual contribution to development but also their potential interactions.

There are several unique strengths in the present design. First, because of the factorial structure, bilingual children were not predominantly Hispanic and poor and middle-class children were not predominantly monolingual and Caucasian (cf., Carlson & Meltzoff, 2008), making conclusions about each of the factors more reliable. Second, the criteria for distinguishing between the two levels of each factor were subtle: All the children attended the
same schools but children differed in level of parental education and affluence and the primary language spoken at home. Thus it would not be surprising to find no significant effect of either factor on children’s performance because of the substantial similarity among all groups. All children performed equivalently on a test of nonverbal reasoning and visual search. However, for the critical tests of language ability and executive functioning, each of SES and bilingualism significantly affected performance but the effect of each applied primarily to different domains and, importantly, the two never interacted.

Consider first the effect of SES. In all tasks for which group differences were found, that is, all tasks excluding nonverbal reasoning and visual search in which there was equivalent performance on basic cognitive skills for all children, the effect of SES was to reduce performance in the WC group relative to the MC group. Thus, WC children had smaller receptive English vocabularies, were (marginally) less accurate on the flanker task, and recalled fewer items across all conditions of the frog matrix working memory task than comparable MC children. These differences are consistent with research reporting SES differences in vocabulary and executive functioning (e.g., Noble et al., 2005) for more disparate levels of SES. In the present study, these effects emerged as main effects across language groups, indicating that the impact of SES was equivalent for both monolingual and bilingual children. Moreover, in tasks that included conditions that manipulated difficulty, specifically the flanker task and the FMT, the main effect of SES appeared across all the manipulations indicating a general depression in the ability tested for the WC children. Therefore, one way of interpreting the effect of SES is that it delays the progress children are making in mastering these skills since all children were equally able to perform the simple conditions but differentially able to succeed in the complex conditions. Thus, WC children are developing these skills in the same manner as MC children, that is, the difficult conditions are still more difficult for these children, but the overall rate of mastery is delayed. Because the WC children were not living in poverty as is usually the case in this type of research, the effect sizes for comparisons of SES were small, but they were significant and demonstrate the effect of even subtle SES differences on children’s language and cognitive development, irrespective of language background.

The effect of bilingualism was also consistent with previous research (Akhtar & Menjivar, 2012). Bilingual children obtained lower scores on a standard test of English vocabulary but made fewer errors on the flanker task and recalled more items on the FMT than did monolingual children irrespective of SES. There were no interactions with either SES or task condition, so again, the interpretation is that bilingualism slows vocabulary acquisition but accelerates executive function development across variations in SES and manipulations of task difficulty.

The extraction of test scores into composite scores for language ability and executive functioning clarified the effect of each of these factors on children’s development. Although both SES and bilingualism were main effects in children’s development, their influence was not the same. Specifically, SES was associated with a decrease in performance for both domains whereas bilingualism was associated with a decrease in language performance but an increase in executive functioning. This is the pattern previously presumed for bilingualism (e.g., Bialystok, 2009) but demonstrated directly in the present results. The important result is that the effect of each factor is independent of the other, so both MC and WC children profit equally from bilingualism, and both monolingual and bilingual children profit equally from higher SES.

Research investigating the effect of naturally-occurring life experiences on defined cognitive outcomes is fraught with the reality that the life situation of each individual is complex, that experiences are not wholly independent of each other, and that life does not provide the opportunity for random assignment to groups. Therefore, there has been considerable discussion regarding the reported research showing effects of SES (which may be potentially confounded with culture and ethnicity) and bilingualism (which may be potentially confounded with SES and culture) on children’s development. By using a factorial design, the present study provides the first evidence that each of these experiences contributes uniquely to development, irrespective of the other factor. The WC children were of widely mixed ethnicity and the bilingual children were of widely mixed language and cultural groups. All experience matters, and the present results demonstrate that they may act through different mechanisms and largely in isolation of each other to shape children’s cognitive ability. What is necessary now is to understand the effects of both SES and bilingualism more deeply so they can be used to improve the developmental outcomes for all children.

Acknowledgements

This work was supported by Grant R01HD052523 from the US National Institutes of Health to Ellen Bialystok. We are grateful to Kornelia Hawrylewicz and Nadeen Sara for their contribution to this study.

References


