

# Audio-visual object search is changed by bilingual experience

Sarah Chabal<sup>1</sup> · Scott R. Schroeder<sup>1</sup> · Viorica Marian<sup>1</sup>

© The Psychonomic Society, Inc. 2015

**Abstract** The current study examined the impact of language experience on the ability to efficiently search for objects in the face of distractions. Monolingual and bilingual participants completed an ecologically-valid, object-finding task that contained conflicting, consistent, or neutral auditory cues. Bilinguals were faster than monolinguals at locating the target item, and eye movements revealed that this speed advantage was driven by bilinguals' ability to overcome interference from visual distractors and focus their attention on the relevant object. Bilinguals fixated the target object more often than did their monolingual peers, who, in contrast, attended more to a distracting image. Moreover, bilinguals', but not monolinguals', object-finding ability was positively associated with their executive control ability. We conclude that bilinguals' executive control advantages extend to real-world visual processing and object finding within a multi-modal environment.

**Keywords** Attention and executive control · Eye movements and visual attention · Visual search

As we navigate the world, we receive information through multiple modalities, including inputs to both our auditory and visual systems. These multiple inputs compete for our attention, and we must selectively focus on the inputs that are most useful to the task at hand. Sometimes, two different sensory inputs provide complementary cues that are both beneficial to the task, and integrating across modalities can improve performance. For

example, imagine you are searching for your keys on a cluttered desk and you hear your keys clink together while opening a drawer. In this scenario, auditory and visual inputs are rapidly integrated to speed search (e.g., Chen & Spence, 2010; Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008; Molholm, Ritter, Javitt, & Foxe, 2004). Often, however, two sensory modalities provide conflicting cues, of which only one is useful. For example, as you search for your keys you may hear papers shuffling on your desk or your dog barking in the background. In this case only visual input (the shape of your keys) provides relevant information, and incompatible cross-modal cues become detrimental (Tellinghuisen & Nowak, 2003).

Because conflicting sensory inputs can negatively impact performance (Tellinghuisen & Nowak, 2003), efficient search requires that misleading auditory information be ignored – a task relying on executive control (Baddeley & Larsen, 2003; Elliott, 2002). As executive control is already needed to manage information from competing visual inputs (i.e., ignoring all distracting items in favor of the target object; Anderson, Vogel, & Awh, 2013; Bleckley, Durso, Crutchfield, Engle, & Khanna, 2003; Poole & Kane, 2009), conducting a visual search within an auditory context places increased demands on the cognitive control system. Given the high executive demands of multi-modal search, strong executive control abilities may be necessary for efficient target identification.

Executive control is a malleable skill that can be improved through experience and practice (e.g., Bialystok, 2006; Green, Sugarman, Medford, Klobusicky, & Bavelier, 2012; Tang et al., 2007). For example, people who speak more than one language develop enhanced executive control relative to their monolingual peers. Because both of a bilingual's languages are simultaneously activated when processing both auditory (e.g., Marian & Spivey, 2003a, 2003b; Spivey & Marian,

---

✉ Sarah Chabal  
schabal@u.northwestern.edu

<sup>1</sup> Department of Communication Sciences and Disorders,  
Northwestern University, 2240 North Campus Drive,  
Evanston, IL 60208-3570, USA

1999) and visual (Chabal & Marian, 2015) inputs, bilinguals must suppress information from the unneeded language and attend only to relevant linguistic information. This practice results in enhanced executive function abilities (e.g., Bialystok, 2006, 2008; Costa, Hernández, & Sebastián-Gallés, 2008; Martin-Rhee & Bialystok, 2008; Prior & MacWhinney, 2009). Bilinguals often outperform their monolingual peers on tasks involving conflict monitoring (e.g., Abutalebi et al., 2012), conflict resolution (Bialystok, 2010), and attentional control (e.g., Martin-Rhee & Bialystok, 2008), and these advantages are observed in auditory (e.g., Moreno, Bialystok, Wodniecka, & Alain, 2010; Soveri, Laine, Hämäläinen, & Hugdahl, 2011), visual (e.g., Bialystok, 2008; Morales, Yudes, Gómez-Ariza, & Bajo, 2015), and audio-visual (e.g., Bialystok, Craik, & Ruocco, 2006; Krizman, Marian, Shook, Skoe, & Kraus, 2012) domains. For example, bilinguals have been shown to outperform monolinguals on the Simon task, a non-linguistic measure of executive control skill (e.g., Bialystok et al., 2005; Bialystok, Craik, Klein, & Viswanathan, 2004). However, the scope of bilingual advantages in executive control has been debated, with some recent studies failing to find differences between monolingual and bilingual groups (e.g., Hilchey & Klein, 2011; Paap & Greenberg, 2013). Indeed, it has been argued that any potential bilingual advantages are confined to very specific task circumstances that are limited in scope (Paap, Johnson, & Sawi, 2015). Therefore, there is a need for further studies extending bilingual executive control research from artificial, laboratory tasks to more ecologically-valid circumstances. Here, we examine bilinguals' executive control performance in a real-world-like, multimodal visual search task.

To examine bilinguals' real-world performance within a multisensory environment, we designed a visual search task that contained multiple and varying auditory contexts. Monolinguals and bilinguals were asked to quickly locate an object while contending with the types of auditory-visual relationships that must be managed in the real world. In a natural environment, some visual search may occur in silence, with no auditory information to aid or hinder performance. However, it is more likely that auditory and visual inputs are simultaneously present. In such cases, sounds may correspond directly with relevant visual information (e.g., jingling keys), they may cue attention to visual items you would like to ignore (e.g., shuffling papers while you search for your keys), or they may signal objects that are not even within your visual field (e.g., a distant siren). The inclusion of all four of these audio-visual contexts ensured that cognitive control was being assessed in the most ecologically-valid settings.

In a previous study exploring visual search ability, bilinguals displayed faster performance than monolinguals under difficult search conditions (Friesen, Latman, Calvo, & Bialystok, 2014). Specifically, when asked to quickly locate a simple colored shape (e.g., turquoise circle) from an array

containing minimally different distractors (e.g., turquoise squares and pink circles), bilingual participants indicated the target's presence faster. We expected that this speed advantage would be observed even when the search contained ecologically-valid stimuli and a more real-world, audio-visual search context.

In order to uncover the locus of the expected bilingual advantage in visual search speed, participants' eye movements were tracked while they completed the multimodal search task. Eye movements can be used to index attentional processing (Hoffman & Subramaniam, 1995), allowing us to draw conclusions about the mechanisms underlying behavioral search performance. For example, group differences in eye-movement patterns (even in the absence of behavioral reaction time differences) may suggest that bilinguals and monolinguals employ distinct search strategies. Because bilinguals excel at focusing on task-relevant information (e.g., Coderre & Van Heuven, 2014), we expected eye-movements to reflect bilinguals' ability to ignore non-target objects.

Finally, because we hypothesize that enhanced cognitive control is the determining factor in bilinguals' expected search advantage, we explored whether multi-modal search performance is directly related to general executive control ability. We compared participants' performance on the search task to their performance on a non-linguistic test of executive control (Simon task). We hypothesized that, because visual search within a multi-modal environment relies directly upon executive control mechanisms, search performance would be correlated with performance on the Simon task. Specifically, we expected that participants with better executive control would complete the search task more quickly and efficiently.

In sum, the present study had three aims. Our first aim was to determine whether bilinguals' advantage in cognitive control extends to real-world, multi-modal settings. Second, we sought to uncover potential mechanisms behind any observed advantage in visual search performance by using fine-grained eye-tracking techniques. Finally, we hoped to provide preliminary evidence for a link between visual search and cognitive control by relating performance on the multi-modal search task to a non-linguistic executive control task.

## Method

### Participants

Thirty-seven participants – 21 bilinguals and 17 monolinguals – were included in the analyses. Participants ranged in age from 18 to 34 (mean age=21.5 years). Language group was determined by responses on the *Language Experience and Proficiency Questionnaire (LEAP-Q)* (Marian, Blumenfeld, & Kaushanskaya, 2007). Bilinguals reported speaking English and at least one other language (second languages

included Cantonese, Chinese, French, German, Hebrew, Italian, Japanese, Korean, Mandarin, Marathi, Polish, and Spanish). To be included in the study, bilinguals were required to have learned both languages by the age of 7 years and to have a self-rated proficiency in both languages of at least 7 on a 0 (none) to 10 (perfect) scale. Five bilinguals reported that English was their first language (L1) and 16 reported that English was their second language (L2) based on age of acquisition. On average, bilinguals reported being slightly more proficient in English ( $M=9.04, SD=0.86$ ) than in their other language ( $M=8.04, SD=1.13; t<0.05$ ). Monolinguals were required to have rated their proficiency in a language other than English as 3 or below on the 0–10 scale.

Table 1 presents demographic data for bilingual and monolingual participants. Language groups did not differ in non-verbal intelligence (performance subtests of the *Wechsler Abbreviated Scale of Intelligence*; Wechsler, 1999), phonological working memory (*Comprehensive Test of Phonological Processing*; Wagner, Torgesen, & Rashotte, 1999), or English receptive vocabulary (*Peabody Picture Vocabulary Test*; Dunn, 1981). Bilinguals and monolinguals also did not differ on a measure of executive function derived from a visual Simon task (Simon Effect, defined as response time on incongruent trials minus response time on congruent trials; Weiss, Gerfen, & Mitchel, 2010). Bilinguals and monolinguals differed in age (bilinguals were slightly younger than their monolingual counterparts); this difference was controlled for in all statistical analyses.

**Materials**

**Object search task** In the object search task, participants were presented with visual displays composed of eight objects. On each trial, one item served as a target object and the other seven items served as distractor objects. Target and distractor images were selected from the same set of 20 objects. Each object served as a target four times (once in each sound condition), creating 80 trials. On each trial, seven objects were randomly selected to serve as distractors.

The objects were represented by colored photographs with a maximum length of 4 cm along the largest dimension. The eight photographs in the display were placed along an iso-acuity ellipse with a plus sign (+) in the center. Viewing distance of the displays was a constant 60 cm; this distance was ensured by having participants place their chins on a chin rest set at a fixed location. While viewing the displays, participants’ eye movement fixations were recorded with an EyeLink 1000 (Version 1.5.2, SR Research Ltd.) eye tracker at a collection rate of 1000 Hz. During the search for the target object in the display, participants heard environmental sounds. The set of environmental sounds represented the same set of objects that were visually depicted. Sounds were, on average, 862 ms in duration (standard deviation=451 ms), and were played through headphones. All sounds and images were identical to those used by Iordanescu and colleagues (Iordanescu et al., 2010; Iordanescu, Grabowecky, & Suzuki, 2011; Iordanescu et al., 2008).

**Design and procedure**

**Object search task** On each of the 80 trials, participants first saw a plus sign (+) in the center of the screen for 1,500 ms. Next, the spoken label of the target object was played. Subsequently, the eight-image visual display was shown and remained on the screen until participants clicked on an image to indicate their selection of the target. At the onset of the visual display, participants saw the target object under one of four auditory conditions: (1) with an environmental sound that was consistent with the *target* object (*Target Consistent* condition; e.g., an audio recording of a barking dog while searching for a dog), (2) with an environmental sound that was consistent with a diagonally-located *distractor* object (*Distractor Consistent* condition; e.g., a flushing toilet while searching for a dog), (3) with an environmental sound consistent with an object not located within the visual display (*Unrelated Sound* condition; e.g., the sound of a train), or (4) in the absence of any environmental sound (*No Sound* condition). Each of the four conditions was represented by

**Table 1** Cognitive and linguistic participant demographics

Measure	Monolinguals	Bilinguals	Comparison
N	17 (6 males)	21 (9 males)	-
Age, years	23.41 (5.08)	20.10 (2.90)	$F(1,36)=6.41, p<0.05$
Performance IQ ( <i>Wechsler Abbreviated Scale of Intelligence</i> ; Wechsler, 1999)	116.00 (9.86)	115.57 (6.07)	$F(1,36)=0.03, n.s.$
Phonological Working Memory ( <i>Comprehensive Test of Phonological Processing</i> ; Wagner et al., 1999)	118.71 (9.69)	113.57 (14.34)	$F(1,36)=1.59, n.s.$
English Vocabulary Standard Score ( <i>Peabody Picture Vocabulary Test</i> ; Dunn, 1981)	114.88 (6.54)	114.81 (11.44)	$F(1,36)=0.00, n.s.$
Simon Effect (incongruent – congruent; Weiss, Gerfen, & Mitchel, 2010)	47.40 (24.75)	42.89 (29.98)	$F(1,36)=0.25, n.s.$

Note. Values represent means, with standard deviations shown in parentheses.

20 trials (conditions were adapted from Iordanescu, Guzman, Grabowecky, & Suzuki, 2008; Iordanescu, Grabowecky, Franconeri, Theeuwes, & Suzuki, 2010; and Iordanescu, Grabowecky, & Suzuki, 2011). Participants were instructed to click on the target object as quickly as possible. See Fig. 1 for a sample trial layout. The experiment was conceptualized as a  $2 \times 4$  mixed design with sound condition as a within-subjects variable and language group (monolingual, bilingual) as a between-subjects variable.

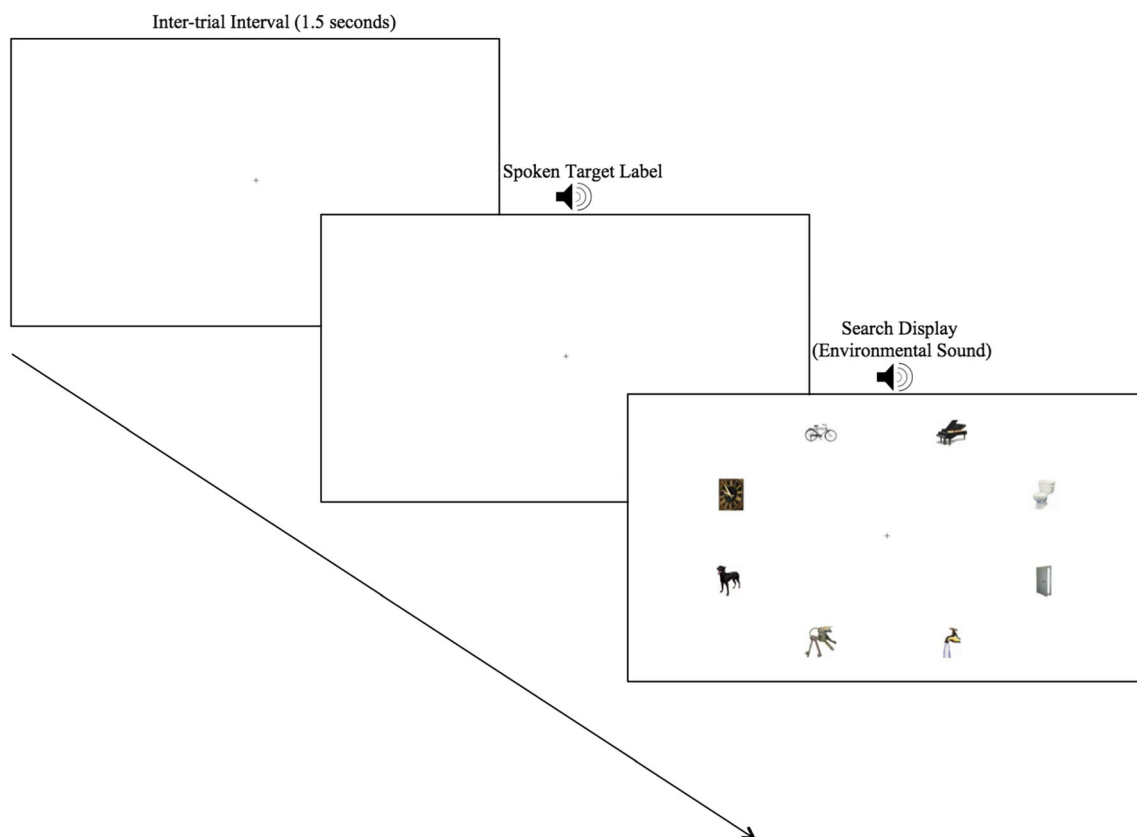
**Simon task** In each of 126 trials, participants first saw a plus sign (+) in the center of the screen for 350 ms, then a blank screen for 150 ms, followed by a blue or brown rectangle for 1,500 ms, and finally a blank screen for 850 ms. Participants pressed a button on the left side of the keyboard when seeing a blue rectangle and a button on the right side when seeing a brown rectangle. The rectangle appeared on the left, middle, or right side of the screen. Trials in which the rectangle appeared on the same side as the response button were called congruent trials, whereas trials in which the

rectangle appeared on the opposite side of the response button were called incongruent trials. Additionally, trials where the rectangle was placed in the middle were called neutral trials. Simon Effect scores were computed by subtracting reaction times on congruent trials from reaction times on incongruent trials.

### Data analysis

Prior to reaction time and eye-fixation analyses of the object search data, incorrect trials (0.71 %) and trials with response times that spanned more than two standard deviations above or below the mean (4.34 %) were excluded.

For the analysis of the eye-tracking data, interest areas were calculated in an  $80 \times 80$  pixel square around the center of each image and only fixations falling within these regions were considered. Within each trial, we calculated the total number of fixations made to each object, beginning at the onset of the search display and terminating with the participant's mouse-click response, and computed empirical logit transformations (Barr, 2008; Jaeger, 2008). These values were subjected to



**Fig. 1** Sample visual display in the object search task. Participants saw a plus sign (+) in the center of the screen for 1,500 ms, followed by the spoken label of the target object. Subsequently, the eight-image visual display was shown and remained on the screen until participants made a response. At the onset of the visual display, participants received one of four auditory conditions: (1) an environmental sound consistent with the

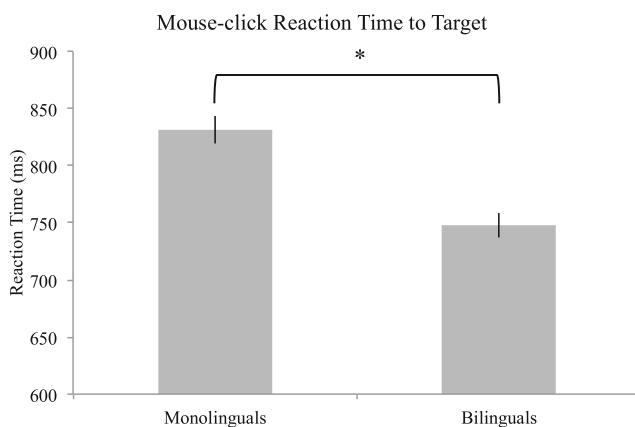
target object (*Target Consistent* condition), (2) an environmental sound consistent with a *distractor* object (*Distractor Consistent* condition), (3) an environmental sound consistent with an object not located within the visual display (*Unrelated Sound* condition), or (4) the absence of any environmental sound (*No Sound* condition)

mixed effects ANOVAs and ANCOVAs (to control for group differences in age). The critical comparison was the number of looks made to the target object versus the number of looks made to the distractor item located diagonally to the target (because competitors in the Distractor Consistent condition were always located in the diagonal position, the diagonal filler was selected to ensure that analyses were conducted identically across conditions).

## Results

### Visual search performance

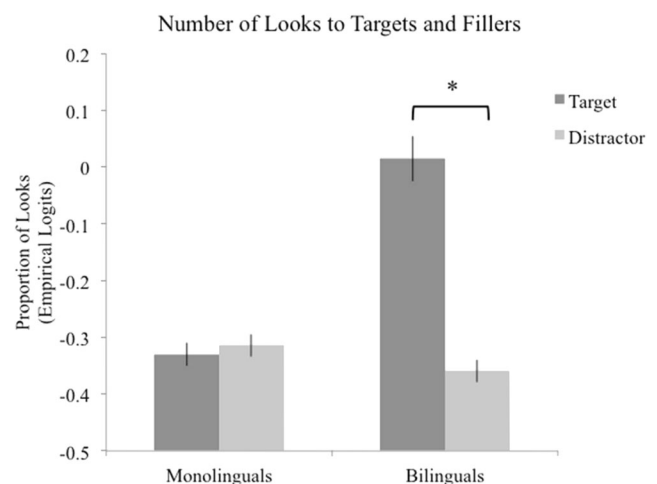
All participants were able to proficiently complete the visual search task, with performance accuracy at ceiling levels ( $M=99.10\%$ ,  $SD=2.46\%$ ). To determine whether monolinguals and bilinguals differed in the speed with which they located the target object, we conducted a  $2 \times 4$  mixed ANOVA with language group (monolingual, bilingual; between-subjects) and sound condition (target consistent, distractor consistent, unrelated, no sound; within-subject) as independent variables and mouse-click response time as the dependent variable. A significant main effect of language group ( $F(1,144)=25.83$ ,  $p<0.05$ ) revealed that bilinguals located the target more quickly than did monolinguals (Fig. 2). No main effect of sound condition or interaction between language group and sound condition emerged (all  $p>0.05$ ). Because age differed between groups (see Table 1) and may impact response times (e.g., Der & Deary, 2006; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994), we conducted a follow-up analysis with age as a covariate in our model. Results remained consistent, with a significant main effect of language group ( $F(1,143)=8.28$ ,  $p<0.05$ ), but no main effect of sound condition or interaction between language group and sound condition (all  $p>0.05$ ).



**Fig. 2** The average click reaction time to the target for monolinguals and bilinguals. Error bars represent the standard error of the mean, and asterisks represent statistical significance at  $p<0.05$

### Fixations to visual objects

To determine whether group differences in response times were due to differences in how the groups directed their attention during the visual search, we compared the number of fixations (logit-transformed) made to targets and diagonally-situated distractors with a  $2 \times 2 \times 4$  (item type, language group, sound condition) mixed ANOVA. Results revealed main effects of language group ( $F(1,288)=14.88$ ,  $p<0.05$ ) and item type ( $F(1,288)=49.98$ ,  $p<0.05$ ), and a language group by item type interaction ( $F(1,288)=49.13$ ,  $p<0.05$ ). No main effect or interaction with sound condition was present (all  $p>0.05$ ). These results remained significant when controlling for age within the ANCOVA model (main effect of language group:  $F(1,287)=11.78$ ,  $p<0.05$ ; main effect of item type:  $F(1,287)=49.82$ ,  $p<0.05$ ; language group  $\times$  item type interaction:  $F(1,287)=48.97$ ,  $p<0.05$ ). Follow-ups to the significant interaction revealed that bilinguals ( $F(1,166)=76.05$ ,  $p<0.05$ ), but not monolinguals ( $F(1,134)=0.42$ , *n.s.*), made more fixations to the target item than to the non-target distractor (Fig. 3). Furthermore, follow-up analyses directly comparing bilinguals to monolinguals revealed that bilinguals made more fixations to the target item ( $F(1,144)=45.18$ ,  $p<0.05$ ; including age covariate:  $F(1,143)=40.89$ ,  $p<0.05$ ) and fewer fixations to the distractor item ( $F(1,144)=7.16$ ,  $p<0.05$ ; including age covariate:  $F(1,143)=8.81$ ,  $p<0.05$ ) relative to monolinguals. Because a  $2 \times 4$  (language group, sound condition) mixed ANOVA confirmed that monolinguals and bilinguals did not differ in the total overall number of looks made within each trial (main effect of group:  $F(1,143)=2.87$ , *n.s.*), the higher proportion of target looks observed in bilinguals cannot be attributed to an overall greater number of eye fixations.



**Fig. 3** The number of looks made to Target and Distractor items. Values are plotted in empirical logits, where a more negative number represents fewer looks. Error bars represent the standard error of the mean, and asterisks represent statistical significance at  $p<0.05$

In order to ensure that our results were not due to an unintentional location bias resulting from selecting the diagonally-located filler object, we created a composite distractor score by averaging looks to all filler objects within the display. Regardless of the filler object used for analysis, results remained consistent. A  $2 \times 2 \times 4$  ANOVA (item type, language group, sound condition) revealed a main effect of language group ( $F(1,288)=25.99, p<0.05$ ), a main effect of item type ( $F(1,288)=46.83, p<0.05$ ), and a language group by item type interaction ( $F(1,288)=44.41, p<0.05$ ). Results were consistent in the ANCOVA model controlling for age (main effect of language group:  $F(1,287)=20.94, p<0.05$ ; main effect of item type:  $F(1,287)=46.69, p<0.05$ ; language group  $\times$  item type interaction:  $F(1,287)=44.27, p<0.05$ ).

### Correlations with executive control

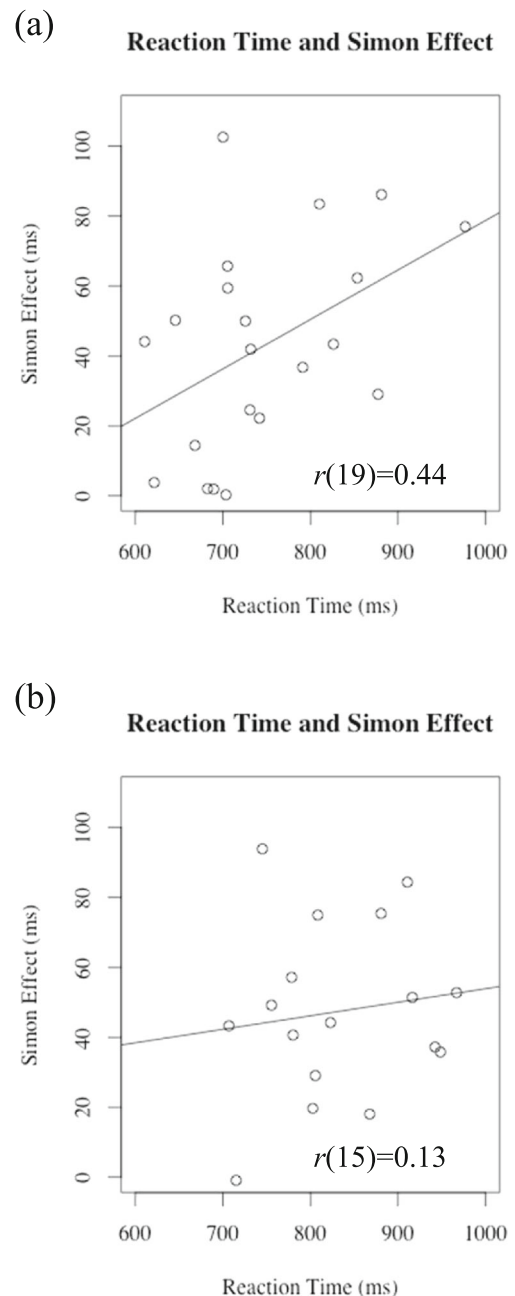
Eye-tracking analyses suggest that bilinguals selectively attended to the relevant target image while inhibiting the distracting items. To examine whether domain-general executive control mechanisms underlie this ability, we explored correlations between performance on our search task and a separate executive control task (the Simon task).

Overall, Simon Effect scores were positively correlated with response times. Participants who were better at inhibiting location information during the Simon task were faster at locating the target object during visual search ( $r(36)=0.33, p<0.05$ ). Within-group analyses, however, revealed that the overall correlations between executive control ability and visual search performance were driven entirely by the bilingual group. Whereas bilinguals' visual search reaction time was significantly correlated with their Simon Effect ( $r(19)=0.44, p<0.05$ ) score, monolinguals' performance on the two tasks was not related ( $r(15)=0.13, n.s$ ). See Fig. 4 for correlation plots.

### Discussion

When completing a multi-modal search task in which auditory input provided inconsistent search cues, bilinguals were better able to focus their attention on the relevant visual object. Specifically, bilinguals fixated the target object more often than did their monolingual peers (who, in contrast, attended more to a distracting image). As a result, bilinguals were able to locate the target visual item faster. Moreover, bilinguals' (but not monolinguals') object-finding ability was positively associated with their executive control ability, as bilinguals who displayed better non-linguistic control also excelled at managing interference during the search task.

Bilinguals' increased speed when searching for a visual object is consistent with recent research in which bilinguals were found to be faster than monolinguals when locating



**Fig. 4** Correlations between search reaction time and performance on a non-linguistic Simon task for bilinguals (a) and monolinguals (b)

simple shapes (Friesen et al., 2014). Our current study extends this object-finding advantage to real-world visual items (e.g., keys and dogs) and to a multi-modal search environment. Across all search contexts (silent, unrelated sound, distractor-consistent sound, and target-consistent sound), bilinguals outperformed monolinguals.

Notably, our study demonstrates that bilinguals' overall speed advantage is not attributed to a faster motor response system. Had monolinguals and bilinguals only differed in their ability to respond after locating the target, we would not have expected the two groups' eye movement patterns to differ.

Instead, we found significant differences in how monolinguals and bilinguals looked at the visual objects within the display. Whereas bilinguals made more looks to the target item, monolinguals looked more often at the distracting item. Therefore, the locus of bilinguals' object-finding advantage seems to be their superior ability to focus on relevant information in the face of distraction.

There are a few mechanisms of executive control that may drive bilinguals' enhanced search abilities. According to some frameworks of the executive control of attention (e.g., Ghatan, Hsieh, Petersson, Stone-Elander, & Ingvar, 1998; Neill, Valdes, & Terry, 1995), relevant information may be up-regulated (facilitation) or irrelevant information may be down-regulated (inhibition). In the current research, the observed bilingual advantage in object-finding may be attributed to either facilitation, inhibition, or to enhancements in both. Whereas effective facilitation would allow for increased focus on the target object, effective inhibition would allow for suppression of the non-target objects and environmental sounds. The idea that inhibition and facilitation may be driving bilinguals' performance advantage is supported by a significant correlation with Simon task performance, a task that also involves inhibition and facilitation. This suggests that the same executive control processes that underlie bilinguals' control of non-linguistic competition similarly affect how attention is delegated and distracting information is managed within a visual scene. In contrast to the bilinguals, there were no significant associations between monolinguals' search performance and their performance on the non-linguistic control task. While this dissociation between how monolinguals and bilinguals recruit domain-general executive control has been demonstrated in linguistic tasks (e.g., Blumenfeld & Marian, 2011; Blumenfeld, Schroeder, Bobb, Freeman, & Marian, 2015; Krizman et al., 2012; Marian, Chabal, Bartolotti, Bradley, & Hernandez, 2014; Yoshida, Tran, Benitez, & Kuwabara, 2011), this study represents the first time that executive control has been shown to be related to bilinguals', but not monolinguals', object search.

One reason that bilinguals' search performance may have been more closely associated with executive control performance than was monolinguals' is that, as in language-based tasks (Marian & Spivey, 2003a, 2003b; Midgley, Holcomb, van Heuven, & Grainger, 2008; Thierry & Wu, 2007), bilinguals may experience greater cognitive demands in non-verbal tasks as well. For example, both visual objects (Chabal & Marian, 2015) and characteristic sounds (Schroeder & Marian, *in preparation*) have been shown to activate associated linguistic information even though they are non-linguistic in nature. With linguistic information activated, bilinguals would necessarily need to contend with greater inhibitory demands (from bilinguals' two languages vs. monolinguals' one), thereby requiring the increased recruitment of domain-general control mechanisms.

Nevertheless, it is likely that monolinguals also recruited executive control to successfully complete the current task. During a visual search task, executive functions such as monitoring, updating, planning, inhibiting, and attending all contribute to efficient selection of the target object. However, the Simon task does not index all aspects of executive control, as it primarily targets inhibition of irrelevant cues and facilitation of informative cues (Zorzi & Umiltá, 1995). Future research should use a more comprehensive battery of executive function tasks in order to identify and understand the cognitive control mechanisms that are required for successful visual search, particularly within a multi-modal environment.

The current finding that bilinguals excel at multi-modal search in a real-world environment adds to the growing debate surrounding bilingual cognitive advantages. In particular, recent discussions have emerged challenging the prevalence of the bilingual advantage in cognitive control (de Bruin, Treccani, & Della Sala, 2014; Paap, 2014; but see Bialystok, Kroll, Green, MacWhinney, & Craik, 2015). While the bilingual advantage has been well documented (e.g., Bialystok, 2006, 2008; Costa et al., 2008; Martin-Rhee & Bialystok, 2008; Prior & MacWhinney, 2009), there are also instances in which bilingual benefits were not observed (e.g., Hilchey & Klein, 2011; Paap & Greenberg, 2013). For example, in the current study, monolinguals and bilinguals did not differ in Simon task performance (see also Kousaie & Phillips, 2012; Morton & Harper, 2007), which could be taken as evidence that the two groups had similar cognitive control abilities. Nevertheless, more sensitive measurement techniques (e.g., eye-tracking) in our search task were able to capture bilingual enhancements in visual processing. We therefore believe that our findings contribute to the bilingual debate not only by providing further support for the bilingual advantage but also by showing how fine-grained measures that capture cognitive processes as they unfold (e.g., eye-tracking) may be more effective at detecting bilingual-monolingual differences than measures that assess only the end state of multi-step cognitive processes (e.g., reaction time).

We also show, for the first time, that bilinguals' cognitive advantages extend to ecologically-valid, naturalistic tasks. For example, the present study involves the common process of searching for an object. It also depicts real-world, colored objects, and involves multiple sensory modalities. This is in contrast to the artificial environments imposed by commonly-used executive control tasks such as the Stroop, flanker, and Simon tasks.

One surprising finding to emerge from our current study is that neither monolinguals nor bilinguals were affected by the presence of an auditory cue. Past research suggested that English speakers' search was faster when the visual display was accompanied by a sound that was consistent with the target (e.g., "meow" while searching for a cat) compared to when the sound was misleading (Iordanescu et al., 2010,

2008). However, we did not find any significant differences in search time across any of our sound conditions. The most likely explanation for this stems from differences in design between our study and those of Iordanescu and colleagues. Unlike in Iordanescu et al.'s work, in which the sound provided a meaningful cue to the target one-third of the time and was distracting one-third of the time, the auditory input in our current study was only helpful on 25 % of the trials and, more importantly, was detrimental on 50 % of the trials. Therefore, participants in our study may have been more likely to ignore the unhelpful auditory information, choosing to focus instead on the visual modality that provided the most consistently meaningful cue. This interpretation is supported by research on non-linguistic interference tasks, showing that as information from a single dimension becomes less and less informative, people become more likely to focus their attention on the more meaningful dimension (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009).

Although the target-consistent sound did not facilitate visual search in the current paradigm, it is possible that audio-visual cues interact to influence visual processing under at least some circumstances. For example, past research suggests that audio-visual effects are heavily influenced by the timing of stimulus presentation. Chen and Spence (2011) found that environmental sounds aid visual processing only when the auditory stimuli precede the onset of the visual stimulus, and not when the audio-visual stimuli are presented concurrently. Future research should therefore manipulate the timing of audio-visual presentations. For instance, in the current paradigm, presenting the auditory cue in advance of the visual stimulus (instead of the simultaneous presentation used here) may potentially show a stronger effect of the auditory cue.

In closing, our results reveal differences in how bilinguals and monolinguals use cognitive control to perform an audio-visual object search. Bilinguals' enhanced executive control ability facilitated search performance, thereby extending bilinguals' advantage to a real-world task that involves object-finding in a multi-modal environment. These results suggest that multi-modal search, a task involved in common experiences such as finding your keys or navigating a cluttered computer desktop, can be improved through experience. We show that multilingual language practice (a form of cognitive training) can improve and optimize performance across sensory modalities, illustrating the interconnectivity and malleability of the human cognitive system.

**Author note** This research was funded by grant NICHD R01 HD059858-01A to Viorica Marian. The authors would like to thank the members of the Northwestern Bilingualism and Psycholinguistics Research Group for helpful comments on this work.

## References

- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernández, M., Scifo, P., Keim, R., ... Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, 22(9), 2076–86. doi:10.1093/cercor/bhr287
- Anderson, D. E., Vogel, E. K., & Awh, E. (2013). A common discrete resource for visual working memory and visual search. *Psychological Science*, 24(6), 929–938. doi:10.1177/0956797612464380
- Baddeley, A., & Larsen, J. D. (2003). The disruption of STM : A response to our commentors. *Quarterly Journal of Experimental Psychology*, 56(8), 37–41. doi:10.1080/02724980343000530
- Barr, D. J. (2008). Analyzing “visual world” eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474. doi:10.1016/j.jml.2007.09.002
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology*, 60(1), 68–79. doi:10.1037/cjep.2006008
- Bialystok, E. (2008). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, 12(01), 3. doi:10.1017/S1366728908003477
- Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: beyond inhibition. *Developmental Psychology*, 46(1), 93–105. doi:10.1037/a0015466
- Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, 24(1), 40–9. doi:10.1016/j.neuroimage.2004.09.044
- Bialystok, E., Craik, F. I. M., Klein, R. M., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, 19(2), 290–303. doi:10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and ageing. *Quarterly Journal of Experimental Psychology*, 59(11), 1968–83. doi:10.1080/17470210500482955
- Bialystok, E., Kroll, J. F., Green, D. W., Macwhinney, B., & Craik, F. I. M. (2015). Publication bias and the validity of evidence: What's the connection? *Psychological Science*, 26(6), 1–8. doi:10.1177/0956797615573759
- Bleckley, M. K., Durso, F. T., Crutchfield, J. M., Engle, R. W., & Khanna, M. M. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic Bulletin & Review*, 10(4), 884–889. doi:10.3758/BF03196548
- Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, 118(2), 245–257. doi:10.1016/j.cognition.2010.10.012
- Blumenfeld, H. K., Schroeder, S. R., Bobb, S. C., Freeman, M. R., & Marian, V. (2015). Auditory word recognition across the lifespan: Links between linguistic and nonlinguistic inhibitory control in bilinguals and monolinguals. *Linguistic Approaches to Bilingualism*. (in press)
- Chabal, S., & Marian, V. (2015). Speakers of different languages process the visual world differently. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000075
- Chen, Y. C., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, 114(3), 389–404. doi:10.1016/j.cognition.2009.10.012
- Chen, Y. C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1554–1568. doi:10.1037/a0024329



- Coderre, E. L., & Van Heuven, W. J. B. (2014). Electrophysiological explorations of the bilingual advantage: Evidence from a Stroop task. *PLoS ONE*, *9*(7). doi:10.1371/journal.pone.0103424
- Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–49. doi:10.1016/j.cognition.2009.08.001
- Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*(1), 59–86. doi:10.1016/j.cognition.2006.12.013
- De Bruin, A., Treccani, B., & Della Sala, S. (2014). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, *26*(1), 99–107. doi:10.1177/0956797614557866
- Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*(1), 62–73. doi:10.1037/0882-7974.21.1.62
- Dunn, L. (1981). *The Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.
- Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition*, *30*(3), 478–487. doi:10.3758/BF03194948
- Fozard, J. L., Verduyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age differences and changes in reaction time: The Baltimore longitudinal study of aging. *Journal of Gerontology*, *49*(4), 179–189.
- Friesen, D. C., Latman, V., Calvo, A., & Bialystok, E. (2014). Attention during visual search: The benefit of bilingualism. *International Journal of Bilingualism*. doi:10.1177/1367006914534331
- Ghatan, P. H., Hsieh, J. C., Petersson, K. M., Stone-Elander, S., & Ingvar, M. (1998). Coexistence of attention-based facilitation and inhibition in the human cortex. *NeuroImage*, *7*(1), 23–29. doi:10.1006/nimg.1997.0307
- Green, C. S., Sugarman, M. A., Medford, K., Klobusicky, E., & Bavelier, D. (2012). The effect of action video game experience on task-switching. *Computers in Human Behavior*, *28*(3), 984–994. doi:10.1016/j.chb.2011.12.020
- Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*(4), 625–58. doi:10.3758/s13423-011-0116-7
- Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795. doi:10.3758/BF03206794
- Iordanescu, L., Grabowecy, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics*, *72*(7), 1736–41. doi:10.3758/APP.72.7.1736
- Iordanescu, L., Grabowecy, M., & Suzuki, S. (2011). Object-based auditory facilitation of visual search for pictures and words with frequent and rare targets. *Acta Psychologica*, *137*(2), 252–9. doi:10.1016/j.actpsy.2010.07.017
- Iordanescu, L., Guzman-Martinez, E., Grabowecy, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*(3), 548–554. doi:10.3758/PBR.15.3.548
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. doi:10.1016/j.jml.2007.11.007
- Kousaie, S., & Phillips, N. A. (2012). Ageing and bilingualism: Absence of a “bilingual advantage” in Stroop interference in a nonimmigrant sample. *The Quarterly Journal of Experimental Psychology*, *65*(2), 356–369. doi:10.1080/17470218.2011.604788
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(20), 7877–81. doi:10.1073/pnas.1201575109
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940. doi:10.1044/1092-4388(2007)067
- Marian, V., Chabal, S., Bartolotti, J., Bradley, K., & Hernandez, A. E. (2014). Differential recruitment of executive control regions during phonological competition in monolinguals and bilinguals. *Brain and Language*, *139*(717), 108–117. doi:10.1016/j.bandl.2014.10.005
- Marian, V., & Spivey, M. J. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*(02), 173–193. doi:10.1017/S0142716403000092
- Marian, V., & Spivey, M. J. (2003b). Competing activation in bilingual language processing: Within- and between-language. *Bilingualism: Language and Cognition*, *6*(2), 97–115. doi:10.1017/S1366728903001068
- Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *11*(01), 81–93. doi:10.1017/S1366728907003227
- Midgley, K. J., Holcomb, P. J., van Heuven, W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research*, *1246*, 123–35. doi:10.1016/j.brainres.2008.09.078
- Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study. *Cerebral Cortex*, *14*, 452–465. doi:10.1093/cercor/bhh007
- Morales, J., Yudes, C., Gómez-Ariza, C. J., & Bajo, M. T. (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia*, *66*, 157–169. doi:10.1016/j.neuropsychologia.2014.11.014
- Moreno, S., Bialystok, E., Wodniecka, Z., & Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *Journal of Neurolinguistics*, *23*, 564–579. doi:10.1016/j.jneuroling.2010.05.002
- 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
- Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and Inhibition in Cognition* (pp. 207–261). San Diego, CA: Academic Press.
- Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology*, *26*(3), 242–255.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–58. doi:10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. a., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 1–14. doi:10.1016/j.cortex.2015.04.014
- Poole, B. J., & Kane, M. J. (2009). Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, *62*(7), 1430–1454. doi:10.1080/17470210802479329
- Prior, A., & MacWhinney, B. (2009). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, *13*(02), 253. doi:10.1017/S1366728909990526
- Schroeder, S. R., & Marian, V. (in preparation). Sound and meaning activation from verbal and nonverbal sounds.

- Soveri, A., Laine, M., Hämäläinen, H., & Hugdahl, K. (2011). Bilingual advantage in attentional control: Evidence from the forced-attention dichotic listening paradigm. *Bilingualism: Language and Cognition*, *14*(2011), 371–378. doi:[10.1017/S1366728910000118](https://doi.org/10.1017/S1366728910000118)
- Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281–284.
- Tang, Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., ... Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 17152–17156. doi:[10.1073/pnas.0707678104](https://doi.org/10.1073/pnas.0707678104)
- Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65*(5), 817–828. doi:[10.3758/BF03194817](https://doi.org/10.3758/BF03194817)
- Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12530–5. doi:[10.1073/pnas.0609927104](https://doi.org/10.1073/pnas.0609927104)
- Wagner, R., Torgesen, J., & Rashotte, C. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.
- Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2010). Colliding cues in word segmentation: The role of cue strength and general cognitive processes. *Language and Cognitive Processes*, *25*(3), 402–422. doi:[10.1080/01690960903212254](https://doi.org/10.1080/01690960903212254)
- Yoshida, H., Tran, D. N., Benitez, V., & Kuwabara, M. (2011). Inhibition and adjective learning in bilingual and monolingual children. *Frontiers in Psychology*, *2*. doi:[10.3389/fpsyg.2011.00210](https://doi.org/10.3389/fpsyg.2011.00210)
- Zorzi, M., & Umiltà, C. (1995). A computational model of the Simon effect. *Psychological Research*, *58*(3), 193–205. doi:[10.1007/BF00419634](https://doi.org/10.1007/BF00419634)