A bilingual disadvantage in metacognitive processing

Tomas Folke [a,*,1], Julia Ouzia [b,*,1], Peter Bright [b], Benedetto De Martino [a], Roberto Filippi [b,c]

[a] University of Cambridge, Department of Psychology, Downing Street, Cambridge CB2 3EB, United Kingdom
[b] Anglia Ruskin University, Department of Psychology, East Road, Cambridge CB1 1PT, United Kingdom
[c] Developmental Neurocognition Lab, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom

A B S T R A C T

Recent research indicating that bilingualism is associated with enhanced executive function suggests that this enhancement may operate within a broader spectrum of cognitive abilities than previously thought (e.g., Stocco & Prat, 2014). In this study, we focus on metacognition or the ability to evaluate one's own cognitive performance (Flavell, 1979). Over the course of two experiments, we presented young healthy adult monolinguals and bilinguals with a perceptual two-alternative-forced-choice task followed by confidence judgements. Results from both experiments indicated that bilingual participants showed a disadvantage in metacognitive efficiency, determined through the calculation of Mratio (Maniscalco & Lau, 2014). Our findings provide novel insight into the potential differences in bilingual and monolingual cognition, which may indicate a bilingual disadvantage. Results are discussed with reference to the balance of advantages versus disadvantages associated with multilanguage learning.

© 2016 Published by Elsevier B.V.

1. Introduction

1.1. Bilingual cognition

Previous research has found bilingual children and adults to outperform their monolingual peers on tasks requiring the inhibition of irrelevant information (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004; Martin-Rhee & Bialystok, 2008), shifting from one set of information to another (e.g., Bialystok, 1999; Bialystok & Martin, 2004), as well as updating information in working memory on tasks with high processing demands (Bialystok et al., 2004; Carlson & Meltzoff, 2008). The bilingual advantage in executive function has been associated with a range of bilingual experiences, cross-culturally, and across the lifespan (Bialystok & Viswanathan, 2009; Bialystok et al., 2004) consistent with claims that both languages of bilingual individuals are concurrently active at all times, even in unilingual contexts (Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra & Van Heuven, 1998, 2002; Van Hell & Dijkstra, 2002). Therefore, for a bilingual speaker, active suppression of

the non-target language may be required (Green, 1986, 1998; though note the alternative explanation of semantic facilitation discussed in Costa, 2005; Costa & Caramazza, 1999; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007) as well as shifting mechanisms (Green & Abutalebi, 2013). As a result of this continuing inhibitory demand, bilingualism is thought to 'train the brain' and enhance executive function beyond the domain of language (Abutalebi & Green, 2007; Hernandez, Bates, & Avila, 1996; Mechelli et al., 2004; Stocco, Yamasaki, Natalenko, & Prat, 2014).

Recently, the 'bilingual advantage hypothesis' has been challenged by reports of no significant group differences or methodological issues (Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2014; Paap, Sawi, Dalibar, Darrow, & Johnson, 2014) including bias towards the publication of confirmatory findings (de Bruin, Treccani, & Della Sala, 2015). According to the work of Paap and colleagues, as well as others (e.g., Duñabeitia et al., 2014; Gathercole et al., 2014; Morton & Harper, 2007), there are no empirical grounds to believe that bilingualism is associated with enhanced executive function.

Broader approaches to bilingual cognition have provided important insights. For example, recent research suggests that bilingualism is associated with advantages in monitoring visual conflict (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Costa, Hernández, & Sebastián-Gallés, 2008), speed of rule-based learning (Stocco & Prat, 2014), Theory of Mind (Rubio-Fernández & Glucksberg, 2012), exercising perceptual-level rather than response-level inhibition (Blumenfeld & Marian, 2014), adjusting

---

* Corresponding authors at: University of Cambridge, Department of Psychology, Brain Decision Modeling Laboratory, Downing Street, Cambridge CB2 3EB, United Kingdom (T. Folke). Anglia Ruskin University, Department of Psychology, Faculty of Science & Technology, East Road, Cambridge CB1 1PT, United Kingdom (J. Ouzia).
E-mail addresses: tomas.folke@gmail.com (T. Folke), jouzia@gmail.com (J. Ouzia).
[1] The first two authors share first authorship. Placement on the manuscript was determined by alphabetical order.

proactive and reactive control (Morales, Gómez-Ariza, & Bajo, 2013), and controlling verbal interference during speech comprehension (Filippi, Leech, Thomas, Green, & Dick, 2012; Filippi et al., 2015). Overall, it appears that research successfully demonstrates a bilingual advantage using paradigms that require the use of multiple components of executive functioning, rather than inhibitory control in particular.

One area of interest that has not, to date, received attention in the literature is whether metacognitive processes are affected through the development of additional linguistic skills (i.e., second or multiple language learning). Past bilingualism research has addressed metalinguistic awareness as well as metacognitive reading strategies in children (García, Jiménez, & Pearson, 1998). However, to our knowledge there have been no attempts reported in the literature to evaluate general metacognitive abilities in bilingual individuals.

1.2. Metacognition

Metacognition is the ability to evaluate one's own cognitive processes, or, more informally, to have 'thoughts about thoughts' (Fernandez-Duque, Baird, & Posner, 2000; Flavell, 1979; Fleming, Ryu, Golfinos, & Blackmon, 2014). On a theoretical level, this is often modelled as a two-level system, with an object level, first order process, and a meta level, second order process (Nelson & Narens, 1994). An important aspect of metacognition is the ability to get a subjective sense of one's cognitive performance (Grimaldi, Lau, & Basso, 2015; Peirce & Jastrow, 1885). For example, when we identify a familiar-looking face on a crowded street we might feel more or less certain that we did see an old friend (or just someone that looked like them). In this case, the face categorisation would be the first order process and our sense of confidence in the categorisation would reflect a second order process, evaluating the fidelity of the first order process. In many cases, subjective confidence judgements are thought to result from imperfect readouts of the uncertainty associated with the first-order decisions (Meyniel, Sigman, & Mainen, 2015).

In experimental psychology, metacognitive performance is often assessed by comparing confidence judgements in relation to an objective measure of task performance, such as error rate (e.g., De Martino, Fleming, Garrett, & Dolan, 2013; Schwartz & Díaz, 2014; Yeung & Summerfield, 2014). When evaluating metacognitive performance three terms are of central importance: accuracy, bias, and efficiency (Maniscalco & Lau, 2012, 2014). Metacognitive accuracy is the extent to which confidence can be used to discriminate between correct trials and error trials (Galvin, Podd, Drga, & Whitmore, 2003). For example, if a participant is shown a set of pictures and has to evaluate whether they have seen them before, good metacognitive accuracy would result in their confidence judgements being consistently higher when they are correct, compared to when they are wrong. Metacognitive accuracy appears to be domain-general in healthy people, in the sense that people have similar metacognitive accuracy across tasks that require different first order abilities (McCurdy et al., 2013; Song et al., 2011; Veenman, Elshout, & Meijer, 1997). However, note that dissociations have been found between metacognition relating to memory and metacognition relating to visual discrimination in patients with brain lesions (Fleming et al., 2014).

In order to gain a complete picture of metacognitive performance one must also account for metacognitive bias. Metacognitive bias refers to the tendency to generally report high- or low confidence, regardless of the quality of the available information, or the accuracy of the first order judgement. For example, people tend to be overconfident in certain memory tasks (i.e., overestimating how often they are correct), whilst still being able to discriminate between correct and incorrect performance (for a review see Hoffrage, 2004). Metacognitive efficiency is a signal theoretic concept that refers to how good a person's metacognitive accuracy is given their first order accuracy. Intuitively, this is straightforward: imagine two people, Susan and John, performing a memory test. Susan produces fewer errors and therefore has better first order accuracy than John. Nevertheless, both participants report high confidence for 80% of the correctly remembered items and report high confidence for 40% of the items when they were wrong. This means that they both demonstrated the same level of metacognitive accuracy, because their confidence judgements were equally good at discriminating between correct and incorrect trials. However, in a sense John is metacognitively superior to Susan, because even though his first order decision process is worse, he still shows equally accurate confidence judgements. In our experiment we controlled for first order performance to get a pure measure of metacognitive efficiency in two ways. First, we used an adaptive staircase procedure to ensure a similar first-order accuracy for the experimental task across all participants. Second, we controlled for differences in first order performance mathematically.

Historically, metacognitive accuracy was computed by correlating confidence with first order performance within each participant (Kornell, Son, & Terrace, 2007; Nelson, 1984). However, this approach has been criticised for its inability to distinguish metacognitive accuracy from metacognitive bias (Masson & Rotello, 2009). This problem has recently been addressed by Maniscalco and Lau (2012, 2014), who applied signal detection theory (SDT) to metacognition, thus providing separate measures for bias and sensitivity. Below follows a non-technical introduction to the SDT framework in relation to first and second order performance, to help the interested reader appreciate how metacognitive efficiency is quantified (for a more in-depth, technical treatment, see Maniscalco & Lau, 2014).

One of the easiest ways to measure first-order performance in a two-alternative discrimination task is simply to compare the proportion of hits to the proportion of false alarms. Hits and false alarms are two of four possible outcomes that can occur within this context: (1) a hit is correctly indicating when a target is present; (2) a miss is failing to indicate when a target is present; (3) a false alarm is indicating that a target is present when it is not; and (4) a correct rejection is indicating that a target is absent when it is. We can calculate the hit rate for the full experiment by dividing the number of hits by the total number of trials when the target was present, and the false alarm rate by dividing the number of false alarms by the total number of trials when the target was absent. (Note that hit rate + miss rate = 1 and false alarm rate + correct rejection rate = 1, so the other two measures are superfluous). The higher the hit rate relative to the false alarm rate, the better the participant's first order performance. This can be visualised by plotting hit rates on the y-axis and false alarm rates on the x-axis.

Now say that we want to determine the participant's discriminatory ability independent of their response bias (i.e., their tendency to prefer one response over the other). One way to do this would be to change the relative rewards offered for hits versus correct rejections and plot different hit rates and false alarm rates for these different incentive structures. Such a plot is called a Receiver Operating Characteristic (ROC) curve. The strength of the SDT framework is that, from a single hit ratio-false alarm ratio pairing, it can estimate ROC curves that closely match ROC curves estimated from multiple pairings (Green & Swets, 1966). Therefore, SDT allows us to separate response bias from discriminatory ability without having to vary the incentive structures of the responses.

SDT assumes that each response is the result of two factors, the strength of evidence on that trial and the response criterion. In the example below, evidence can be ranked from "target is definitely absent" to "target is definitely present" (see the x-axis in Fig. 1). For each trial of a given difficulty, the strength of evidence is drawn
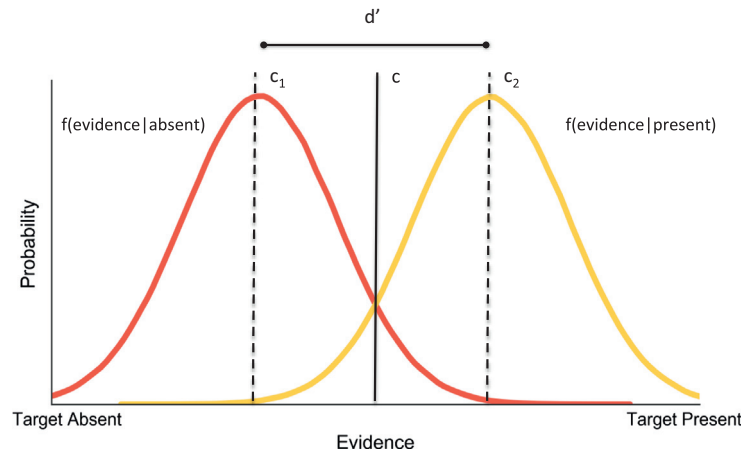
Fig. 1. Illustration of signal detection theory (adapted from Maniscalco & Lau, 2014). According to signal detection theory, the response of a given trial is determined by two factors: the strength of evidence (x-axis) and the response criterion (c). The strength of evidence is drawn from one of two Gaussian distributions: the left distribution if the target is absent, and the right distribution if the target is present. Participants respond "present" if the evidence value drawn is to the right of the response criterion and "absent" if it is to the left. A participant's discriminative ability is quantified as d', the difference between the means of the distributions relative to their joint standard deviation. Confidence can be added by introducing additional criteria that distinguish between high and low confidence (c_1 for when participants respond "absent," and c_2 for when participants respond "present").

from one of two Gaussian distributions, based on whether the target is present or not. The intersection of the two distributions represents the indecision point, where the evidence is equally strong for the target's presence as for its absence. The evidence on a given trial is evaluated in relation to decision criterion, c. If the evidence is greater than c, participants respond that the target is present, and if the evidence is less than c, participants respond that the target is absent. c is unbiased if it rests on the indecision point (as in the graph), but can be at any position, depending on how strongly participants value hits relative to correct rejections. A participant's overall discriminative ability is quantified as d', the difference in means of the two probability distributions divided by their joint standard deviation.

SDT can be extended to explain confidence judgements by adding decision criteria for different levels of confidence. In Fig. 1, any evidence drawn to the left of c but to the right of c_1 would lead to a low-confidence response that the target is absent, whereas any evidence drawn to the left of c_1 would be a high-confidence response that the target is absent. Conversely, any evidence drawn to the right of c but to the left of c_2 would lead to a low-confidence response that the target is present, and any evidence drawn to the right of c_2 would lead to a high-confidence response that the target is present. The confidence judgements also allow us to draw an ROC curve from a single incentive structure. We can mimic a liberal response criterion by only treating confident absence judgments as absence responses, and everything else as present responses. Similarly, we can mimic a conservative response criterion by only treating confident "present" judgements as "present" responses and everything else as absent responses.

However, while d' provides a good estimate of an empirical ROC curve for first-order decisions, there is a poor fit between the theoretical ROC curve from d' and the empirical ROC curve from confidence judgments. This implies that information is lost between the first-order (present-absent) and second-order (confidence) discrimination. Meta-d' quantifies this information loss by estimating a d' that fits the ROC curve drawn for the confidence judgments. In other words, meta-d' estimates the first-order accuracy the participants would have had if there were no information loss between first- and second-order judgments. The Mratio provides an easy readout of metacognitive efficiency because it is the fraction between meta-d' and d', so that a participant with confidence responses that perfectly reflect their first order accuracy would have an Mratio of 1 (Fleming & Lau, 2014). This SDT approach to

metacognitive performance has been demonstrated to outperform alternatives and to give robust measures of metacognitive accuracy and metacognitive efficiency (Barrett, Dienes, & Seth, 2013).

Fernandez-Duque et al. (2000) argue that, on a conceptual level, there are similarities between metacognitive processes and executive function. Empirical findings, supporting the notion that there may be a connection between metacognition and executive function, demonstrate that inhibitory control correlates with the ability to accurately apply rules in decision making tasks (Del Missier, Mäntylä, & Bruine de Bruin, 2010; Souchay & Isingrini, 2004). It has also been suggested that attentional shifting skills supports the ability to provide consistent judgements of performance (Del Missier et al., 2010). Nevertheless, unless a firmer evidence base for such links is established, interdependency between metacognitive ability and components of executive function will remain speculative.

1.3. The current study

The aim of this study was to explore metacognitive abilities in bilingual individuals. Given considerable uncertainty in the literature regarding i. the relationship between metacognitive abilities and executive function and ii. the balance of cognitive advantage versus disadvantage associated with bilingualism, our key prediction was necessarily tentative: that bilinguals would demonstrate a metacognitive advantage. We compared metacognitive efficiency (as expressed by Mratio) between bilinguals and monolinguals in relation to a perceptual two-alternative-forced choice task. Because metacognitive performance tends to be associated with task performance (Galvin et al., 2003; Maniscalco & Lau, 2012), we used a perceptual task that allowed us to adjust task difficulty online for each participant, titrating performance at around 71% for all participants. This standardisation ensured severely restricted variation in task performance across participants, implying that any variation in metacognitive performance could not be accounted for by differences in task performance.

2. Methods

2.1. Participants

We tested sixty-two healthy young adults, thirty-one English monolinguals (M_{age} = 22.3, SD = 3.7, range 18.3–34.4; 12 males),

and thirty-one bilinguals from a range of linguistic backgrounds (M_{age} = 25.3, SD = 4.5, range 19.6–38.3; 13 males). Whilst all participants were considered to be 'young adults' and recruited with corresponding age restrictions, the bilingual group was found to be significantly older than the monolingual group, $t(60) = -2.87$, $p = 0.006$, $d = -0.74$.

The majority of participants were undergraduate students ($n = 41$), others were postgraduates ($n = 13$) or professionals ($n = 8$), and all but one participant had attended university. All participants gave informed consent prior to testing, had normal or corrected-to-normal vision and did not report to have a history of mental or neurological illness. All bilinguals completed a language history questionnaire adapted from Li, Sepanski, and Zhao (2006) with this information summarised in Table 1. Based on the self-rated proficiency scores, the bilingual group was characterised as highly proficient.

2.2. Materials

We administered standardised measures of working memory and non-verbal reasoning to all participants in order to ensure that the groups were comparable with regard to general cognitive function.

2.2.1. Working memory test

Working memory was assessed using the digit span task of the Wechsler Adult Intelligence Scale IV (Wechsler, 2008). In this task, participants are asked to repeat a set of single digits (between two and nine) after the experimenter. During the first round (eight sets of two trials) they are asked to repeat the numbers in the same order; in the second round (seven sets of two trials) they have to repeat the numbers in reverse. Each round is terminated once a participant has failed to correctly repeat both trials of one set, and a total score is calculated with a maximum of thirty points.

2.2.2. Non-verbal reasoning

Non-verbal reasoning was measured using the Raven's Advanced Progressive Matrices (RM; Raven & Court, 1998). In this task participants were presented with twelve trials. In each trial they were shown an incomplete matrix of black and white abstract figures. Participants were asked to identify the missing piece from a selection of eight alternatives and complete all 12 trials in no more than 10 min. None of the participants reached this time limit.

2.2.3. English language proficiency

In addition to the language history questionnaire, we also measured English language proficiency in bilinguals using the picture naming scale of the Bilingual Verbal Ability Tests (BVAT; Muñoz-Sandoval, Cummins, Alvarado, & Ruef, 1998).

2.3. The Dot Discrimination task

Participants completed a two-alternative-forced-choice task programmed in PsychoPy v. 1.82 (Peirce, 2009) presented on a 24-in. widescreen monitor using a standard keyboard. A MATLAB version of a similar task has previously been used in Fleming et al. (2014). On each trial participants saw two white circles on a black background, and indicated whether the left or the right circle contained the most dots by pressing the appropriate arrow key on a standard computer keyboard. For every trial, one circle was randomly assigned to have 50 dots; the other circle contained a variable number of dots that was either larger than, or smaller than 50. The difference in dots between the two circles was modified throughout the experiment by a staircase procedure, so that whenever participants correctly responded to two successive trials the task increased in difficulty (one less dot difference between the

Table 1
Bilingual participants' language history information, experiment 1.

Linguistic background	First language	Bulgarian ($n = 1$)
		Creole ($n = 1$)
		Dutch ($n = 2$)
		Farsi ($n = 1$)
		French ($n = 1$)
		German ($n = 2$)
		Hindi ($n = 1$)
		Hungarian ($n = 1$)
		Italian ($n = 2$)
		Lithuanian ($n = 1$)
		Malayalam ($n = 2$)
		Polish ($n = 7$)
		Portuguese ($n = 2$)
		Romanian ($n = 2$)
		Sinhalese ($n = 1$)
	Second language	Afrikaans ($n = 1$)
		English ($n = 26$)
		Frisian ($n = 1$)
		Greek ($n = 1$)
		Gujarati ($n = 1$)
		Twi ($n = 1$)
	Third language	English ($n = 1$)
Other linguistic background information	Age of first exposure	Birth–6 years ($n = 15$)
		7–12 years ($n = 9$)
		Teenage years ($n = 7$)
	Time spent in the UK	0–5 years ($n = 16$)
		5–10 years ($n = 9$)
		10+ years ($n = 6$)
	Switch	Rarely ($n = 14$)
		Sometimes ($n = 15$)
		Frequently ($n = 2$)
Self-rated proficiency (1–6)	Reading	$M = 5.1$; SD = 0.7
	Writing	$M = 4.6$; SD = 0.9
	Speaking	$M = 4.8$; SD = 0.8
	Listening	$M = 5.2$; SD = 0.7

options) and for every failed trial the task became easier (one more dot difference between the options). The purpose of the staircase was to normalise first order accuracy at 71% across the sample. After each trial participants were asked to indicate their confidence on a sliding scale. For a graphical representation of the trial structure see Fig. 2.

Participants completed 8 blocks with 25 trials in each, making up a total of 200 trials. Response times were unlimited for both the perceptual choices and the confidence judgments.

Prior to beginning the main task, participants were provided with three practice phases. In the first phase they were shown pairs of circles with the number of dots indicated in writing below the circles. In the second phase participants started making perceptual choices without conducting any confidence judgments. These trials started with a 20-item dot difference, which first changed in increments of four, then in successively smaller increments down to one; this was performed to calibrate the difficulty to each participant. The second phase terminated after 8 reversals (i.e., when participants had switched between picking the correct and the incorrect option 8 times). Participants received feedback on their choices in the second calibration phase. The final phase consisted of 10 trials that simulated the main experimental trials in every way, i.e., without performance feedback, and they were asked to indicate their confidence in their choice after each trial. All practice trials were excluded from all analyses.

2.4. Procedure

All participants were tested in one hour-long session at Anglia Ruskin University, Department of Psychology in the same room

Fig. 2. Trial structure of the Dot Discrimination task, experiment 1.

Table 2
Descriptive statistics for control measures, experiment 1.

	Monolinguals		Bilinguals	
	M	SD	M	SD
Working memory (maximum score: 30)	17.97	4.85	16.03	3.73
Nonverbal reasoning (maximum score: 20)	9.94	1.65	10.26	1.79

using the same equipment. After informed consent was given they completed a short demographics questionnaire and the bilingual participants were also asked to complete an adapted version of the language history questionnaire by Li et al. (2006). We then administered the digit span task, Raven's Matrices as well as the Dot Discrimination task, with task ordering counter-balanced across participants. The practice phases of the Dot Discrimination task were presented with extensive instructions and participants were encouraged to ask any questions prior to task commencement. All participants were given the option of entering into a raffle (using rafflecopter.com and blogger.com) for a £25 Amazon voucher that was sent to one randomly chosen participant upon completion of the study.

3. Results section

3.1. Equivalence of groups on control measures

An analysis of the control measures revealed that both groups performed comparably on measures of working memory, $t(56.40) = 1.71$, $p = .09$, $d = 0.43$ and nonverbal reasoning, $t(35.98) = -0.54$, $p = .59$, $d = 0.14$. Means and standard deviations are reported in Table 2. Therefore, any differences found in metacognitive abilities are unlikely to be attributable to group differences in general cognitive functioning. Because working memory differences were marginally significant we included working memory as a covariate in all analyses.

3.2. Dot Discrimination task: First order performance

We compared the bilinguals' and monolinguals' performance with regard to their first order accuracy (measured by percentage of correct responses), the difficulty of the trials (measured by dot difference) and response time of the choice and the confidence judgment (both measured in seconds). The results of all these analyses are summarised in Table 3.

The monolingual group had a mean accuracy of 70.98%, with a standard deviation of 1.06%, whilst the bilingual group had mean accuracy of 70.79% with a standard deviation of 1.23%. This indicates that the staircase procedure successfully standardised accuracy across participants. There were no significant group differences with regard to trial difficulty, with a mean dot difference of 4.64 for the monolinguals and 4.34 for the bilinguals. Additionally, with regards to response time for the confidence judgments there was no difference between the groups: the mono-

lingual group took, on average, 1179 ms to respond compared to 1112 ms for the bilinguals.

However, the groups did differ with regards to choice response time; an independent samples t-test showed that bilinguals ($M = 2679$ ms, SD = 923 ms) were significantly faster than monolinguals ($M = 3360$ ms, SD = 1475 ms, $t(50.38) = -2.18$, $p = .03$, $d = 0.55$). A random slopes multilevel model (MLM) revealed that this relationship was significantly mediated by block (see Fig. 3; for more detailed information about the MLM fitting see Appendix A). Monolinguals were set to be the reference category for this analysis and all subsequent MLMs. The model tells us that the main-effect of group became statistically non-significant when the block-group interaction was accounted for ($b = 283.11$, $t(64.03) = 0.71$, $p = 0.48$). The main effect of block was also non-significant ($b = 19.23$, $t(64.03) = 0.61$, $p = 0.54$), meaning that the response speed of the monolinguals did not change significantly over time, when individual variation in intercepts and slopes were accounted for. The bilingual group ∗ block interaction was significant ($b = -88$, $t(64.03) = -2.01$, $p = .05$), meaning that bilinguals, as a group, became faster as the task progressed.

Two participants in the monolingual group displayed outlying values for one variable (difficulty and response time, respectively). In order to ensure that these outliers did not unduly influence the group-wise comparisons they were both capped at 3 standard deviations above the group mean. Capping these values did not change the results for difficulty ($t(59.78) = 1.29$, $p = .20$, $d = 0.33$) or for response time ($t(55.17) = 2.24$, $p = .03$, $d = 0.57$).

### 3.3. Dot Discrimination task: Second order performance

To estimate metacognitive efficiency we used the Mratio. An Mratio was fitted to each participant using a hierarchical Bayesian estimation method (MATLAB code available at https://github.com/smfleming/HMM). The confidence data was binned into three quantiles prior to the meta-d' estimation. A two-tailed independent-samples t-test showed that the monolingual group had significantly higher Mratios than the bilingual group, $t(37.61) = 5.09$, $p < 0.001$, $d = 1.29$ (monolinguals' $M = 0.58$, SD = 0.09; bilinguals' $M = 0.49$, SD = 0.03).

Given that the response time of the choices varied across groups over the course of the experiment (see Fig. 3) and metacognitive processing has been associated with first-order response time (Baranski & Petrusic, 1998), we explored whether metacognitive efficiency also varied as a function of time. In order to test this we calculated meta-d' separately for the first and second half of

Table 3
Descriptive and inferential statistics for first order performance, experiment 1.

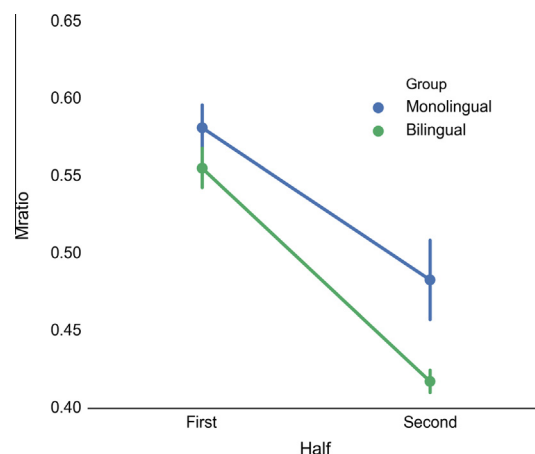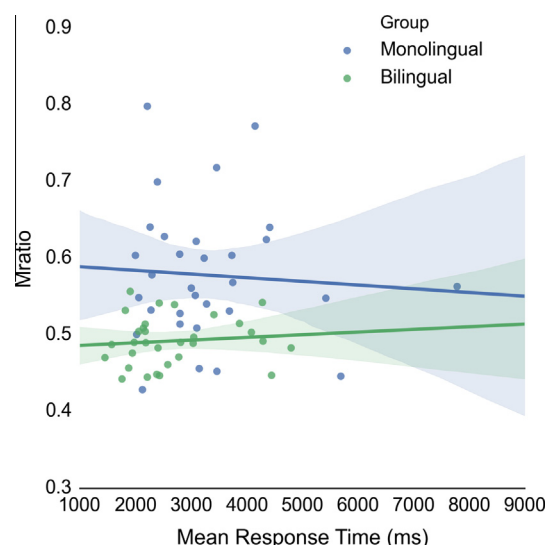|	Monolinguals		Bilinguals		*t*-statistic	*df*	*p*-value	Cohen's *d*
	Mean	SD	Mean	SD				
Accuracy (% correct)	70.98	1.06	70.79	1.23	0.66	58.73	0.51	0.17
Difficulty (Δ dot)	4.33	1.07	4.64	1.03	1.25	59.90	0.26	0.29
Response time confidence (ms)	1179	318	1112	280	0.89	58.77	0.37	0.23
Response time choice (ms)	3360	1475	2679	922	2.18	50.38	0.03	0.55
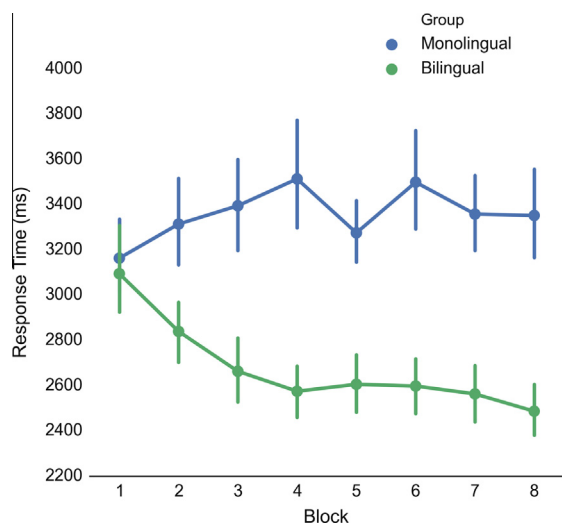
Fig. 3. Mean response times for each group by experimental block, experiment 1. Error bars show 95% group-wise confidence intervals.

Fig. 4. Mean Mratio for each group for the first and second half of experiment 1. Error bars represent group-wise 95% confidence intervals.

Fig. 5. Relationship between participant wise mean response times and Mratios. Shaded areas show boot-strapped 95% confidence intervals.

the experiment and compared the results across the groups. A fixed-slopes MLM showed that the main effect of group remained significant, when the other effects were accounted for ($b = -0.03$, $t(123.70) = 2.15$, $p = .03$) and both groups showed reduced metacognitive awareness over time ($b = -0.10$, $t(123.70) = -10.05$, $p < 0.001$). Most importantly, the interaction effect between group and time was significant ($b = -0.04$, $t(123.70) = -2.85$, $p = .005$), showing that bilinguals had a greater decrease in metacognitive efficiency between T1 and T2 than monolinguals, as illustrated in Fig. 4. For further information about the MLM model fitting, see Appendix A.

To further explore whether Mratios are influenced by response time, we tested whether a participant's mean response time predicted their Mratio (Fig. 5). We found that response times did not predict Mratios ($b > 0.01$, $t(58) = 0.26$, $p = .79$) when group affiliation was accounted for.

3.4. Effects of second language proficiency and working memory

Given the marked difference between monolinguals and bilinguals, we wanted to test whether second language proficiency related to metacognitive ability in the bilingual group. We tested this by predicting Mratios from BVAT scores for the bilingual group in a regression model. We found that BVAT scores did not significantly predict Mratios in the bilingual group ($b < 0.01$, $t(29) = 0.07$, $p = 0.94$). Additionally we tested whether accounting for the significant group difference in age and the trending difference in working memory scores (favouring monolinguals) influenced the results reported here. Our analyses revealed that the effects we found remained stable in both cases; these control analyses are reported in Appendix B.

4. Experiment 1 discussion

We presented young, healthy bilingual and monolingual adults with a two-alternative-forced-choice task that required them to make a perceptual judgement and subsequently to indicate their level of confidence in their choice (metacognitive performance). Task difficulty was calibrated for each participant individually and adjusted online throughout the experiment so that overall accuracy was set at 71% for each participant. This was implemented to eliminate potential effects of first-order task performance on second-order metacognitive performance.

The findings of this experiment were threefold. First, the evaluation of first-order performance revealed a bilingual advantage in choice response time despite similar levels of accuracy and difficulty across both groups, suggesting that bilinguals were better at basic perceptual discrimination: bilinguals became faster over time, whilst monolinguals maintained a stable response time throughout the Dot Discrimination task (see Section 3.2). Second, the analysis of second-order performance revealed that monolinguals had significantly higher Mratios compared to bilinguals. This indicates that the monolingual group demonstrated better metacognitive efficiency (they tended to feel more confident in trials they completed correctly and less confident in trials where their performance was incorrect, compared to their bilingual peers). Finally, we assessed metacognition over the progression of the task, revealing a significant interaction: whilst both groups experienced a reduction in metacognitive efficiency over time, this reduction was greater for the bilingual group compared to the monolingual group.

Research published by Baranski and Petrusic (1998, 2001) offers a possible interpretation for our findings. These authors showed that in tasks where there were no external time constraints on choices (as was the case with the current task), choice response times increased when participants were asked to make a subsequent confidence judgement (compared to only making a choice). Importantly, these choice response times correlated positively with metacognitive accuracy (Petrusic & Baranski, 2003). This indicates that, in conditions without time constraints, processing of metacognitive judgement takes place during the primary decision-making process.

The potential association between the bilingual response time advantage and metacognitive deficit requires two assumptions: First, bilinguals have a processing advantage in the networks associated in the first-order response compared to monolinguals. Second, both groups have similar levels of efficiency with regard to metacognitive processing. If these two assumptions are true, participants who have a more efficient first-order system will make decisions more quickly, but their speed will allow the metacognitive system less time to map the uncertainty of the decision, thus impairing metacognitive performance. Because bilinguals responded faster whilst displaying the same level of accuracy as monolinguals at comparable levels of difficulty, we can conclude that their speed advantage reflects a genuine processing advantage with regards to the perceptual choice. Consequently, bilinguals have less information to support their confidence judgements unless their metacognitive network has a processing speed advantage similar to that of their first-order network. A model that assumes equivalent metacognitive processing speeds but different first-order processing speeds thus accounts for our findings. Additionally, this model can explain the widening group difference, both in metacognitive performance and response speed, over the course of the task.

However, there is also a case against this interpretation of our findings. First of all, mean individual response times did not correlate with Mratios for either group (see Fig. 4). This means that people who responded faster on average were equally metacognitively efficient as those that responded slower on average, so between people there was no relationship between faster response time and lower-metacognitive accuracy. Second, the response times for both groups were excessively long for a perceptual discrimination task, so it is conceivable that the difference we found in response times reflects a difference in response thresholds rather than a genuine processing advantage for bilinguals. In order to address these problems, we ran a second experiment where we capped response times at 1.5 s, whilst maintaining the general structure of the first experiment.

Table 4
Bilingual participants' language history information, experiment 2.

Linguistic background	First language	Bengali (n = 2)
		Cantonese (n = 1)
		Chinese (n = 2)
		English (n = 4)
		French (n = 1)
		German (n = 1)
		Gujarati (n = 1)
		Greek (n = 1)
		Italian (n = 2)
		Korean (n = 1)
		Mandarin (n = 1)
		Nepalese (n = 2)
		Polish (n = 1)
		Portuguese (n = 2)
		Setswana (n = 1)
		Spanish (n = 2)
		Turkish (n = 4)
	Second language	English (n = 23)
		Farsi (n = 1)
		French (n = 1)
		Malay (n = 1)
		Punjabi (n = 3)
	Third language	English (n = 2)
		Urdu (n = 1)
Other linguistic background information	Age of first exposure	Birth–6 years (n = 20)
		7–12 years (n = 6)
		Teenage years (n = 3)
	Time spent in the UK	0–5 years (n = 13)
		5–10 years (n = 2)
		10+ years (n = 14)
	Switch	Rarely (n = 13)
		Sometimes (n = 14)
		Frequently (n = 2)
Self-rated proficiency (1–6)	Reading	M = 5.0; SD = 0.9
	Writing	M = 4.7; SD = 1.1
	Speaking	M = 4.8; SD = 1.1
	Listening	M = 5; SD = 1.0

5. Experiment 2

5.1. Method

5.1.1. Participants

For the second experiment, we recruited sixty-one participants: thirty-two English monolinguals (M_{age} = 20.4, SD = 0.7, range 19.3–22; 7 males) and a group of twenty-nine highly proficient bilinguals (M_{age} = 22.3, SD = 4.7, range 18.2–41.4; 6 males). Similarly to experiment one, the bilingual group was significantly older than the monolingual group, t(29.123) = −2.20, p = 0.04, d = −0.81, although this effect was reduced to a trend when one participant was excluded from analysis (see below). All of the participants were undergraduate students except for one, who was a postgraduate student. Informed consent was provided prior to testing and participants had normal or corrected-to-normal vision and did not report to have a history of mental or neurological illness. All bilinguals completed a language history questionnaire adapted from Li et al. (2006). The information deriving from this questionnaire is summarised in Table 4.

5.1.2. Materials

General cognitive functioning and English language proficiency were assessed using the same measures as outlined in Section 2.2.

5.1.2.1. The Dot Discrimination task. The Dot Discrimination task was identical to the task in experiment 1, with the exception that

participants now had to respond within 1.5 s after first seeing the dots. We also introduced slightly longer inter-trial intervals which featured a fixation cross in the centre of the screen (see Fig. 6). If participants took longer than 1.5 s to respond, the trial was terminated and the words "Too Slow", presented for one second.

5.1.3. Procedure

Testing took place at Anglia Ruskin University using the same equipment as in experiment 1. Participants completed the Dot Discrimination task, as well as all the control measures in a counterbalanced order. Participants were offered participation credits, which are part of the undergraduate course requirements.

5.2. Results

5.2.1. Participant exclusion

One participant from the bilingual group was excluded because they were found to have reported a confidence of 50% on 88% of the trials of the Dot Discrimination task. Because the confidence marker for the second order performance started at 50%, it is likely that this participant simply neglected to provide a confidence judgement for the majority of trials. Including this participant in the non-confidence analyses did not alter the direction or magnitude of any of the effects reported. Therefore, we proceeded to analyse the data provided by a sample of 60 participants, 28 bilinguals and 32 monolinguals.

5.2.2. Equivalence of groups on control measures

Both groups performed comparably on measures of working memory, $t(58) = 1.01$, $p = 0.32$, $d = 0.27$ and nonverbal reasoning, $t(58) = -0.86$, $p = 0.40$, $d = -0.23$, indicating that the groups were matched on general cognitive functioning (see Table 5). However, the bilinguals were marginally older than the monolinguals ($t(28.29) = 1.88$, $p = 0.07$, $d = .50$).

5.2.3. Dot Discrimination task: First order performance

Descriptive and inferential statistics are provided in Table 6. There was no difference in response times between the groups, either for the choices ($t(52.91) = 1.30$, $p = .20$, $d = 0.33$) or the confidence judgements ($t(56.98) = 0.28$, $p = .78$, $d = .07$). Similarly, participants had a comparable proportion of invalid trials for responding too slow, namely 2% for monolinguals and 3% for bilinguals ($t(51.88) = 1.74$, $p = .09$, $d = 0.45$). Additionally, both groups had a similar average difficulty ($t(54.82) = 1.02$, $p = .31$, $d = 0.26$),

with 5.85 average dot difference for the monolinguals and 5.45 average dot difference for the bilinguals. As in experiment 1, the staircase procedure successfully standardised accuracy across participants. The monolingual group had a mean accuracy of 71.34%, with a standard deviation of 1.38%, whilst the bilingual group had mean accuracy of 70.27% with a standard deviation of 1.04%. Given that response time changed over the course of experiment 1, we tested for similar patterns in experiment 2. We found that bilinguals were significantly slower than monolinguals when the block ∗ group interaction was accounted for ($b = 86.84$, $t(57.98) = 2.23$, $p = .03$), but that bilinguals became faster over time, while monolinguals became slower ($b = -10.03$, $t(57.98) = -2.32$, $p = .02$), such that the groups converged over the course of the experiment (see Fig. 7). There was no significant main effect of block ($b = 4.19$, $t(57.98) = 1.42$, $p = .16$).

5.2.4. Dot Discrimination task: Comparing first order performance across experiments

The participants in experiment 2 ($M = 875$ ms, SD = 128 ms) were significantly faster than the fastest group in experiment 1 (the bilinguals; $M = 2679$ ms, SD = 922 ms; $t(30.6) = 10.82$, $p < .0001$, $d = 2.73$). Response times for confidence judgements did not differ across experiments ($t(114.48) = 1.61$, $p = 0.11$, $d = 0.29$). However, incorporation of the speeded response requirement affected task difficulty such that in experiment 2 a larger dot difference was required ($M_{experiment1} = 4.48$, $SD_{experiment1} = 1.06$; $M_{experiment2} = 5.66$, $SD_{experiment2} = 1.55$; $t(103.82) = 4.90$, $p < .0001$, $d = 0.89$).

5.2.5. Dot Discrimination task: Second order performance

The monolingual group had significantly higher Mratios than the bilingual group ($t(48.62) = 2.70$, $p = .001$, $d = 0.71$), although the magnitude of the effect was smaller than experiment 1 (Cohen's d experiment 1 = 1.29, Cohen's d experiment 2 = 0.71). This suggests that monolinguals have a genuine advantage in metacognitive efficiency, and that the effect observed in experiment 1 is not solely due to differences in response times. Next, we ran an MLM model to test if metacognitive efficiency decreased over time, as was observed in experiment 1. We computed Mratio for both groups, and we found that bilinguals performed significantly worse on metacognitive ability ($b = -22.09$, $t(123.98) = -3.16$, $p = .002$). Additionally, we examined performance in the first and the second half of the task to determine whether the groups performed differently as a function of time. The main effect

Fig. 6. Trial structure of the Dot Discrimination task, experiment 2.

**Table 5**
Descriptive statistics for control measures, experiment 2.

	Monolinguals		Bilinguals	
	M	SD	*M*	SD
Working memory (maximum score: 30)	15.66	3.55	14.93	2.62
Nonverbal reasoning (maximum score: 12)	8.66	2.47	9.21	1.91

Table 6
Descriptive and inferential statistics for first order performance, experiment 2.

	Monolinguals		Bilinguals		t-statistic	df	p-value	Cohen's d
	Mean	SD	Mean	SD				
Difficulty (Δ dot)	5.85	1.79	5.45	1.22	1.02	54.82	0.31	0.26
Response time confidence (ms)	1061	443	1035	222	0.28	56.98	0.78	0.07
Response time choice (ms)	856	151	898	95	1.30	52.91	0.20	0.33
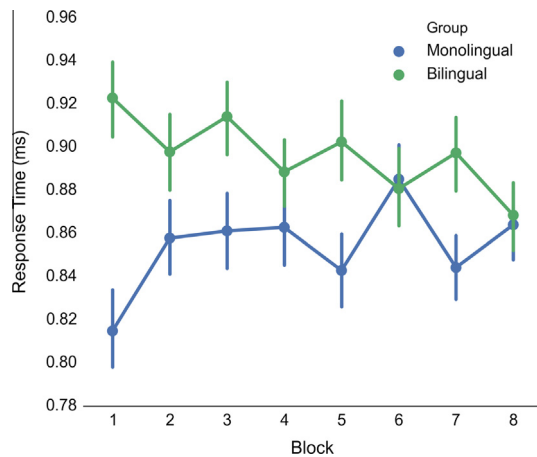
Fig. 7. Mean response times of each group by experimental block, experiment 2. Error bars show 95% group-wise confidence intervals.
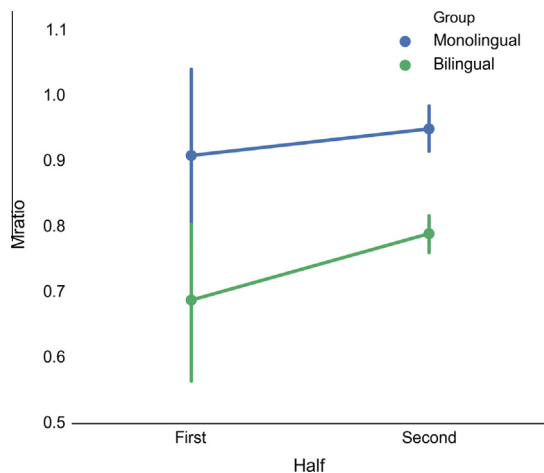
Fig. 8. Mean Mratio for each group for the first and second half of experiment 2. Error bars represent group-wise 95% confidence intervals.

of time was not significant ($b = 0.04$, $t(123.98) = −0.63$, $p = .53$), and neither was the group $*$ time interaction effect ($b = 0.06$, $t(123.98) = −0.65$, $p = .52$), as illustrated in Fig. 8.

5.2.6. Effects of second language proficiency

Consistent with experiment 1 we examined whether second language ability predicted metacognitive efficiency in the bilingual group. We tested this by predicting Mratios from BVAT scores for the bilingual group in a regression model. We found that levels of English proficiency did not significantly predict Mratios in the bilingual group ($b = 0.01$, $t(26) = 1.07$, $p = 0.30$). Additionally, we tested whether accounting for the marginal group difference in age influenced the results reported here. Our analyses revealed that the effects we found remained stable in either case; these control analyses are reported in Appendix B.

5.3. Discussion

We compared young, healthy monolinguals and bilinguals in a perceptual metacognition paradigm, in which participants chose which of two circles contained the most dots and subsequently made a confidence judgement about the accuracy of their choice. In contrast to experiment 1, where response time was unconstrained, we implemented a 1.5-s response threshold on first order responses. We did this to determine whether the differences between the language groups in response times and metacognitive efficiency observed in experiment 1 would replicate with constrained response times.

Because of the new design, participants' response times were significantly faster in both groups relative to experiment 1. This increase in response speed came at a cost to performance as expected by the trade-off between speed and accuracy. However, given the staircase procedure we implemented to keep the error-rate constant, this resulted in easier trials (i.e., a greater dot difference) compared to experiment 1. Critically, unlike in experiment 1, both language groups showed comparable response times during the visual discrimination. Additionally, both groups missed a similar number of trials because of responding too slowly (3% or less), and the two groups did not differ significantly with regard to average trial difficulty.

With regard to second-order performance, the monolingual group had significantly higher Mratios than the bilingual group, in line with experiment 1. As was the case in experiment 1, these differences could not be explained in terms of differences in non-verbal reasoning, working memory or age. However, unlike in experiment 1, there was no general decrease in Mratios over the course of experiment 2 in either group. Together, these findings therefore replicate our key observation in experiment 1, with monolinguals demonstrating higher metacognitive abilities compared with the bilingual group. Critically, the new design allowed us to confirm that this effect was not due to a difference in first-order performance given the undifferentiated response times and error rates across groups observed in experiment 2.

6. Conclusions

Our study demonstrates a bilingual disadvantage in metacognitive efficiency. In the first experiment we showed that bilinguals had reduced metacognitive efficiency but were also significantly faster in their response time compared with the monolingual individuals. Nevertheless, this response speed advantage did not reflect superior first-order discrimination performance. Given the possibility that the reduced metacognitive efficiency in the bilingual group was a consequence of the difference in reaction times (first-order performance) we conducted a second experiment in which we implemented a speeded response requirement. This new design resulted in both groups responding at similar speeds, but a metacognitive efficiency advantage was again observed in monolingual individuals in comparison to bilingual individuals.

Our findings might seem at odds with many previous studies that report a bilingual advantage in executive function (e.g., Adesope, Lavin, Thompson, & Ungerleider, 2010; Bialystok, 2009) and influential early work which has suggested links between

metacognitive and executive abilities (Fernandez-Duque et al., 2000; Shimamura, 2000). However, whether or not there is direct link between metacognitive monitoring and other aspects of executive function is not clear from these earlier experiments (and unpublished data from our lab have not provided conclusive results). Other studies have failed to find any relationship between task switching and meta-cognitive performance measured by the first order discrimination task (Boduroglu, Tekcan, & Kapucu, 2014; Palmer, David, & Fleming, 2014). However, Boduroglu and colleagues did find a relationship between set shifting and metacognitive abilities measured in a memory task in which participants were required to predict their ability to remember a target word from a visual cue, prior to making the memory judgement. It has been proposed that metacognitive abilities might be qualitatively different when measured in different domains such as perception or memory (Fleming et al., 2014) and therefore a link between executive function and metacognition might be domain specific.

The difference in metacognitive performance demonstrated here also opens up the possibility that monolinguals and bilinguals operate differently with regard to higher-order decision making. For example, metacognition has been implicated in tasks that require people to judge the value of staying with a current, known option, versus the value of exploring the environment and trying something new, the so-called exploration-exploitation trade-off (Cohen, McClure, & Yu, 2007; Kolling, Behrens, Mars, & Rushworth, 2012).

Studying bilingual speakers/learners is notoriously problematic due to the difficulty in controlling the wide range of possible extraneous variables that may frequently operate in group comparisons. Nevertheless, in two separate studies conducted across unique participant groups, we have identified a robust disadvantage in the bilingual groups operating on higher-level cognition. These effects warrant further exploration in the service of promoting a more comprehensive account of how language impacts on cognition.

Acknowledgements

Appendix A. Multilevel model fitting

The multilevel regression analyses reported in this paper were conducted using the *lme4* package in R (Bates, Maechler, & Bolker, 2011) Degrees of freedom and *p*-values were obtained using the Kenward-Roger approximation, as implemented in the *pbkertest* package (Halekoh & Hojsgaard, 2011).

In all regression analyses reported here the monolingual group serves as the reference category, so if the *t*-value for the group variable is positive, bilinguals have higher values than the monolinguals, and vice versa.

A.1. Experiment 1

A.1.1. Response time model

Table A1 lists the various models we attempted to fit, Fig. A1 shows the BIC scores for all the response time models. Table A2 show the full model specification for model 3, which is the best-

Table A1
Response time models, experiment 1.

| Models	Formulas
Empty	RT $\sim$ 1
Model 1	RT $\sim$ Language Group
Model 2	RT $\sim$ Language Group + Block + Group $*$ Block
Model 3	RT $\sim$ Language Group + Block + Group $*$ Block (Block as a random variable)
Model 4	RT $\sim$ Language Group + Block + Age + Group $*$ Block (Block as a random variable)
Model 5	RT $\sim$ Language Group + Block + WM + Group $*$ Block (Block as a random variable)

Table A2
Response time by language group and experimental block, random slopes model, experiment 1.

Effects	Variance	Coefficients	SE	*t*-values	DF	*p*
Random effects						
Participant						
Intercept	2,380,701					
Block	26,127					
Fixed effects						
Intercept		3273	283	11.57	64.03	<0.0001
Group (Bilingual)		−283	400	−0.71	64.03	0.48
Block		19	31	0.62	64.03	0.54
Group $*$ Block		−88	44	−2.01	64.03	0.05

Table A3
Mratio models, experiment 1.

Models	Formulas
Empty	Mratio $\sim$ 1
Model 1	Mratio $\sim$ Language Group
Model 2	Mratio $\sim$ Language Group + Half
Model 3	Mratio $\sim$ Language Group + Half + Group $*$ Half
Model 4	Mratio $\sim$ Language Group + Half + Age + Group $*$ Half
Model 5	Mratio $\sim$ Language Group + Half + WM + Group $*$ Half
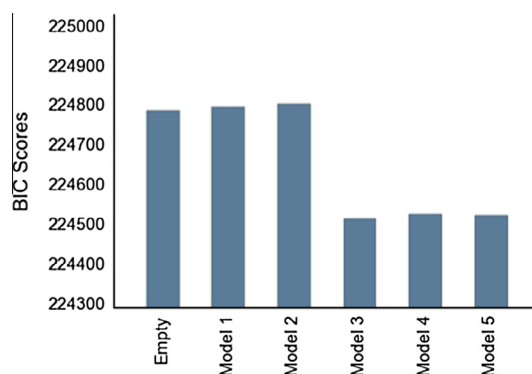
Fig. A1. BIC scores for all response time models, experiment 1.

fitting model, and the model that is reported in the main text of the paper.

A.1.2. Mratio model

Table A3, lists the multilevel models we used to predict Mratio for experiment 1. Fig. A2 shows the BIC scores for all of these models. Table A4 shows the full specifications for the best-fitting model
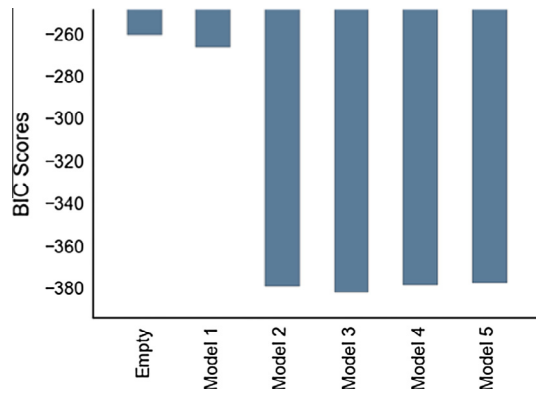
Fig. A2. BIC scores for all Mratio models, experiment 1.

Table A4
Mratio by language group and time point, fixed slopes, experiment 1.

Effects	Variance	Coefficients	SE	*t*-values	DF	*p*
Random effects						
Participant						
Intercept	0.0008					
Fixed effects						
Intercept		0.58	0.009	67.50	123.70	<.0001
Group		−0.03	0.01	−2.15	123.70	.034
Time		−0.10	0.01	−10.05	123.70	<.0001
Group ∗ Time		−0.04	0.01	−2.85	123.70	.005

Table A5
Response time models, experiment 2.

Models	Formulas
Empty	RT ∼ 1
Model 1	RT ∼ Language Group
Model 2	RT ∼ Language Group + Block + Group ∗ Block
Model 3	RT ∼ Language Group + Block + Group ∗ Block (Block as a random variable)
Model 4	RT ∼ Language Group + Block + Age + Group ∗ Block (Block as a random variable)

Table A6
Response time by language group and experimental block, random slopes model, experiment 2.

Effects	Variance	Coefficients	SE	*t*-values	DF	*p*
Random effects						
Participant						
Intercept	21,450					
Block	242					
Fixed effects						
Intercept		837	26	31.69	58.01	<.0001
Group (Bilingual)		84	38	2.17	58.01	0.03
Block		4	3	1.42	58.01	0.16
Group ∗ Block		−10	4	−2.21	58.01	0.03

(that is reported in the main text of the paper). There are no random-slopes models in this comparison because there were too few data points to support such complex models.

A.2. Experiment 2

A.2.1. Response time model

Table A5 lists the various response time models we attempted to fit, Fig. A3 shows the BIC scores for all the response time models in experiment 2. Table A6 shows the full model specification for model 3, which is the best-fitting model, and the model that is reported in the main text of the paper.

Table A7
Mratio models, experiment 2.

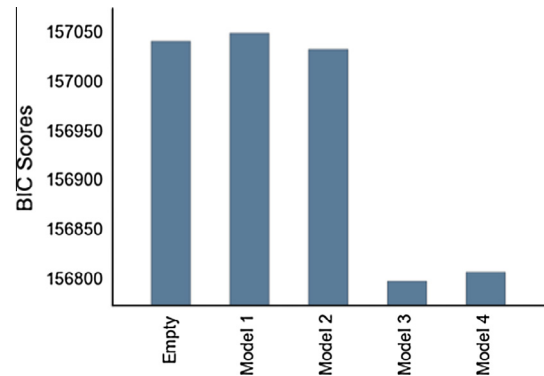| Models	Formulas
Empty	Mratio ∼ 1
Model 1	Mratio ∼ Language Group
Model 2	Mratio ∼ Language Group + Half
Model 3	Mratio ∼ Language Group + Half + Group ∗ Half
Model 4	Mratio ∼ Language Group + Half + Age + Group ∗ Half

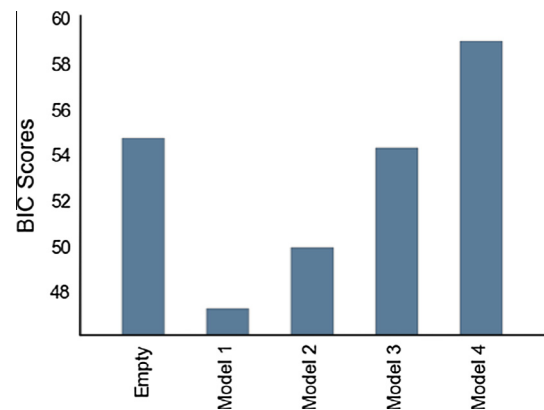**Fig. A3.** BIC scores for all response time models, experiment 2.

Fig. A4. BIC scores for all Mratio models, experiment 2.

A.2.2. Mratio model

Table A7, lists the multilevel models we used to predict Mratio. Fig. A4 shows the BIC scores for all of these models. Table A8 shows the full specifications for model 3, which is reported in the main text of the paper. Note that model 3, was not the best-fitting model in this case, but it was still included in the paper to allow for clear comparisons with experiment 1. There are no random-slopes models in this comparison because there were too few data points to support such complex models.

Appendix B. Analyses controlling for potential confounds

B.1. Experiment 1

Because age differed significantly between the groups in experiment 1, and working memory differed on the trend level, we reran all the analyses with these two covariates, to ensure that these differences did not influence our results.

Table A8
Mratio by language group and time point, fixed slopes, experiment 2.

Effects	Variance	Coefficients	SE	t-values	DF	p
Random effects						
Participant						
Intercept	0.007					
Fixed effects						
Intercept		0.91	0.047	19.05	123.98	<0.0001
Group		−0.22	0.07	−3.16	123.98	0.002
Time		0.04	0.06	0.63	123.98	0.53
Group * Time		0.06	0.01	0.65	123.98	0.52

Table A9
First order analyses, age covariate (df = 2, 59), experiment 1.

DV	Group		Age	
	t-value	p-value	t-value	p-value
Accuracy	−1.72	0.09	3.06	0.003
Difficulty	1.71	0.09	−1.83	0.07
Confidence RT	−1.19	0.24	1.02	0.31
Choice RT	−2.51	0.01	1.31	0.20

Table A10
First order analyses, WM covariate (df = 2, 59), experiment 1.

DV	Group		WM	
	t-value	p-value	t-value	p-value
Accuracy	−0.62	0.54	0.1	0.93
Difficulty	1.05	0.3	−0.19	0.85
Confidence RT	−0.62	0.54	1.22	0.23
Choice RT	−1.89	0.06	1.15	0.26

B.2. First order analyses

To see if age or working memory influenced our first order analyses we ran two sets of regression models that predicted the first order variables from group and adding either age or working memory as covariates. Neither covariate altered the findings reported in the main text (see Tables A9 and A10).

B.2.1. The bilingualism * block interaction in predicting RT

To control for the potential influence of age and working memory on the bilingualism * block interaction we ran two MLMs, one adding age as covariate, and one adding working memory as a covariate. For the age model, bilinguals did not differ significantly from monolinguals (b = −412, $t(72.95)$ = −1.01, p = .32), neither was age a significant predictor of response time (b = −43, $t(72.95)$ = 1.18, p = .24), nor was block (b = −19, $t(72.95)$ = 0.62, p = .53). However, the block * bilingual interaction showed that bilinguals responded significantly faster as the task progressed (b = −88, $t(72.95)$ = −2.01, p = .05). The exact same pattern was true for the working memory model, with the working memory (b = 35, $t(74.93)$ = −0.99, p = .32), group (b = −218, $t(74.93)$ = −0.54, p = .59) and block (b = 19, $t(74.93)$ = −0.61, p = .53) main effects being insignificant, but the block * group interaction being significant (b = −88, $t(74.93)$ = −2.01, p = .05). Both of these models show the same pattern of effects as is reported in the main text of the paper.

B.3. Mratio analyses

B.3.1. The main effect of bilingualism

To examine whether the main effect of bilingualism on Mratio was influenced by working memory or age, we constructed a set of regression models, as we did for response time. We found that age was not a significant predictor of Mratio (t = 0.92, df = 2, 59, p = .33), nor did including it in the model alter the effect of group on Mratios (t = −5.11, df = 2, 59, p < .0001). Similarly, working memory scores did not predict Mratios (t = −1.16, df = 2, 59, p = .25), nor did including them in the model, change the influence of language group (t = −4.72, df = 2, 59, p < .0001). Including the covariates in our analysis did not change the effect of group on Mratios that is reported in the paper.

B.3.2. The bilingualism * time interaction

To control for the potential influence of age and working memory on the bilingualism * time interaction we ran two additional MLMs, one adding age as covariate, and one adding working memory as a covariate. For the age model, we found that age did not significantly predict Mratios (b = 0.002, $t(71.53)$ = 1.24, p = .22), and the effects remained significant for group (b = −0.03, $t(71.53)$ = −2.43, p = .02) time (b = −0.10, $t(71.53)$ = −10.05, p < .0001) and the group time interaction (b = −0.04, $t(71.53)$ = −2.85, p = .006). Similarly, working memory did not predict Mratios (b = 0.001, $t(74.93)$ = 0.86, p = .39), and including working memory did not greatly modify the effects of group (b = 0.02, $t(74.93)$ = −1.96, p = .05) time (b = −0.10, $t(74.93)$ = −10.05, p < .0001) or the group time interaction (b = .04, $t(74.93)$ = −2.85, p = .006). Accounting for the covariates did not change the direction or magnitude of the effects reported in the main text.

B.3.3. Predicting Mratio from Mean RT

In the main text we found that Mean RT did not predict a participant's Mratio. For completeness, we ran two additional models to ensure that adding age or WM as covariates would not change this null-effect. Age did not predict Mratio (t = 1.04, df = 4, 57, p = .30), and neither did mean RT (t = 0.18, df = 4, 57, p = .85). Furthermore, there was no significant Group * Mean RT interaction (t = −0.48, df = 4, 57, p = .63) but the effect of group was significant (t = 2.37, df = 4, 57, p = .02). The same pattern held for the WM model: WM did not predict Mratio (t = 1.25, df = 4, 57, p = .22), and neither did mean RT (t = 0.26, df = 4, 57, p = .80), nor the Group * Mean RT interaction (t = −0.56, df = 4, 57, p = .57). However, the effect of group was significant (t = 2.26, df = 4, 57, p = .03). All of these effects are in line with what is reported in the main body of the text.

B.4. Experiment 2

Because age differed significantly between the groups in experiment 2, we reran all the analyses with age as a covariate, to ensure that the age difference did not cause our results.

B.5. First order analyses

To see if age influenced our first order analyses we ran regression models where we predicted the first order variables from group and adding age as a covariate. We note that age is significant predictor of confidence response time and a marginally significant

Table A11
First order analyses, age covariate (df = 2, 57), experiment 2.

DV	Group		Age	
	t-value	p-value	t-value	p-value
Accuracy	−3.83	0.0005	1.94	0.06
Difficulty	−1.26	0.21	1.15	0.25
Confidence RT	−0.97	0.33	2.68	0.01
Choice RT	1.09	0.28	0.47	0.63

predictor of accuracy, but including age in the model does not change the effect of group on any first-order variable (see Table A11).

B.5.1. The bilingualism ∗ block interaction in predicting RT

To control for the potential influence of age on the bilingualism ∗ block interaction we ran an additional MLM, adding age as covariate. Age was not a significant predictor of response time ($b = 3.20$, $t(58.12) = 0.55$, $p = .58$) and neither was block ($b = 4.19$, $t(58.12) = 1.42$, $p = .24$). However, there was a significant main effect of group ($b = 79$, $t(58.12) = 1.97$, $p = .05$) and a group ∗ block interaction that showed that bilinguals responded significantly faster as the task progressed ($b = -9$, $t(58.12) = -2.21$, $p = .03$). All of these effects are in line with what was reported in the main text of the study.

B.6. Mratio analyses

B.6.1. The main effect of bilingualism

To examine whether the main effect of bilingualism on Mratio was influenced by age, we constructed a regression model predicting Mratio form group and age. We found that age was not a significant predictor of Mratio ($t = 0.01$, $df = 2, 57$, $p = .99$), nor did including it in the model modify the influence of group on Mratios ($t = -2.64$, $df = 2, 57$, $p = .01$).

B.6.2. The bilingualism ∗ time interaction

To control for the potential influence of age on the bilingualism ∗ time interaction model we ran an additional model, adding age as covariate. For the age model, we found that age did not significantly predict Mratios ($b = -0.003$, $t(69.75) = -0.34$, $p = .73$), nor did time ($b = 0.04$, $t(69.75) = 0.63$, $p < .53$), nor did the group ∗ time interaction ($b = -0.06$, $t(69.75) = 0.65$, $p < .52$). However, the main effect for group was significant ($b = -0.22$, $t(69.75) = -3.04$, $p = .003$), showing that bilinguals had lower Mratios than monolinguals. These findings are in line with the effects reported in the main text of the paper.

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2016.02.008.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275. http://dx.doi.org/10.1016/j.neuroling.2006.10.003.

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207–245.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 929–945. http://dx.doi.org/10.1037/0096-1523.24.3.929.

Baranski, J. V., & Petrusic, W. M. (2001). Testing architectures of the decision–confidence relation. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 55*(3), 195–206. http://dx.doi.org/10.1037/h0087366.

Barrett, A. B., Dienes, Z., & Seth, A. K. (2013). Measures of metacognition on signal-detection theoretic models. *Psychological Methods, 18*(4), 535–552. http://dx.doi.org/10.1037/a0033268.

Bates, D., Maechler, M., & Bolker, B. (2011). Linear mixed-effects models using S4 classes. R package version 0999375–42. *R: A language and environment for statistical computing.*

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70*(3), 636–644. http://dx.doi.org/10.1111/1467-8624.00046.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(01), 3–11.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290–303. http://dx.doi.org/10.1037/0882-7974.19.2.290.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science, 7*(3), 325–339. http://dx.doi.org/10.1111/j.1467-7687.2004.00351.x.

Bialystok, E., & Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition, 112*(3), 494–500.

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in stimulus–stimulus inhibition. *Bilingualism: Language and Cognition, 17*(03), 610–629.

Boduroglu, A., Tekcan, A. İ., & Kapucu, A. (2014). The relationship between executive functions, episodic feeling-of-knowing and confidence judgements. *Journal of Cognitive Psychology, 26*(3), 333–345.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282–298. http://dx.doi.org/10.1111/j.1467-7687.2008.00675.x.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1481), 933–942.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 308–325). New York, NY: Oxford University Press.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(03), 231–244. http://dx.doi.org/10.1017/S1366728999000334.

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), 135–149. http://dx.doi.org/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. http://dx.doi.org/10.1016/j.cognition.2006.12.013.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism an example of publication bias? *Psychological Science, 26*(1), 99–107.

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience, 16*(1), 105–110. http://dx.doi.org/10.1038/nn.3279.

Del Missier, F., Mäntylä, T., & Bruine de Bruin, W. (2010). Executive functions in decision making: An individual differences approach. *Thinking & Reasoning, 16*(2), 69–97. http://dx.doi.org/10.1080/13546781003630117.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language, 42*(4), 445–464. http://dx.doi.org/10.1006/jmla.1999.2697.

Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(03), 175–197. http://dx.doi.org/10.1017/S1366728902003012.

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., & Carreiras, M. (2014). The inhibitory advantage in bilingual children revisited: Myth or reality? *Experimental Psychology, 61*(3), 234–251. http://dx.doi.org/10.1027/1618-3169/a000243.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*(2), 288–307. http://dx.doi.org/10.1006/ccog.2000.0447.

Filippi, R., Leech, R., Thomas, M. S., Green, D. W., & Dick, F. (2012). A bilingual advantage in controlling language interference during sentence comprehension. *Bilingualism: Language and Cognition, 15*(04), 858–872. http://dx.doi.org/10.1017/S1366728911000708.

Filippi, R., Morris, J., Richardson, F. M., Bright, P., Thomas, M. S., Karmiloff-Smith, A., & Marian, V. (2015). Bilingual children show an advantage in controlling verbal interference during spoken language comprehension. *Bilingualism: Language and Cognition, 18*(03), 490–501. http://dx.doi.org/10.1017/S1366728914000686.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8.*

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain: A Journal of Neurology, 137*(10), 2811–2822. http://dx.doi.org/10.1093/brain/awu221.

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review, 10*(4), 843–876. http://dx.doi.org/10.3758/BF03196546.

García, G. E., Jiménez, R. T., & Pearson, P. D. (1998). Metacognition, childhood bilingualism, and reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 193–219). Mahwah, NJ: Erlbaum.

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., ... Jones, L. (2014). Does language dominance affect cognitive performance in

bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology, 5*, 11. http://dx.doi.org/10.3389/fpsyg.2014.00011.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*(2), 210–223. http://dx.doi.org/10.1016/0093-934X(86)90016-7.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(02), 67–81. http://dx.doi.org/10.1017/S1366728998000133.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*(5), 515–530.

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. *Neuroscience and Biobehavioral Reviews, 55*, 88–97.

Halekoh, U. and Højsgaard, S., Linear mixed models fitted with lmer () in R: p–values based on a Kenward-Roger modification of the F-statistic or on parametric bootstrap methods, 2011.

Hoffrage, U. (2004). 13 Overconfidence. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory, 235*.

Hernandez, A. E., Bates, E. A., & Avila, L. X. (1996). Processing across the language boundary: A cross-modal priming study of Spanish–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 846–864. http://dx.doi.org/10.1037/0278-7393.22.4.846.

Kolling, N., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2012). Neural mechanisms of foraging. *Science, 336*(6077), 95–98.

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science, 18*(1), 64–71. http://dx.doi.org/10.1111/j.1467-9280.2007.01850.x.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods, 38*(2), 202–210.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535. http://dx.doi.org/10.1037/0278-7393.33.3.503.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition, 21*(1), 422–430. http://dx.doi.org/10.1016/j.concog.2011.09.021.

Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: Meta-d′, response-specific Meta-d′, and the unequal variance SDT model. In S. M. Fleming & C. D. Frith (Eds.), *The cognitive neuroscience of metacognition* (pp. 25–66). Berlin: Springer.

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. http://dx.doi.org/10.1017/S1366728907003227.

Masson, M. E., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 509.

McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience, 33*(5), 1897–1906. http://dx.doi.org/10.1523/JNEUROSCI.1890-12.2013.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature, 431*(7010), 757. http://dx.doi.org/10.1038/431757a.

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron, 88*(1), 78–92.

Morales, J., Gómez-Ariza, C. J., & Bajo, M. T. (2013). Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology, 25*(5), 531–546.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*(6), 719–726. http://dx.doi.org/10.1111/j.1467-7687.2007.00623.x.

Muñoz-Sandoval, A. F., Cummins, J., Alvarado, C. G., & Ruef, M. L. (1998). *Bilingual verbal ability tests: Comprehensive manual*. Itasca, IL: Riverside.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*(1), 109. http://dx.doi.org/10.1037/0033-2909.95.1.109.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: The MIT Press.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258. http://dx.doi.org/10.1016/j.cogpsych.2012.12.002.

Paap, K. R., Johnson, H. A., & Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *Journal of Cognitive Psychology, 26*(6), 615–639.

Paap, K. R., Sawi, O. M., Dalibar, C., Darrow, J., & Johnson, H. A. (2014). The brain mechanisms underlying the cognitive benefits of bilingualism may be extraordinarily difficult to discover. *AIMS Neuroscience, 1*(3), 245–256.

Palmer, E. C., David, A. S., & Fleming, S. M. (2014). Effects of age on metacognitive efficiency. *Consciousness and Cognition, 28*, 151–160. http://dx.doi.org/10.1016/j.concog.2014.06.007.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2*, 10. http://dx.doi.org/10.3389/neuro.11.010.2008.

Peirce, C. S., Jastrow, J. (1885). *On small differences of sensation* (pp. 73–83). US Government Printing Office.

Petrusic, W. M., & Baranski, J. V. (2003). Judging confidence influences decision processing in comparative judgments. *Psychonomic Bulletin & Review, 10*(1), 177–183. http://dx.doi.org/10.3758/BF03196482.

Raven, J. C., & Court, John Hugh. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Psychologists Press.

Rubio-Fernández, P., & Glucksberg, S. (2012). Reasoning about other people's beliefs: Bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(1), 211.

Schwartz, B. L., & Díaz, F. (2014). Quantifying human metacognition for the neurosciences. In *The cognitive neuroscience of metacognition* (pp. 9–23). Berlin, Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-45190-4_2.

Shimamura, A. P. (2000). Toward a cognitive neuroscience of metacognition. *Consciousness and Cognition, 9*(2), 313–323.

Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Consciousness and Cognition, 20*(4), 1787–1792. http://dx.doi.org/10.1016/j.concog.2010.12.011.

Souchay, C., & Isingrini, M. (2004). Age related differences in metacognitive control: Role of executive functioning. *Brain and Cognition, 56*(1), 89–99. http://dx.doi.org/10.1016/j.bandc.2004.06.002.

Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language, 137*, 50–61. http://dx.doi.org/10.1016/j.bandl.2014.07.005.

Stocco, A., Yamasaki, B., Natalenko, R., & Prat, C. S. (2014). Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. *International Journal of Bilingualism, 18*(1), 67–92. http://dx.doi.org/10.1177/1367006912456617.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*(4), 780–789. http://dx.doi.org/10.3758/BF03196335.

Veenman, M. V., Elshout, J. J., & Meijer, J. (1997). The generality vs domain-specificity of metacognitive skills in novice learning across domains. *Learning and Instruction, 7*(2), 187–209. http://dx.doi.org/10.1016/S0959-4752(96)00025-4.

Wechsler, D. (2008). *WAIS-IV, Wechsler Adult Intelligence Scale–Fourth edition (WAIS–IV)*. San Antonio, TX: Psychological Corporation.

Yeung, N., & Summerfield, C. (2014). Shared mechanisms for confidence judgements and error detection in human decision making. In *The cognitive neuroscience of metacognition* (pp. 147–167). Berlin, Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-45190-4.