The role of bilingualism in executive functions in healthy older adults: A systematic review

Merve Gul Degirmenci
Judith Alina Grossmann
Heidelberg University, Germany

Patric Meyer
Heidelberg University, Germany
SRH University Heidelberg, Germany

Birgit Teichmann
Heidelberg University, Germany

Abstract
Aims and Objectives: Bilinguals have been claimed to develop superior executive functioning compared to monolinguals due to their continuous experience of controlling two languages. Given the developmental trajectory of executive functions, a bilingual advantage could be more pronounced at an advanced age. To gain a clearer understanding, we reviewed the effect of bilingualism executive functions in healthy older adults.

Methodology: The present paper systematically examines the methods and the results of 24 studies from 22 articles comparing healthy older monolinguals and bilinguals in at least one domain of executive functions.

Data and Analysis: Data of each study were extracted for sample characteristics, country, language background and measures, controlled confounders and task paradigms. Study quality was also calculated for each study.

Findings and Conclusions: In general, nine out of the 24 studies fully supported the notion of a bilingual advantage. Four studies showed a bilingual disadvantage. The rest of the studies challenged the existence of a bilingual advantage, as neither full support for bilingual advantages nor bilingual disadvantages were seen in various domains. The available data did not clearly support the widespread notion that bilingualism is related to a general advantage in executive control. However, when looking at the domains of executive functions separately, bilingualism was reliably associated with an advantage in inhibition, especially in two commonly applied tasks: the Stroop test and the Simon task.

Corresponding author:
Merve Gul Degirmenci, Network Aging Research, Heidelberg University, Bergheimer Strasse 20, Heidelberg, 69115, Germany.
Email: degirmenci@nar.uni-heidelberg.de
Originality: This is the first systematic review aimed at exploring the link between bilingualism and executive functions in healthy older adults.

Significance/Implications: Heterogeneity in study characteristics and control of confounding variables may partially explain some of the inconsistencies found between studies. Therefore, well-designed studies that measure all core domains of executive functions and consider confounding variables are urgently needed.

Keywords
Aging, bilingual advantage, cognitive control, shifting, switching, inhibition, updating, working memory

Introduction
In its most general form, bilingualism refers to the ability to use two languages in everyday life (Grosjean, 2010) and can be considered as a complex mental activity (Adesope et al., 2010; Bialystok et al., 2007). In the field of cognitive sciences, many researchers are motivated to study the effect of bilingualism on cognition to explore whether there is a bilingual advantage. It is assumed to stem from the experience of the use of two languages that require efficient cognitive control. Even if bilinguals do not use both languages actively for communication, they need efficient cognitive control functions to enable fluent use of the appropriate language and to inhibit the other language (Rodriguez-Fornells et al., 2006; Wu & Thierry, 2010). The findings regarding young adults, however, have been mixed, with some studies showing a bilingual advantage (Vega-Mendoza et al., 2015) and others documenting comparable performance between bilinguals and monolinguals (Paap & Greenberg, 2013). In general, the bilingual advantage seems to be greater among children and older adults than among young or middle-aged adults (Bak et al., 2014; Bialystok et al., 2004, 2008; Calvo & Bialystok, 2014; Hilchey & Klein, 2011). This finding is in line with the observation that bilingualism attenuates the age-related decline of cognitive processes (Bialystok et al., 2008).

Recent evidence for a bilingual advantage in cognition has led to the investigation of whether bilingualism affects improved executive functions (EFs). It is a broad umbrella term and incorporates diverse cognitive operations such as inhibition, shifting or updating, which are inter-related as well as inter-dependent (Miyake et al., 2000). These complex cognitive processes are closely related to the maturation of the prefrontal cortex (PFC). Due to the delayed maturation of the PFC, EFs are among the last functions to reach maturity. Moreover, the development of EFs is bell-shaped with regard to their acquisition and their later loss, highlighting the existence of nonlinear trajectories (Anderson et al., 2008).

Inhibition is the ability to deactivate irrelevant or distracting information. It is an active cognitive suppression of the habitual response to prevent task-irrelevant information from entering working memory (WM) (Hasher et al., 1999). It has been proposed that a decline in the efficiency of inhibitory processes is an important cause of age-related changes in a wide range of cognitive tasks (Persad et al., 2002). Correspondingly, aging results in poorer performance in a variety of well-established paradigms dependent on inhibitory processing, such as the stop-signal task or the Stroop test (Anderson et al., 2008).

Shifting is the cognitive ability to detect inappropriate responses during changing or unexpected events and make the appropriate changes to adapt behaviour and thoughts to new situations (Miyake et al., 2000). According to results of task-switching paradigms typically used in the aging literature, age is associated with difficulties in maintaining two competing mental...
sets in mind, rather than the difficulty in the specific execution of a mental switch (Kray & Lindenberger, 2000).

Updating is closely linked to the notion of WM, which holds information in mind and works mentally with it. Overall, the literature on aging suggests that WM declines early in the normal aging process (Bowles & Salthouse, 2003; Burke & Barnes, 2006; Cappell et al., 2010).

In light of the literature, we can assume that the bilingual advantage may be more pronounced in the inhibition domain (Bialystok, 2001; Bialystok et al., 2004). However, not only inhibiting the irrelevant language but also switching between languages is required for the efficient use of two languages (Green, 1998). Bilingual advantages have been shown to be pronounced in the shifting domain, revealing a relationship between task switching and language switching (Garbin et al., 2010; Prior & Gollan, 2011). Bilinguals switch more efficiently than monolinguals on non-linguistic tasks, which might be explained by a shared switch mechanism. Thus, this might be a cognitive ability that monolingual speakers do not develop to the same degree as bilinguals (Prior & MacWhinney, 2010). Furthermore, efficient WM capacity is required to monitor and activate the two languages and choose the appropriate target language (Bialystok et al., 2008; Luo et al., 2013). Although the findings for non-verbal EF tasks mostly support a bilingual advantage, bilingual participants also showed a disadvantage in verbal tasks compared to monolinguals (Bialystok, 2009).

Whether EFs are modified by bilingualism is highly debated among researchers. Despite the extensive efforts of previous systematic reviews and meta-analysis, the evidence regarding a bilingual advantage remains inconclusive and controversial, as these reviews have reported varying results (Adesope et al., 2010; de Bruin, Treccani, et al., 2015; Donnelly, 2016, 2019; Grundy & Timmer, 2017; Hilchey & Klein, 2011; Hilchey et al., 2015; Lehtonen et al., 2018; Noort et al., 2019; Zhou & Krott, 2016).

To summarize, Adesope et al. (2010) performed the first meta-analysis in this field and found a bilingual advantage in various cognitive domains, including metalinguistic and meta-cognitive awareness, abstract and symbolic representation, attentional control, problem-solving and WM. Hilchey and Klein (2011) found little evidence for a bilingual advantage in non-verbal inhibitory tasks (the Simon task, Flanker task and Attentional Network Task [ANT]) in children and adults. In their following work, contrary to their earlier review, Hilchey et al. (2015) did not find any association between bilingualism and improved cognitive abilities. De Bruin, Treccani, et al. (2015) reported a small effect on several EF tasks. Zhou and Krott (2016) reported a bilingual advantage in the three most common non-verbal interference tasks (the Simon task, spatial Stroop test and Flanker task), which employ shorter cut off criteria that is more likely to report null effects. They also claimed that instead of an enhanced inhibitory control ability, bilinguals might have increased an attentional control ability. Donnelly (2016) did not report the consistent effects of bilingualism on global reaction time (RT) and interference costs, which refers to the difference in RT between incongruent and congruent trials. Grundy and Timmer (2017) found a small to moderate positive effect size for a difference in WM between bilinguals and monolinguals. Donnelly et al. (2019) found a very small but significant bilingual advantage in global RT in interference-control tasks. However, the effects became non-significant once corrected for publication bias. Lehtonen et al. (2018) found a bilingual advantage in inhibition, switching and WM. A small disadvantage for bilinguals was only found for verbal fluency. There were no differences in attention or monitoring. However, the evidence for a bilingual advantage in all domains could not be held after adjustment for publication bias. In a recent systematic review, Noort et al. (2019) found that the majority (54.3%) of studies reported a bilingual advantage and the rest showed mixed results (28.3%) and no advantage (17.4%) on various EF tasks (e.g. the Flanker task, Simon task and Stroop test).
The present study

The interaction between bilingualism and its impact on EFs cannot be reduced to a simple polar question. Even if there is a bilingual advantage, its impact may often be obscured given the number of variables and experiences that can modulate EFs and that need to be considered.

As mentioned earlier, the bilingual advantage might be more pronounced in older adults because of the trajectory of EFs (Bak et al., 2014; Bialystok et al., 2008; Luo et al., 2010). To our knowledge, there is no published systematic review or meta-analysis that reported only the results of older adults. Although almost all of the former reviews included aging studies, only two studies reported separate results related to older adult participants (Donnelly et al., 2016; Lehtonen et al., 2018). Donnelly et al. (2016) reported significantly larger effect sizes for interference costs than global RT amongst older adults. Lehtonen et al. (2018) reported a larger difference in a task-switching paradigm favouring older bilinguals over younger ones. Instead of expecting to see whether there is a bilingual advantage without sorting different age groups that have different levels of EFs, we set age as an exclusion criterion in our review, different from previous reviews. To address this goal, we reviewed the currently available literature on bilingualism and EFs in healthy older adults.

Moreover, research on bilingualism is challenging due to participants’ various linguistic profiles and socio-demographic backgrounds (Bialystok, 2001). In that way, depending on their life experience, bilinguals can be different from each other in many aspects. Heterogeneity in linguistic experiences has been shown to result in diverse cognitive consequences (Coderre & van Heuven, 2014; Goral et al., 2015; Kalia et al., 2014; Mishra et al., 2012; Soveri et al., 2011). Larger advantages would be expected in bilinguals with higher second language (L2) proficiency, with the assumption that higher L2 proficiency will be associated with stronger demands and enhanced cognitive control than bilinguals with lower proficiency. Furthermore, researchers have speculated that the effects of bilingualism on interference-control tasks might be influenced by the age of acquisition (AoA) of the L2; however, effects have been mixed. Pelham and Abrams (2014) reported a bilingual advantage of similar magnitude for early and late Spanish-English, relative to monolinguals, on the ANT. In contrast, other studies suggest larger effects for bilinguals with early AoA on the Flanker task, such as smaller interference costs or faster global RTs in children and adults (Luk, De Sa, et al., 2011; Kapa & Colombo, 2013; Tao et al., 2011).

At the same time, socio-demographic factors such as the country in which a study is conducted, socioeconomic status (SES) and immigration status can also drive different language experiences and are related to general cognitive performance (Morton & Harper, 2007; Yang et al., 2011). To some extent, performance differences of bilingual participants could also be the result of the testing language. In a meta-analysis conducted by Grundy and Timmer (2017), the small to moderate bilingual advantage in WM was reduced when bilinguals performed the tasks in their L2.

Although all EF domains are both inter-related and inter-dependent, a benefit of bilingualism might not be equally seen in all domains. Hence, it is important to consider a large number of domains and tasks. Unlike some of the previous systematic reviews and meta-analyses that usually examined only one or two domains of EFs (Donnelly et al., 2016, 2019; Grundy & Timmer, 2017; Hilchey & Klein, 2011; Hilchey et al., 2015; Zhou & Krott, 2016), we decided to report results for a wide range of commonly addressed domains and tasks. To this end, we have adopted the “Unity and Diversity Model” of EFs (Miyake et al., 2000), which is one of the most frequently cited models of EFs and suggests three domains: inhibition, shifting and updating. In addition, it would not be expected to observe similar effects in different tasks, even if they measure the same function. For this reason, we aimed to report the results of specific task paradigms used in the studies.
The goal of the present systematic review is to investigate whether a bilingual advantage exists among bilingual older adults compared to their monolingual peers on EF tasks that measure inhibition, shifting and updating. The study aims to address the following questions.

1. Are there any differences between bilingual and monolingual older adults in EF performance?
2. In which EF domain exactly can a bilingual advantage be observed?
3. Are possible advantages specific to particular task paradigms?
4. How do the effects of bilingualism vary across different variables, such as language proficiency, testing language, AoA, country, SES and immigration background?

**Method**

This review is registered in the International Prospective Register of Systematic Reviews PROSPERO network (https://www.crd.york.ac.uk/prospero/), registration no. CRD42018113093, and is described according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (Moher et al., 2009). The PRISMA checklist is available in Additional file 1.

**Search strategy**

The following databases were periodically searched up to 21 January 2020: PubMed, PsycINFO and Web of Science. The search strategy was developed and run by a librarian with expertise in systematic reviews. No date criteria were imposed on the search. The search was based on a predetermined series of keywords that are related as follows: bilingual, monolingual, executive function, cognitive control, set-shifting, task-switching, inhibition, cognitive flexibility, switching, shifting and response inhibition. Furthermore, references including reviews or relevant original works were scanned for an additional bibliography.

**Selection criteria**

The following inclusion criteria were implemented: to be included, a study should have quantitatively compared healthy bilingual and monolingual older adults aged 60 years and above with respect to cognitive outcomes of inhibition, shifting or updating. Studies that included different age groups were only included if the behavioural results or at least the data from older adult groups were reported separately. For practical reasons, only published journal articles in English were included. Studies that included multilingual participants and individuals with psychiatric or neurological disorders were excluded. Also, studies that did not compare monolingual to bilingual participants or that compared two monolingual groups with one bilingual group were excluded.

**Screening strategy**

All extracted articles were imported into the Covidence tool, and duplicates were automatically removed. Titles and abstracts of potentially relevant articles were screened by two independent reviewers. Full-text copies of articles were then obtained for those meeting the initial screening criteria. If an article was included by one reviewer and not the other, the article was obtained for further review. Two independent reviewers screened all full-text articles. For discrepancies between the reviewers, eligibility was decided based on discussions among the reviewers or by a third reviewer if it was required.
Quality assessment

Study quality was assessed by two independent reviewers using a modified version of the “QualSyst” tool (Kmet et al., 2004). Each study was assessed based on 11 criteria according to the degree to which they met each criterion (yes = 2, partial = 1, no = 0). Final ratings were calculated based on the mean of independent review scores among the reviewers. All studies received a quality score of 0.6 or higher, indicating moderate to good quality of the studies. Two studies met all the rating criteria and thus received a score of 1.

Data extraction

Data extraction was conducted in March and April 2019. Information was extracted for reference, country, sample, mean age, language of monolinguals, first language (L1) and L2 of bilinguals, language measures, dementia screenings, cognitive measures, controlled confounders, the quality score and outcomes. All extracted data were checked for correctness and completeness by the second reviewer. For studies that included both younger and older adult participants, where no significant Age × Language interaction was reported, follow-up t-test analyses were conducted to determine any significant group difference between older bilinguals and monolinguals. A fourth reviewer assigned each cognitive outcome to a cognitive domain of EFs.

Results

Search results

The flow diagram for the inclusion of studies is presented in Figure 1. A total of 9357 records were identified through the search of the three databases. An additional record was identified through the search of reference lists. After removing duplicates, titles and abstracts of 5973 records were screened. From these, 53 records were identified as potentially meeting the inclusion criteria. After a detailed assessment of the full-text articles, 22 articles published in 2004 or later were found to meet the inclusion criteria. Reasons for excluding articles included outcomes (n = 7), population (n = 19) and study design (n = 5).

Characteristics of included studies

Participant characteristics. Study characteristics are presented in Table 1. The 24 studies involved a total of 1130 monolingual and bilingual older adults, ranging from 20 to 121 participants from 10 different countries, namely Canada, France, Germany, India, the Netherlands, Portugal, Singapore, Spain, the UK and the USA.

Age of the participants ranged from 64.5 to 80.9 years. Except for three studies (Bialystok et al., 2014, Studies 1 and 2; Luo et al., 2013), all studies reported that language groups were comparable in terms of age.

Languages. The most commonly spoken language among monolinguals and bilinguals either as a L1 or L2 was English (n = 19). Other commonly spoken languages were, for example, Basque, French, Spanish, Tamil or Welsh. Except for one study (Kirk et al., 2014), monolinguals always spoke the same language. One study did not report the language of monolingual participants (Zunini et al., 2019). Regarding the languages of the bilinguals, only in six studies did bilingual participants speak the same languages as their L1 and L2 (Ansaldo et al., 2015;
Bialystok et al., 2004, Study 1; Billig & Scholl, 2011; de Bruin, Bak, et al., 2015; Kousaie & Phillips, 2012; Sundaray et al., 2018), while in the rest of the studies bilinguals differed in their L1s or L2s. Among these 18 studies, nine did not indicate which of the languages was the L1 or L2 of the bilinguals (Anderson et al., 2017; Anton et al., 2016; Bialystok et al., 2004, Study 2; Grady et al., 2015; Houtzager et al., 2017; Kirk et al., 2014; Kousaie & Phillips, 2017; Sullivan et al., 2016; Zunini et al., 2019). Moreover, in 14 studies monolingual and bilingual participants differed in their L1 (Bialystok et al., 2004, Study 1; Bialystok et al., 2006, Study 2; Bialystok et al., 2008, 2014, Studies 1 and 2; Billig & Scholl, 2011; Blumenfeld et al., 2016; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Luo et al., 2013; Massa et al., 2020; Schroeder & Marian, 2012; Sundaray et al., 2018). Overall, only two studies clearly stated that the L1 was the same for all monolinguals and bilinguals included (Ansaldo et al., 2015; Kousaie & Phillips, 2012). Moreover, the testing language was only reported in three studies (Anderson et al., 2017; Ansaldo et al., 2015; Anton et al., 2016). In another study, although not reported by the research team, the testing language could be estimated (Kousaie & Phillips, 2012).
Table 1. Study characteristics.

| Reference                  | Country       | Sample | Mean age in years | Language of ML | L1, L2(s) of BL | Language measures | Dementia screening | Cognitive measures | Controlled confounders | Quality score |
|----------------------------|---------------|--------|-------------------|----------------|----------------|-------------------|--------------------|---------------------|----------------------|---------------|
| Anderson et al. (2017)     | Canada        | 35 BL-ML | ML = 74.9 | English | English, Mixed | LSBQ              | MMSE               | D-KEFS TMT subtest, Stroop test | Immigration status | 0.79 |
| Ansaldo et al. (2015)      | Canada        | 10 ML  | ML = 74.5 | French | French, English | LEAP-Q, Part C of the Bilingual Aphasia Test | MoCA               | Digit span test, Simon task, Stroop test, TMT | AoA, immigration status | 0.68 |
| Anton et al. (2016)        | Spain         | 24 ML  | ML = 68.7 | Spanish | Basque, Spanish | Individual interview | MMSE               | Verbal/numerical Stroop test | AoA, immigration status | 1.00 |
| Bialystok et al. (2004), Studies 1 and 2 | Canada (ML), India (BL) | 10 ML  | ML = 71.6 | English | Tamil, English | Language Background Questionnaire, PPVT | N/A                | Simon task | AoA, SES | 0.71 |
| Bialystok et al. (2006), Study 2 | Canada | 24 BL  | ML = 66.9 | English | Mixed, English | Language Background Questionnaire | N/A                | Modified anti-saccade task | AoA | 0.78 |
| Bialystok et al. (2008)    | Canada        | 24 ML  | ML = 67.2 | English | Mixed, English | BNT, self-reported proficiency and usage, PPVT | N/A                | Forward and backward Corsi blocks test, Self-ordered pointing task, Simon arrows task, Stroop test | AoA, immigration status | 0.81 |
| Bialystok et al. (2014), Studies 1 and 2 | Canada | 25 ML  | ML = 71.3 | English | Mixed, English | LSBQ, SVT | N/A | Recent probe task | AoA | 0.8 |
| Billig & Scholl (2011)     | Portugal      | 21 ML  | ML = 65.2 | Brazilian | Brazilian | Language and Health Background Questionnaire | MMSE               | Simon arrows task, Simon squares task, Stroop test | AoA | 0.63 |
| Blumenfeld et al. (2016)   | USA           | 15 ML  | ML = 70.3 | English | Spanish, English | LEAP-Q, PPVT | N/A | Digit span forward task, Non-linguistic Stroop test | AoA | 0.88 |
| Clare et al. (2016)        | UK            | 49 ML  | ML = 72.5 | English | Welsh, English | BPVS 3, Language Questionnaire – short version, BNT, Prawf | MMSE               | D-KEFS Design Fluency and TMT subtests, Go-No Go task, Keep Track task, Simon task, Stroop test, TEA Visual Elevator and Elevator Counting with Distraction subtests, WMS Spatial span test | AoA, immigration status, SES | 0.88 |
| de Bruin, Bak, et al. (2015) | UK          | 24 ML  | ML = 70.2 | English | Gaelic, English | Picture-word matching task, self-report questionnaire | ACE-III | Simon arrows task, task-switching paradigm | AoA, immigration status, SES | 1.00 |
| Reference               | Country          | Sample | Mean age in years | Language of ML | LI, L2(s) of BL | Language measures                                      | Dementia screening            | Cognitive measures                                    | Controlled confounders | Quality score |
|------------------------|------------------|--------|-------------------|----------------|----------------|-------------------------------------------------------|------------------------------|--------------------------------------------------------|------------------------|--------------|
| Grady et al. (2015)    | Canada           | 14 ML  | ML = 70.4, BL = 70.3 | English, Mixed | English, Mixed | SVT, self-reported usage                              | Short MMSE, Simon task       | N/A, Corsi blocks test, task-switching paradigm       | AoA, SES               | 0.77         |
| "Houtzager et al. (2017) | Germany (ML)   | 25 ML  | ML = 73.7, BL = 73.6 | German, Dutch, Frisian | Language background questionnaire, self-assessment report | N/A                          | MoCA, Stroop test                                       | AoA, immigration status, SES | 0.9          |
| Kirk et al. (2014)     | UK               | 32 ML  | ML = 69.5, BL = 71.2 | SSE, Anglo-English | Gaelic, SSE, English, Mixed | Anamnestic judgement task, self-reported proficiency and usage, Corsi blocks test, self-assessment report | AoA, SES                   | 0.73         |
| Kousie & Phillips (2012)| Canada          | 25 ML  | ML = 68.9, BL = 71.9 | Not reported | English, French | Anamnestic judgement task, self-reported proficiency and usage | MoCA, Stroop test, Simon task | AoA, immigration status, SES | 0.93        |
| Kousie & Phillips (2017)| Canada          | 21 ML  | ML = 71.7, BL = 68.7 | English, French | Anamnestic judgement task, self-reported proficiency and usage | MoCA, Stroop test, Task-switching paradigm | AoA, immigration status, SES | 0.97        |
| Luo et al. (2013)      | Canada           | 61 ML  | ML = 71.7, BL = 67.8 | Mixed, English | English, French | Self-report questionnaire, SVT                         | N/A                          | Corsi blocks test, Word and alpha span task           | AoA                    | 0.88         |
| "Massa et al. (2020)   | France           | 16 ML  | ML = 71.1, BL = 72.3 | French, Italian | French, Italian | Self-reported proficiency and usage                   | MMSE, Anti-saccade task, BCST, Stop-signal task, Stroop test, TMT | AoA                    | 0.81         |
| Schroeder & Marian (2012)| USA             | 18 ML  | ML = 80.9, BL = 80.6 | English, Mixed | English, Mixed | LEAP-Q, PPVT                                            | N/A                          | Simon task                                            | AoA                    | 0.88         |
| Sullivan et al. (2016) | Canada           | 23 ML  | ML = 70.4, BL = 71.9 | English, Mixed | English, Mixed | LSBRQ                                                 | N/A                          | Modified Flanker task, recent probe task, Star counting task | AoA                    | 0.84         |
| Sundaray et al. (2018) | UK (ML) Singapore | 20 ML  | ML = 69.9, BL = 67.0 | Tamil, English | Language History and Use Questionnaire, The Complex Ideational Materials Subtest of the BDAE | N/A                          | Stroop arrows task, WMS-R Digit span subtest           | AoA                    | 0.86         |
| "Zunini et al. (2019)  | Canada           | 18 ML  | ML = 71.7, BL = 71.3 | Not reported | English, French | Self-reported proficiency                              | MoCA, Task-switching paradigm, WAIS-III Letter Number Sequencing and Digit span subtest, WCST | N/A                    | 0.93         |

ACE-III: Addenbrooke's Cognitive Examination-III; AoA: age of acquisition; BCST: Berg's Card Sorting Test; BL: bilinguals; BDAE: Boston Diagnostic Aphasia Examination; BNT: Boston Naming Test; BPVS 3: British Picture Vocabulary Scale 3rd edition; D-KEFS: Delis–Kaplan Executive Function System; LI: first language; L2(s): second language(s); LEAP-Q: Language Experience and Proficiency Questionnaire; LSBQ: Language and Social Background Questionnaire; ML: monolinguals; MMSE: Mini Mental State Examination; MoCA: Montreal Cognitive Assessment; N/A: not applicable; PPVT: Peabody Picture Vocabulary Test; SES: socioeconomic status; SSE: Standard Scottish English; SVT: Shipley Vocabulary Test; TEA: Test of Everyday Attention; TMT: Trail Making Test; WAIS: Weschler Adult Intelligence Scale; WCST: Wisconsin Card Sorting Test; WMS(-R): Weschler Memory Scale(-Revised).

*Studies that include both younger and older participant groups. Here, only older adult participants’ characteristics were extracted.

*Allocation of LI and L2 not indicated.
Language measures. In every study, there was an assessment to determine subjective proficiency and usage of the L1 and L2. Nine studies conducted a standardized language background questionnaire: the Language and Social Background Questionnaire (LSBQ; Luk & Bialystok, 2013) (Anderson et al., 2017; Bialystok et al., 2014, Studies 1 and 2; Sullivan et al., 2016) and the Language Experience and Proficiency Questionnaire (LEAP-Q) (Ansaldo et al., 2015; Blumenfeld et al., 2016; Kirk et al., 2014; Schroeder & Marian, 2012), whereas another study used the Language Questionnaire – short version (LQ-SV; Gathercole & Thomas, 2009) (Clare et al., 2016). Eight studies conducted a self-developed or adapted language background questionnaire (Bialystok et al., 2004, Studies 1 and 2; Bialystok et al., 2006, Study 2; Billig & Scholl, 2011; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Luo et al., 2013; Sundaray et al., 2018). Moreover, except for nine studies (Anderson et al., 2017; Anton et al., 2016; Bialystok et al., 2006, Study 2; Billig & Scholl, 2011; Houtzager et al., 2017; Kirk et al., 2014; Massa et al., 2020; Sullivan et al., 2016; Zunini et al., 2019), the rest of the studies also assessed language abilities by conducting standardized verbal tests.

Dementia screening. Only 12 out of the 24 studies conducted a cognitive screening by using a dementia screening tool. Seven studies conducted the Mini-Mental State Examination (MMSE; Folstein et al., 1975) (Anderson et al., 2017; Anton et al., 2016; Billig & Scholl, 2011; Clare et al., 2016; Grady et al., 2015; Massa et al., 2020; Sundaray et al., 2018), four studies conducted the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) (Ansaldo et al., 2015; Kousaie & Phillips, 2012, 2017; Zunini et al., 2019) and one study conducted the Addenbrooke’s Cognitive Examination-III (ACE-III; Hodges & Larner, 2017) (de Bruin, Bak, et al., 2015) as a cognitive screening tool.

Controlled confounders. Except for three studies (Anderson et al., 2017; Kirk et al., 2014; Zunini et al., 2019), all studies reported AoA of the participants. Since distinguishing between early and late bilingualism is quite challenging, it is difficult to generalize studies. In light of the literature, we grouped bilinguals according to their AoA. Individuals exposed to two languages in the first seven years of life can be referred to as early bilinguals. Others who are exposed to their L2 after seven years of age can be referred to as late bilinguals (Berens et al., 2013; Kovelman et al., 2008). Six studies included early bilingual participants (Bialystok et al., 2004, Studies 1 and 2; Billig & Scholl, 2011; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al. 2017), three studies included late bilingual participants (Ansaldo et al., 2015; Blumenfeld et al., 2016; Grady et al., 2015) and the rest had a mixture of early and late bilingual participants (Anton et al., 2016; Bialystok et al., 2006, Study 2; Bialystok et al., 2008, 2014, Studies 1 and 2; Kousaie & Phillips, 2012, 2017; Luo et al., 2013; Massa et al., 2020; Schroeder & Marian, 2012; Sullivan et al., 2016; Sundaray et al., 2018).

Nine studies reported immigration status (Anderson et al., 2017; Ansaldo et al., 2015; Anton et al., 2016; Bialystok et al., 2008; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Kirk et al., 2014; Kousaie & Phillips, 2012) and six studies from five articles reported the SES of participants (Bialystok et al., 2004, Studies 1 and 2; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Kirk et al., 2014). Moreover, except for three studies (Bialystok et al., 2014, Studies 1 and 2; Kirk et al., 2014), all other studies reported that participants were equivalent in terms of educational level.

The role of bilingualism in executive functions

Task paradigms are coded and grouped into three domains and report the general outcome in terms of whether a bilingual advantage is observed in studies. Domains and the number of tasks
conducted in each domain are presented in Table 2. Eleven studies conducted cognitive tests of a single domain of EFs, such as inhibition (Anton et al., 2016; Bialystok et al., 2004, Study 1; Bialystok et al., 2006, Study 2; Bialystok et al., 2014, Study 1; Grady et al., 2015; Kirk et al., 2014; Kousaie & Phillips, 2012, 2017; Schroeder & Marian, 2012) or updating (Bialystok et al., 2014, Study 2; Luo et al., 2013). Eleven studies measured two domains in different combinations, such as inhibition and shifting (Anderson et al., 2017; de Bruin, Bak, et al., 2015; Billig & Scholl, 2011; Blumenfeld et al., 2016; Sullivan et al., 2016; Sundaray et al., 2018) or shifting and updating (Houtzager et al., 2017; Zunini et al., 2019). Only two studies measured all included domains of EFs (Ansaldo et al., 2015; Clare et al., 2016). These two studies did not find a clear overall bilingual advantage in EFs.

In general, out of the 24 studies, nine studies fully supported the bilingual advantage in inhibition (Bialystok et al., 2004, Studies 1 and 2; Bialystok et al., 2006, Study 2; Bialystok et al., 2008,
2014, Study 1; Schroeder & Marian, 2012), shifting (Ansaldo et al., 2015; Houtzager et al., 2017) and updating domains (Bialystok et al., 2014, Study 2). Three studies showed partial support for the bilingual advantage and reported not only a bilingual advantage but also no positive effect of bilingualism on various tasks in inhibition (Ansaldo et al., 2015; Kousaie & Phillips, 2017) and shifting (Zunini et al., 2019). Two studies showed mixed results, such as a bilingual advantage, disadvantage or no positive effect of bilingualism (Luo et al., 2013; Massa et al., 2020). The rest of the studies challenged the existence of a bilingual advantage, as neither positive effects of bilingualism nor bilingual disadvantages were seen in various domains. In the following section, we will outline the chosen measures, grouped by the EF domain. As expected, none of the studies reported a bilingual advantage or disadvantage in all domains.

**Inhibition.** The results of the inhibition domain are presented in Table 3. Twenty studies conducted 31 tasks to measure inhibition. Eleven tasks from nine studies showed a bilingual advantage in this domain. A bilingual advantage was reported in the following EF tests: the Stroop test, Simon task, Simon arrows task, Flanker task and anti-saccade task. One study found a bilingual disadvantage in the inhibition domain by using the Stroop test.

In the Stroop test, which assesses the so-called interference condition, participants are required to name the colour of the word when the word is congruent with the colour (the word “Red” is presented in red) or when there is a mismatch between colour and word (the word “Red” is presented in blue; Stroop, 1935). Regarding the results of the Stroop test, monolingual older adults were slower than bilingual peers on the interference condition and they had higher interference costs (Bialystok et al., 2014, Study 1). They also made fewer errors compared to their monolingual peers (Massa et al., 2020). During the incongruent condition, bilingual older adults made fewer errors (Ansaldo et al., 2015) and they also outperformed monolinguals on RT and accuracy (Kousaie & Phillips, 2017). In another study that showed a bilingual disadvantage, monolinguals outperformed bilinguals in all following measurements of the Stroop test: colour naming, inhibition, inhibition switching, inhibition errors (Anderson et al., 2017).

The Simon task includes congruent (stimulus and response are both located on the same side) and incongruent (stimulus and response are located on opposite sides) stimuli to which participants are required to respond (Simon & Wolf, 1963). Bialystok et al. (2004, Study 1) reported higher accuracy rates for older bilinguals compared to monolingual peers in the incongruent condition of the Simon task. In study 2, accuracy rates were higher for older bilinguals than monolinguals. Again, bilinguals reacted significantly faster in the four-colour congruent condition (Bialystok et al., 2004, Study 2). The Simon effect refers to a greater increase in RT for incongruent trials compared to congruent trials. Bilinguals showed a smaller Simon effect than monolinguals both on the Simon task (Bialystok et al., 2004, Studies 1 and 2; Schroeder & Marian, 2012) and on the Simon arrows task (Bialystok et al., 2008).

In the Flanker task, participants are required to indicate the direction of the target arrow when it is “flanked” by irrelevant stimuli (Eriksen & Eriksen, 1974). Results showed that bilinguals were more accurate than monolinguals on the Flanker task (Kousaie & Phillips, 2017).

In the anti-saccade task, which investigates the voluntary and flexible control of eye movements, participants are required to make saccadic eye movements away from a target. The results of the anti-saccade task showed that there is an overall advantage for bilingual older adults. They reacted significantly faster than older monolinguals in a modified anti-saccade task, especially in conditions involving the greatest degree of conflict (gaze shift and mixed blocked) (Bialystok et al., 2006, Study 2).

**Shifting.** The results of the shifting domain are presented in Table 3. Seven studies conducted 11 tasks to measure the shifting domain, and three tasks from three studies showed a bilingual
Table 3. Overview of the results of the studies.

| Reference                 | Domain                                                                 |
|---------------------------|------------------------------------------------------------------------|
|                          | Inhibition | Shifting | Updating |
| Anderson et al. (2017)    | ML outperformed BL in the following measures of the Stroop test: colour naming, inhibition, inhibition switching and inhibition errors. | There were no significant performance differences between BL and ML in the D-KEFS TMT subtest. | N/A |
| Ansaldo et al. (2015)     | BL made fewer errors than ML on the incongruent condition of the Stroop test. There were no significant performance differences between the two groups on the Simon task. | BL showed a smaller B/A ratio in the TMT than ML. | N/A |
| Anton et al. (2016)       | In study 1, there were no significant performance differences between BL and ML on either the verbal or numerical Stroop tests. | N/A | N/A |
| aBialystok et al. (2004), Studies 1 and 2 | In study 1, accuracy rates were higher for older BL than their ML peers in the incongruent condition of the Simon task. Also, older BL reacted significantly faster in the congruent and incongruent condition and had a significantly smaller Simon effect. In study 2, accuracy rates were higher for older BL. Again, BL reacted significantly faster in the four-colour congruent condition. In both the two- and the four-colour incongruent conditions, BL reacted significantly faster and had a significantly smaller Simon effect. | N/A | N/A |
| aBialystok et al. (2006), Study 2 | In study 2, older BL reacted significantly faster than older ML in a modified anti-saccade task using keypress responses, especially in conditions involving the greatest degree of conflict (gaze shift and mixed blocked). | N/A | N/A |
| aBialystok et al. (2008)  | Older ML showed a larger Simon effect on the Simon arrows task than older BL. Older BL had larger facilitation effects than their ML peers in the Stroop test. | N/A | N/A |
| aBialystok et al. (2014), Studies 1 and 2 | In study 1, older ML were slower than their BL peers on the interference condition of the Stroop test and had higher interference costs. | N/A | N/A |
| Billig and Scholl (2011)  | There were no significant performance differences between older BL and ML in the Simon arrows task and Stroop test. | N/A | N/A |
| Blumenfeld et al. (2016)  | There were no significant performance differences between older BL and ML on a non-linguistic Stroop test. | N/A | N/A |
| Clare et al. (2016)       | There were no significant performance differences between BL and ML in the Simon task, Stroop test, TEA Elevator Counting with Distraction subtest or the Go-No Go task. | ML performed significantly better in the switching condition of the D-KEFS Design Fluency subtest. There were no significant performance differences between BL and ML in the D-KEFS TMT subtest and the TEA Visual Elevator subtest. | N/A |

(Continued)
Table 3. (Continued)

| Reference                  | Domain                        | Inhibition                                                                 | Shifting                                                                 | Updating                                                                 |
|---------------------------|-------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------|--------------------------------------------------------------------------|
| de Bruin, Bak, et al. (2015) | There were no significant performance differences between BL and ML in the Simon arrows task. | There were no significant performance differences between BL and ML in the Simon arrows task. | N/A                                                                       | N/A                                                                       |
| Grady et al. (2015)       | There were no significant performance differences between BL and ML in the Simon task. | N/A                                                                       | N/A                                                                       | N/A                                                                       |
| "Houtzager et al. (2017)  | N/A                           | Older BL showed significantly lower switching costs than their ML peers in a task-switching paradigm. | There were no significant performance differences between older BL and ML in the Corsi blocks test. | N/A                                                                       |
| Kirk et al. (2014)        | There were no significant performance differences between BL and ML in the Simon task. | N/A                                                                       | N/A                                                                       | N/A                                                                       |
| "Kousaie and Phillips (2012) | There were no significant performance differences between older BL and ML in the Stroop test. | N/A                                                                       | N/A                                                                       | N/A                                                                       |
| Kousaie and Phillips (2017) | BL outperformed ML on RT and accuracy measures of incongruent trials of the Stroop test. BL were also more accurate in the Flanker task. There were no significant performance differences between the two groups in the Simon task. | N/A                                                                       | N/A                                                                       | N/A                                                                       |
| "Luo et al. (2013)        | N/A                           | Older ML outperformed BL in an Alpha span task. Older BL outperformed older ML in both forward and backward conditions of the Corsi blocks test. There were no significant performance differences between the two groups in a Word span task. | N/A                                                                       | N/A                                                                       |
| "Massa et al. (2020)      | Older BL made fewer errors than their ML peers in the Stroop test. There were no significant performance differences between the two groups in the stop-signal task and an anti-saccade task. | Older ML showed a smaller B−A difference than their BL peers in the TMT. There were no significant performance differences between the two groups in the BCST. | N/A                                                                       | N/A                                                                       |
| Schroeder and Marian (2012) | BL exhibited a smaller Simon effect in the Simon task than ML. | N/A                                                                       | N/A                                                                       | N/A                                                                       |
| Sullivan et al. (2016)    | There were no significant performance differences in RT between older BL and ML in a modified Flanker task. | N/A                                                                       | There were no significant performance differences between older BL and ML in a Star counting and a recent probe task. | N/A                                                                       |
| "Sundaray et al. (2018)   | There were no significant performance differences between older BL and ML in the Stroop arrows task. | N/A                                                                       | There were no significant performance differences between older BL and ML in the WMS-R Digit span subtest. | N/A                                                                       |
| "Zanini et al. (2019)     | N/A                           | Older BL showed smaller mixing costs than their ML peers in a task-switching paradigm. There were no significant performance differences between older BL and ML in the WCST. | There were no significant performance differences between older BL and ML in the WAIS-III Letter Number Sequencing and Digit span subtests. | N/A                                                                       |

BCST: Berg’s Card Sorting Test; BL: bilinguals; D-KEFS: Delis–Kaplan Executive Function System; ML: monolinguals; N/A: not applicable; RT: reaction time; TEA: Test of Everyday Attention; TMT: Trail Making Test; WAIS-III: Weschler Adult Intelligence Scale; WCST: Wisconsin Card Sorting Test; WMS(-R): Wechsler Memory Scale(-Revised).

*Studies that include both younger and older participant groups. Here, only older adult participants’ characteristics were extracted.
advantage on this domain by using a task-switching paradigm and the Trail Making Test (TMT). Two tasks from two studies showed a bilingual disadvantage in the shifting domain on the Delis–Kaplan Executive Function System (D-KEFS) Design Fluency subtest and the TMT.

In the task-switching paradigm, according to a regular experimental schedule, the participant switches between two or more tasks. Decreased accuracy and a slower performance might occur when the participant switches from one task to another. This difference in accuracy and performance after switching is known as the switching cost. The difference between the single-task condition and repeat trials in the mixed-task condition is described as the mixing cost. In this task, older bilinguals showed lower switching costs (Houtzager et al., 2017) and lower mixing costs than their monolingual peers (Zunini et al., 2019).

The TMT is a connect-the-dot task that consists of two parts. Part A requires the connection in the sequence of dots labelled by numbers. Part B requires the connection in the sequence of dots labelled by alternating numbers and letters (1–A–2–B–3–C). Studies showed heterogeneous results on the TMT. In one study, monolinguals had a smaller difference between the TMT-A and the TMT-B (B–A) (Massa et al., 2020), indicating an advantage of monolinguals over bilinguals. In another study, bilinguals showed a smaller B/A ratio compared to their monolingual peers on the TMT, suggesting a bilingual advantage (Ansaldo et al., 2015).

The D-KEFS Design Fluency subtest measures problem-solving skills by generating visual patterns, drawing new designs according to certain rules and restrictions and inhibiting previous patterns (Delis et al., 2004). Monolinguals performed significantly better on the switching condition of the D-KEFS Design Fluency subtest with a medium effect size (Clare et al., 2016).

**Updating.** The results of the updating domain are presented in Table 3. Twelve studies conducted 18 tasks to measure the updating domain. Fifteen tasks from nine studies did not show either a bilingual advantage or disadvantage. Two tasks from two studies showed a bilingual advantage by using the recent probe task and the Corsi blocks test. One study showed a bilingual disadvantage on the Wechsler Memory Scale – Third Edition (WMS-III; Wechsler, 1997).

In a recent probe task, a trial begins with central fixation, followed by a small number of items called the target set. After an interval, a recognition probe appears and the participant must indicate whether the probe is one of the items from the target set. A recent probe belongs to the target set from the previous trial. Bilinguals showed a bilingual advantage for letter task Accuracy-Facilitation and also for figure task RT-Interference (Bialystok et al., 2014, Study 1).

The Corsi blocks test consists of a platform with 10 blocks mounted on top of it. The examiner taps block sequences of increasing length that have to be repeated in the same (forward) or reverse (backward) order. Bilinguals outperformed their monolingual peers in both forward and backward conditions of the Corsi blocks test (Luo et al., 2013).

In another study, monolinguals showed a significant advantage in the WMS-III Spatial Span subtest forward condition, which has a similar procedure to the Corsi blocks test (Clare et al., 2016).

**The role of confounders on inconsistencies**

Although AoA is an important factor, no clear information was obtained from the included studies. The outcome remained inconsistent when we compared results of the individual studies that included only early bilinguals (Bialystok et al., 2004, Studies 1 and 2; Billig & Scholl, 2011; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017) and late bilinguals (Ansaldo et al., 2015; Blumenfeld et al., 2016; Grady et al., 2015), as there is no one common direction that both groups follow in terms of a bilingual advantage or disadvantage in any domain. It was only possible to have a clear outcome if we sort the results based on the common tasks conducted in studies.
In the inhibition domain, while early AoA was positively related to the bilingual advantage on the Simon task (Bialystok et al., 2004, Studies 1 and 2), there were no differences in the Stroop test. In the shifting domain, late AoA was positively related to a bilingual advantage on the TMT (Ansaldo et al., 2015).

Our results indicated that some studies did not report immigration status as a variable or did not use the data during analysis as a covariate. Again, the outcome remained inconsistent when we compared results of the studies that included non-immigrant (Ansaldo et al., 2015; Anton et al., 2016; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Kousaie & Phillips, 2012) and mixed samples (Anderson et al., 2017; Bialystok et al., 2008; Kirk et al., 2014), as there is no one common direction that both groups follow regarding a bilingual advantage and disadvantage in any domain.

Four studies tested participants in different countries. As they did not all measure the same domain, we decided to focus only on the most commonly measured domains. Three studies measured the updating domain, and none of the studies found a bilingual advantage (Bialystok et al., 2004, Study 2; Houtzager et al., 2017; Sundaray et al., 2018). Three of these studies also measured inhibition and two of them showed a bilingual advantage (Bialystok et al., 2004, Studies 1 and 2), whereas the other found no difference between the language groups (Sundaray et al., 2018). However, this effect might also be attributable to AoA. While two studies included early bilinguals (Bialystok et al., 2004, Studies 1 and 2), the other one had a mixed group of early–late bilinguals (Sundaray et al., 2018).

Discussion

Despite a significant amount of research conducted, the question of whether bilinguals outperform monolinguals in EFs is still debated. In this systematic review, we comprehensively summarized all peer-reviewed published articles that showed evidence for or against a bilingual advantage in healthy older adults, with a closer look at core EF domains separately and the task paradigms used, as well as at moderator variables. Taken together, our systematic review includes 24 studies from 22 articles. Overall, we found no clear evidence for a general bilingual advantage in EFs in healthy older adults. However, we observed that a bilingual advantage tended to be more pronounced in the inhibition domain, particularly expressed in the Stroop test and Simon task. Less clear results were obtained for the shifting domain, while no clear evidence of a bilingual advantage was found in the updating domain.

Based on our results, the bilingual advantage in EFs seems to be restricted to specific EF domains, and its magnitude and extent are related to both participant characteristics and methodological issues. In the following part, we will take some core variables into account that might be responsible for some of the heterogeneities observed in the results of the individual studies.

It has been proposed that age-related changes lead to difficulties in maintaining a wide range of cognitive functions. As mentioned earlier, it is expected that the bilingual advantage might be more pronounced in older adults because of the bell-shaped curve of EFs. These age-related changes are the main reason to review studies that include older adult participants. All included studies reported that mono- and bilingual participants were equivalent in terms of age, except for three studies (Bialystok et al., 2014, Studies 1 and 2; Luo et al., 2013). In these rare cases, any group differences could be driven by age instead of bilingualism.

Another important factor is the proficiency level in the L2. It is assumed that higher L2 proficiency is associated with stronger cognitive demands and correspondingly enhanced cognitive control than lower language proficiency. As a result, larger advantages such as better interference control would be expected in bilinguals with a higher L2 proficiency. However, in our review, a
large number of studies only conducted self-report and self-developed questionnaires to estimate L2 proficiency. Using self-report questionnaires for evaluations has some limitations, such as memory failures. That may become particularly problematic when self-evaluation is used to determine specific aspects of language experience, such as frequency of language use or rate of language switching, which may be difficult for participants to accurately report. In addition, self-disclosure is susceptible to misreporting. For example, participants might under- or overestimate their language proficiency and thus mislead the researcher in order to be included in a study. Thus, the exclusive use of self-report measures limits the validity of results concerning a bilingual advantage. For this reason, it would be best to additionally assess language proficiency with standardized tests, as already done in some studies.

In the studies that conducted verbal tasks, the testing language was either the L1 or the L2 of the bilingual sample. Tasks performed in the L2 could lead to an unfair comparison between bilingual and monolingual participants in verbal tasks, especially for late bilinguals. To an extent, performance differences of bilingual participants between verbal and non-verbal tasks might not only be caused by task type but also the result of the testing language. For instance, as mentioned before, Grundy and Timmer (2017) observed that the small to moderate bilingual advantage in WM was reduced when bilinguals performed the tasks in their L2. In our review, only three studies reported the testing language (Anderson et al., 2017; Ansaldo et al., 2015; Anton et al., 2016). If we assume that the verbal tasks were conducted in the native languages of the monolingual participants, this would mean that in some studies the bilingual sample was tested in their L2, which might lead to an underestimation of the effect of bilingualism.

Because previous studies reported a bilingual advantage in EFs in both early and late bilinguals, we did not set restricted inclusion criteria for the definition of bilingualism regarding AoA. However, it is assumed that early bilinguals might show a larger advantage than late bilinguals because of their long-standing experience of controlling two languages (Luk, De Sa, et al., 2011). Being early bilingual might be advantageous because of two reasons. Firstly, early bilingual individuals tend to have used the two languages in a more balanced manner than late bilinguals, leading to higher proficiency, which causes enhanced inhibitory abilities. However, some previous meta-analyses did not find any evidence supporting that the early acquisition of a L2 would result in enhanced EF performance (Donnelly, 2016; Lehtonen et al., 2018). In our review, some studies did not report AoA data as a variable or did not use the data during analysis as a covariate. The outcome remained inconsistent when we compared the results of the studies that included only early or late bilinguals.

Besides linguistic aspects, also socio-demographic aspects might influence cognitive outcomes when participants are measured in different countries and environmental settings (Hamers & Blanc, 2000). The sociolinguistic environment of the participant might moderate the effects related to the country in which the study is conducted. It has been suggested that findings may be associated with general attitudes regarding bilingualism in different countries (Bak & Alladi, 2016). In Australia, for example, official policies support cultural diversity, although public opinion towards immigrant groups is divided (Dandy & Pe-Pua, 2010). On the other hand, in the USA, a common model of language assimilation, the three-generation model of language assimilation is observed (Waters & Jiménez, 2005). In contrast to the USA’s unilingual “melting pot” approach, Canada’s multicultural attitude towards immigration encourages immigrants to maintain their native language while acquiring at least one of Canada’s official languages (Esses & Gardner, 1996). Furthermore, French language policy, as in many other European countries, is rooted in the nation-state ideology and opts for linguistic assimilation of immigrants (Archibald, 2002). Thus, socio-demographic aspects may potentially be associated with EFs in studies that compare bilingual and monolingual participants from different cultures. In our review, four studies tested participants in
different countries. While three studies found no bilingual advantage in updating (Bialystok et al., 2004, Study 2; Houtzager et al., 2017; Sundaray et al., 2018), one study showed a bilingual advantage in shifting (Houtzager et al., 2017). Regarding inhibition, two studies found a bilingual advantage (Bialystok et al., 2004, Studies 1 and 2), whereas one study found no difference in inhibition (Sundaray et al., 2018). The mixed results observed in the inhibition domain might partly be attributable to AoA. AoA is influenced, among other things, by the sociolinguistic environment, which is also shaped by the country’s language policy. As mentioned before, early bilingualism might cause more enhanced inhibitory control. While the two studies that showed a bilingual advantage in the inhibition domain included early bilinguals (Bialystok et al., 2004, Studies 1 and 2), the other one that found no difference had a mixed group of early–late bilinguals (Sundaray et al., 2018). Thus, bilingual language policies of countries might have an influence on performance differences of individuals. Since there were only a limited number of studies conducted in different countries, there is not enough evidence to give a clear result.

Previous studies have commonly assessed how certain individual characteristics influence the bilingual influence in EFs (DeKeyser, 2010; Yow & Li, 2015). For instance, it has been suggested that immigration status may confound the link between bilingualism and EFs (Fuller-Thomson & Kuh, 2014). Thus, we examined the possible effect of bilinguals’ immigration status on the bilingual advantage. Our results indicated that some studies did not report immigration status as a variable or did not use the data during analysis as a covariate. Considering studies including only participants without a migration background (Ansaldo et al., 2015; Anton et al., 2016; Clare et al., 2016; de Bruin, Bak, et al., 2015; Houtzager et al., 2017; Kousaie & Phillips, 2012) and studies with mixed samples separately (Anderson et al., 2017; Bialystok et al., 2008; Kirk et al., 2014), still a large heterogeneity in findings remained. Even studies with participants from different immigration backgrounds that conducted the same task reported different results. While one found a bilingual advantage on the Stroop task (Ansaldo et al., 2015), the other one reported a bilingual disadvantage (Anderson et al., 2017).

Since all the studies did not control the same confounding variables, it was quite challenging to investigate the dynamics behind the studies that conducted the same EF tasks and demonstrated heterogeneous results. Critically, it is still not clear whether these effects were moderated by individual characteristics or not. The Stroop test, which is the most conducted task, is a robust example to explain the inconsistency in this dynamic. For example, studies that reported different results included participants from the same immigration background (Ansaldo et al., 2015; Anton et al., 2016; Clare et al., 2016; Kousaie & Phillips, 2012) differed in AoA or vice versa (Anton et al., 2016; Bialystok et al., 2008; Kousaie & Phillips, 2012). Yet again, it was difficult to sort studies regarding controlled variables and making an assumption based on them.

**Limitations and future directions**

The present systematic review has several limitations. Firstly, only published articles in peer-reviewed journals were included in this review. Different from some previous reviews (de Bruin, Treccani, et al., 2015; Lehtonen et al., 2018), unpublished data or conference materials were excluded. Therefore, non-peer-reviewed data was not included to keep the scientific standards of included articles at a high level.

It would have also been interesting to enter the available data into a meta-analysis. However, due to the high heterogeneity of the tasks applied for the single domains and the reduced amount of data available for each of the tasks and outcomes, we refrained from conducting a meta-analysis.

Furthermore, it was claimed that the bilingual advantage hypothesis could result from a publication bias in favour of studies with positive results. Studies that show strong evidence for the
bilingual advantage hypothesis could be more likely to be published than others (de Bruin, Treccani, et al., 2015). However, the evidence for a publication bias remains inconclusive and controversial. While some previous meta-analyses indicated a publication bias (de Bruin, Treccani, et al., 2015; Lehtonen et al., 2018), others found no evidence (Adesope et al., 2010; Donnelly et al., 2016; Grundy & Timmer, 2017).

Moreover, the validity and reliability of the conducted tasks is another limitation of bilingualism research that should be taken into consideration (Barkley, 2012). Both variability in task selection and task versions might lead to some problems related to validity and reliability (Valian, 2015). Also, heterogeneity in the results concerning the existence of a bilingual advantage may be related to the task impurity problem. One EF task does not necessarily measure exclusively the targeted EF domain. To an extent, results reflect non-executive variance and it might also engage other domains (Paap et al., 2016). Because of these task-related problems, developing valid and reliable tasks that are sensitive to the inter-related and inter-dependent structure of EFs is an important factor. Future studies should include tasks that measure different domains of EFs to gain a clear and broad outcome of the role of bilingualism on EFs.

To gain a more clear-cut understanding of the effects of bilingualism, we advocate for the identification and stronger control of confounding factors in future studies. The participant-related linguistic and socio-demographic factors mentioned above, such as AoA, language proficiency and immigration status, might have a significant impact on the results. Not only different bilingual populations but also different language combinations, given their structural or lexical distance, might influence the impact on specific EF domains, such as inhibition. For instance, speaking similar languages should lead to competition, and, in turn, might result in improved inhibitory control. While some studies reported cross-linguistic transfer occurs easily for linguistically close language pairs (i.e. Spanish-English) (Prior & Gollan, 2011; Sörman et al., 2019), others reported the same effect for linguistically distant language pairs (i.e. Cantonese-English) (Bialystok et al., 2005; Wierzbicki, 2013). Identifying potential connections between individual differences of bilingual participants and EFs may help to overcome those problems (Bialystok, 2017; Laine & Lehtonen, 2018).

Conclusion

In summary, our systematic review provides no evidence for a general bilingual advantage in EFs in healthy older adults. However, when focusing on single EF domains, bilingualism seems to be reliably associated with an advantage in inhibition, most pronounced for the Stroop test and Simon task. For the shifting domain, less clear results were obtained, but this domain was also least studied. For the updating domain, no evidence of a bilingual advantage was found at all. Individual participant characteristics and methodological problems in studies may have caused the heterogeneity of the results and might explain some of the inconsistencies between the studies. Therefore, well-designed studies that measure all core domains of EFs with valid and reliable tasks and that consider confounding variables are urgently needed.

Acknowledgements

The authors wish to thank Volker Braun from the Library for the Medical Faculty of Mannheim for the consultation on developing a search strategy. We thank Taisiya Baysalova for her expertise and assistance throughout with language editing and proofreading. We are also indebted to colleagues from Network Aging Research for valuable discussions.

Portions of these findings were presented as a poster at the International Association of Gerontology and Geriatrics European Region Congress 2019, Gothenburg, Sweden.
Author contributions

The title, abstract and full texts of articles were screened by MGD and JAG. For discrepancies between reviewers, eligibility was decided based on discussions among the reviewers or by BT. Quality assessment was completed by MGD and JAG. MGD extracted the data and drafted the manuscript. All extracted data were checked for correctness and completeness by JAG. PM supervised the assignment of each cognitive outcome to a cognitive domain. BT, PM and JAG critically reviewed the manuscript for relevant intellectual content. All authors read and approved the final manuscript. PM and BT share last authorship.

Declaration of conflicting interests

The authors have no conflicts of interest to declare.

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by a scholarship from the Klaus Tschira Foundation GmbH.

ORCID iD

Merve Gul Degirmenci https://orcid.org/0000-0002-3412-6571

Supplemental material

Supplemental material for this article is available online.

References

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.
Anderson, V., Anderson, P. J., Jacobs, R., & Smith, M. S. (2008). Development and assessment of executive function: From preschool to adolescence. In V. Anderson, R. Jacobs, & P. J. Anderson (Eds.), Neuropsychology, neurology, and cognition. Executive functions and the frontal lobes: A lifespan perspective (pp. 123–154). Taylor & Francis.
Anderson, J. A. E., Saleemi, S., & Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. Journal of Neurolinguistics, 43, 17–27.
Ansaldo, A. I., Ghazi–Saidi, L., & Adrover-Roig, D. (2015). Interference control in elderly bilinguals: Appearances can be misleading. Journal of Clinical and Experimental Neuropsychology, 37, 455–470.
Anton, E., Garcia, Y. F., Carreiras, M., & Dunabeitia, J. A. (2016). Does bilingualism shape inhibitory control in the elderly? Journal of Memory and Language, 90, 147–160.
Archibald, J. (2002). Immigrant integration: The ongoing process of reform in France and Québec. In S. J. Baker (Ed.), Language policy: Lessons from global models (pp. 30–58). Monterey Institute.
Bak, T. H., Vega-Mendoza, M., & Sorace, A. (2014). Never too late? An advantage on tests of auditory attention extends to late bilinguals. Frontiers in Psychology, 5, 485.
Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? Annals of Neurology, 75(6), 959–963.
Bak, T. H., & Alladi, S. (2016). Bilingualism, dementia and the tale of many variables: Why we need to move beyond the Western World. Commentary on Lawton et al. (2015) and Fuller-Thomson (2015). Cortex, 74, 315–317.
Barkley, R. A. (2012). Executive functions: What they are, how they work, and why they evolved. Guilford Press.
Berens, M., Kovelman, I., & Petitto, L. A. (2013) Should bilingual children learn reading in two languages at the same time or in sequence? Bilingual Research Journal, 36(1), 35–60.
Bialystok, E. (2001). Bilingualism in development: Language, literacy and cognition. Cambridge University Press.
Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin, 143*(3), 233–262.

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459–64.

Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantel, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage, 24*(1), 40–49.

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.

Bialystok, E., Craik, F. I., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859–873.

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(6), 1341–1354.

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging, 29*(3), 696–705.

Billig, J. D., & Scholl, A. P. (2011). The impact of bilingualism and aging on inhibitory control and working memory. *Organon, 26*(51), 39–52.

Blumenfeld, H. K., Schroeder, S. R., Bobb, S. C., Freeman, M. R., & Marian, V. (2016). Auditory word recognition across the lifespan: Links between linguistic and nonlinguistic inhibitory control in bilinguals and monolinguals. *Linguistic Approaches to Bilingualism, 6*(1–2), 119–146.

Bowles, R. P., & Salthouse, T. A. (2003). Assessing the age-related effects of proactive interference on working memory tasks using the Rasch model. *Psychology and Aging, 18*, 608–615.

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the aging brain. *Nature Reviews Neuroscience, 7*, 30–40.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition, 130*, 278–288.

Cappell, K. A., Gmeindl, L., & Reuter-Lorenz, P. A. (2010). Age differences in prefrontal recruitment during verbal working memory maintenance depend on memory load. *Cortex, 46*, 462–473.

Clare, L., Whitaker, C. J., Martyr, A., Martin-Forbes, P. A., Bastable, A. J. M., Pye, K. L., Quinn, C., Thomas, E. M., Gathercole, V. C. M., & Hindle, J. V. (2016). Executive control in older Welsh monolinguals and bilinguals. *Journal of Cognitive Psychology, 28*(4), 412–426.

Coderre, E. L., & van Heuven, W. J. (2014). Electrophysiological explorations of the bilingual advantage: Evidence from a Stroop task. *PLoS One, 9*(7), 1–15.

de Bruin, A., Bak, T. H., & Della Sala, S. (2015). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. *Journal of Memory and Language, 85*, 15–26.

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

Dandy, J., & Pe-Puía, R. (2010). Attitudes to multiculturalism, immigration and cultural diversity: Comparison of dominant and non-dominant groups in three Australian states. *International Journal of Intercultural Relations, 34*, 34–46.

DeKeyser, R. (2010). Practice for second language learning: Don’t throw out the baby with the bathwater. *International Journal of English Studies, 10*(1), 155–165.

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Letter to the editor: Reliability and validity of the Delis-Kalplan Executive Function System: An update. *Journal of the International Neuropsychological Society, 10*(2), 301–303.

Donnelly, S. (2016). *Re-examining the bilingual advantage on interference control and task-switching tasks: A meta-analysis* [Doctoral dissertation]. CUNY Academic Works.

Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review, 26*, 1122–1147.
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.*, 16, 143–149.

Esses, V. M., & Gardner, R. C. (1996). Multiculturalism in Canada: Context and current status. *Canadian Journal of Behavioural Science*, 28(3), 145–152.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.

Fuller-Thomson, E., & Kuh, D. (2014). The healthy migrant effect may confound the link between bilingualism and delayed onset of Alzheimer’s disease. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 52, 128–130.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., Hernández, M., Costa, A., & Avila, C. (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *Neuroimage*, 53, 1272–1278.

Gathercole, V. C. M., & Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12(2), 213–237.

Goral, M., Campanelli, L., & Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults. *Bilingualism-Language and Cognition*, 18(1), 79–89.

Grady, C. L., Luk, G., Craik, F. I., & Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia*, 66, 170–181.

Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism, 1*, 67–81.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory: A comprehensive meta-analysis. *Second Language Research*, 33(3), 325–340.

Hamers, J. F., & Blanc, M. H. A. (2000). *Bilinguality and bilingualism*. Cambridge University Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 653–675). MIT Press.

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, 625–658.

Hilchey, M. O., Saint-Aubin, J., & Klein, R. M. (2015). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. H. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 586–613). Cambridge University Press.

Hodges, J. R., & Larner, A. J. (2017). Addenbrooke’s cognitive examinations: ACE, ACE-R, ACE-III, ACEapp, and M-ACE. In A. J. Larner (Ed.), *Cognitive screening instruments: a practical approach* (pp. 109–137). Springer.

Houtzager, N., Lowie, W., Sprenger, S., & de Bot, K. (2017). A bilingual advantage in task switching? Age-related differences between German monolinguals and Dutch-Frisian bilinguals. *Bilingualism: Language and Cognition*, 20(1), 69–79.

Kaila, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology*, 26(7), 699–713.

Kapa, L. L., & Colombo, J. (2013). Attentional control in early and later bilingual children. *Cognitive Development*, 28(3), 233–246.

Kirk, N. W., Fiala, L., Scott-Brown, K. C., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Psychology*, 26(6), 640–648.

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. Alberta Heritage Foundation for Medical Research (AHFMR).

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.

Kousaie, S., & Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia*, 94, 23–35.

Kovelman, I., Baker, S. A., & Petitto, L. A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism (Cambridge, England)*, 11(2), 203–223.
Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*, 126–147.

Laine, M., & Lehtonen, M. (2018). Cognitive consequences of bilingualism: where to go from here? *Language, Cognition and Neuroscience, 33*(9), 1205–1212.

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*(4), 394–425.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605–621.

Luk, G., De Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition, 14*(4), 588–595.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2011). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes, 27*(10), 1479–1488.

Luo, L., Craik, F. I., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychology and Aging, 28*(1), 28–34.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*, 29–41.

Massa, E., Kopke, B., & El Yagoubi, R. (2020). Age-related effect on language control and executive control in bilingual and monolingual speakers: Behavioral and electrophysiological evidence. *Neuropsychologia, 138*, 107336.

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: Higher proficiency in a second language is associated with more rapid endogenous disengagement. *Quarterly Journal of Experimental Psychology, 65*(8), 1502–1510.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howarter, A., & Wager, T. D. (2000) The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*, 719–726.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699.

Noort, M. V., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences, 9*(27), 1–30.

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*, 232–258.

Paap, K. R., Johnson, H. A., & Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue? *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 74*, 305–314.

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 313–325.

Persad, C. C., Abeles, N., Zacks, R. T., & Denburg, N. L. (2002). Inhibitory changes after age 60 and their relationship to measures of attention and memory. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 57*(3), 223–232.

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish-English and Mandarin-English bilinguals. *Journal of International Neuropsychological Society, 17*, 682–691.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism Language and Cognition, 13*, 253–262.
Degirmenci et al.

Rodriguez-Fornells, A., de Diego Balaguer, R., & Münte, T. F. (2006). Executive functions in bilingual language processing. *Language Learning, 56*, 133–190.

Schroeder, S. R., & Marian, V. (2012). A bilingual advantage for episodic memory in older adults. *Journal of Cognitive Psychology, 24*(5), 591–601.

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus response correspondence and age. *Ergonomics, 6*(1), 99–105.

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Frontiers in Psychology, 2*, 183.

Sörman, D. E., Hansson, P., & Ljungberg, J. K. (2019). Different features of bilingualism in relation to executive functioning. *Frontiers in Psychology, 10*, 269.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Sullivan, M. D., Prescott, Y., Goldberg, D., & Bialystok, E. (2016). Executive control processes in verbal and nonverbal working memory: The role of aging and bilingualism. *Linguistic Approaches to Bilingualism, 6*(1–2), 147–170.

Sundaray, S., Marinis, T., & Bose, A. (2018). Comprehending non-literal language: Effects of aging and bilingualism. *Frontiers in Psychology, 9*, 2230.

Tao, L., Marzecova, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: The role of age of acquisition. *Frontiers in Psychology, 2*, 123.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3–24.

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition, 137*, 40–46.

Waters, M. C., & Jiménez, T. R. (2005). Assessing immigrant assimilation: New empirical and theoretical challenges. *Annual Review of Sociology, 31*, 105–125.

Wechsler, D. (1997). *Wechsler Memory Scale – (WMS–III) administration and scoring manual* (3rd ed.). The Psychological Corporation.

Wierzbicki, S. (2013). *An investigation into the relationship between bilingualism and attentional networks with a focus on linguistic distance and language proficiency* [Master’s dissertation]. Department of Psychology, University of Witwatersrand, Johannesburg, South Africa.

Wu, Y. J., & Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 30*(22), 7646–7651.

Yang, S., Yang, H., & Lust, B. (2011). Early Childhood Bilingualism Leads to Advances in Executive Attention: Dissociating Culture and Language. *Bilingualism: Language and Cognition, 14*(3), 412–422.

Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experience matter. *Frontiers in Psychology, 6*(164), 1–12.

Zhou, B., & Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychonomic Bulletin & Review, 23*, 1221–1230.

Zunini, R. A. L., Morrison, C., Kousaie, S., & Taler, V. (2019). Task switching and bilingualism in young and older adults: A behavioral and electrophysiological investigation. *Neuropsychologia, 133*, 107186.

**Author biographies**

Merve Gul Degirmenci is a psychologist and a doctoral student at the Network Aging Research, Heidelberg University.

Judith Alina Grossmann is a psychologist and a doctoral student at the Network Aging Research, Heidelberg University.

Patric Meyer is a professor at the Department of Applied Psychology, SRH University Heidelberg.

Birgit Teichmann, Dr. rer. nat., is a biologist (diploma) and a gerontologist (M. Sc.), has held the position of the scientific manager of the Network Aging Research since 2007.