REVIEW ARTICLE

The interpreter advantage in executive functions—A systematic review and meta-analysis

Minxia Hu*, Wenjiao Fan

College of Foreign Languages and Cultures, Sichuan University, China

Abstract: Given the heavy cognitive load inherent in language interpreting, interpreters may develop cognitive advantages from managing frequent switching of linguistic codes and working modes. Based on a systematic review of executive functions of inhibiting, shifting and working memory (WM) updating by Nour et al. (2020) and meta-analysis of working memory by Wen and Dong (2019) and Mellinger and Hanson (2019), this research follows the PICOS framework and the PRISMA guideline to synthesize findings from 98 tasks of 29 original studies from International and Chinese databases with a cut-off date of 1st October, 2020. Substantial evidence for an interpreter advantage in shifting was found, but not for inhibition or updating. The meta-analysis showed 1) a moderate to high effect in shifting (g = 0.68, seven WCST effects; g = -0.32, eight switching cost effects); 2) no effect in inhibiting (g = 0.13, six Stroop effects); 3) mixed effects in WM updating. Subgroup analysis on WM updating revealed significant training effects from within-group comparisons (g = 0.58, five 2-back effects; g = 0.71, two L2 listening span effects), but insignificant difference from between-group comparisons (g = -0.03 , five 2-back effects; g = 0.18, five L2 listening span effects). More reproducible behavioral research with scientific and consistent designs is needed for a clearer understanding of the relationship between interpreting experience and EFs.

Keywords: interpreting experience; interpreter advantage; executive functions; inhibition; shifting; working memory updating

***Corresponding author:** Minxia Hu, College of Foreign Languages and Cultures, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, Sichuan province, China; 1301573016@qq.com

Received: May 20, 2021; **Accepted:** July 16, 2021; **Available online:** September 9, 2021

Citation: Hu M and Fan W (2021) The interpreter advantage in executive functions—A systematic review and meta-analysis. *Forum for Linguistic Studies* 3(1): 131–161. DOI: 10.18063/fls.v3i1.1251

1. Introduction

Interpreting is a concurrent process that involves listening and comprehension of speech segments in the source language (SL), attention and retention of the incoming segments and (re) production of equivalents in the target language (TL) with little time tag (Gerver, 1975; Liu et al., 2004). Therefore, interpreting is effortful (Christoffels et al., 2006; Babcock and Vallesi, 2017) and filled with “problem triggers”, such as those caused by dense information, strong accent, thick terminology and asymmetrical SL-TL structures (Gile, 2009: 161–178). It relies on systematic training of interpreting strategies (Li, 2013; Dong et al., 2019) and use of interpreting technologies to reduce cognitive saturation (Gile, 2008; 2011; Fantinuoli, 2018). Given the high cognitive load from frequent switching of codes and modes (Pöchhacker, 2016; Chen, 2017; Stachowiak-Szymczak, 2019), interpreters may face more competitions for their cognitive resources than general bilinguals or non-bilinguals, giving rise to an advantage in cognitive control (García, 2014).

Executive functions (also EFs; cognitive control, executive control) cover a set of mental capacities to formulate goals, execute plans and monitor performances (Lezak, 1982). These functions include inhibiting irrelevant information, storing and updating information in distraction status, or working memory, switching between mental sets, as well as self-initiation, strategy application, multitasking, planning and monitoring (Gilbert and Burgess, 2008; Diamon, 2013; Friedman and Miyake, 2017;). In addition, executive functioning is effortful and trainable (Diamond, 2013: 154).

Working Memory (WM) is a closely related concept. While EFs are top-down goal-oriented mental capacities to coordinate behavior by keeping information active while restraining interferences (Baddeley, 1996; 2012; Kane and Engle, 2002; Friedman and Miyake, 2017), WM is a limited-capacity system supporting cognitive processes by simultaneously storing and processing information (Kane and Engle, 2002; Conway et al., 2005; Baddeley, 2012; Stachowiak-Szymczak, 2019). The central executive in Baddeley's multi-component model of working memory (Baddeley, 1996) supervises, manages and coordinates slave systems, rather than simply maintains information (Baddeley, 2012), which work similarly to EFs. Nevertheless, most cognitive psychologists consider WM (updating) to be one of the EFs (e.g., Miyake et al., 2000; Diamond, 2013; Dong et al., 2018; Lehtonen et al., 2018).

From perception to articulation and from rendition to correction, language interpreting is a complex operation that triggers the activation, manipulation and inhibition of mental representations (Stachowiak-Szymczak, 2019). Memory systems are needed to store (Long-term memory, LTM; Short-term memory, STM) and process (WM) these mental representations (Pöchhacker, 2016: 113–117). Given the high cognitive load (Seeber, 2011), researchers posit that interpreters may develop transferable advantages on behavioral tasks (e.g. García, 2014; Rosiers et al., 2019).

However, such transferable advantages in executive functions have not been consistently reported over the years. For instance, while the interpreter advantage in inhibition was not found in most studies (e.g. Yudes et al., 2011; Dong and Xie, 2014; Dong and Liu, 2016; Aparicio et al., 2017; Babcock and Vallesi, 2017; Van der Linden et al., 2018), others revealed some interpreter superiority, at least for some interpreter groups and tasks (e.g. Köpke and Nespoulous, 2006; Timarová et al., 2014; Woumans et al., 2015; Henrard and Van Daele, 2017). This is also the case for shifting, with supporting evidence from some researchers (e.g. Yudes et al., 2011; Macnamara and Conway, 2014; Dong and Liu, 2016; Liu and Dong, 2017) and mixed evidence from others (e.g. Babcock and Vallesi, 2017; Zhao and Dong, 2020). While the interpreter advantage in updating was revealed in multiple studies (e.g. Timarová et al., 2014; Morales et al., 2015; Dong and Liu, 2016; Dong et al., 2018), others had null findings (e.g. Liu and Dong, 2017; Van der Linden et al., 2018; Rosiers et al., 2019; Liu and Dong, 2020).

Nour et al. (2020) adopted the unity and diversity model of executive functions proposed in

Miyake et al. (2000) for a systematic review of seventeen studies of the interpreter advantage before December 1, 2016. The framework included shifting or switching between tasks, and mental sets (henceforth "Shifting"), updating and monitoring of working memory representations ("Updating") and inhibition of prepotent responses ("Inhibition"). (Diamond, 2013; Gilbert and Burgess, 2008; Miyake et al., 2000). Nour et al. (2020) found evidence for the interpreter advantage in shifting and updating, but not in inhibition. In the mean times, the meta-analysis by Wen and Dong (2019, cut-off before Oct.30, 2018) and Mellinger and Hanson (2019, cut-off before Dec. 2016) revealed significant effects for the interpreter advantage in STM and WM.

Given that prior findings are far from consistent on the presumed interpreter advantage in EFs, important patterns and moderators may be revealed in the systematic review and meta-analysis. Therefore, this research aims to synthesize existing evidence on the impact of interpreting training and/or experience on EFs, expanding on the systematic review by Nour et al. (2020) and the meta-analysis by Wen and Dong (2019) and Mellinger and Hanson (2019).

2. Methods

2.1. PICOS and PRISMA

This study assumes an interpreter advantage in executive functions due to interpreting training or experience by replicating the only published and latest systematic review by Nour et al. and meta-analysis in working memory by Wen and Dong, and conducting research under the "unity and diversity" model of executive functions by Miyake et al. The research is set within the PICOS (*Participants, Intervention, Controls, Outcome and Study Design*) framework (Higgins and Green, 2008; Liberati et al., 2009), with the *Participants* being (more advanced) interpreters, *Intervention* being interpreting training or experience, *Controls* being non- (or less advanced) interpreters, *Outcome* being an interpreter advantage, and *Study Design* being cross-sectional or longitudinal.

This systematic and meta-analytic review follows the *Preferred Reporting Items for Systematic reviews and Meta-Analysis* (PRISMA). Specifically, this systematic and meta-analytic review aims to answer the following three questions:

(1) Do interpreters exhibit EF advantages over non-interpreters or professional interpreters over novice interpreters? This question will be answered by reviewing cross-sectional correlational or between-group comparisons;

(2) Do interpreters enhance EFs with interpreting training? This question will be answered by reviewing longitudinal studies;

(3) Do interpreters exhibit EF advantages on specific tasks? This question will be answered by meta-analysis of five replicable tasks.

2.2. Search strategy

To be as inclusive as possible, both published peer-reviewed studies and unpublished data in grey literature are hand searched in domestic and foreign electronic database (Google Scholar, ResearchGate, ScienceDirect, CNKI, Wanfang Data and Baidu Scholar) with subject heading and key words, i.e., "interpret (er) (ing) experience (training)", "interpreter advantage", "(working) memory", "executive function (s) (ing)", "cognitive control". These keywords are combined using Boolean operators, mainly AND because the operator AND narrows the scopes of search with all concepts searched together (Atkinson and Cipriani, 2018). Besides, the present study also scans bibliographies or references, and conducts backward and forward searches (Card, 2012: 42–52).

2.3. Inclusion criteria

The present systematic review specifies eligibility criteria as follows:

(1) Data Information: Study included must be empirical with statistical analysis.

(2) Study Design: Published and unpublished original articles, including doctoral and master dissertations, both cross-sectional or longitudinal designs.

(3) Sample Characteristics: At least one group of professional interpreters or interpreter trainees should be compared with controls, and at least one EF task be contained; Language of the study should be either English or Chinese; Study included must take interpreting training or experience as the intervention.

(4) Task Inclusion: An EF task by its nature rather than its label.

(5) Definitions of Constructs of Interest: EF components and interpreting should be clearly defined in the included study.

2.4. Extraction criteria

Those excluded are: 1) duplicates; 2) theoretical research, reviews or articles that are unable to trace full-text; 3) not mentioning the moderating factor, i.e. interpreting training or experience; 4) studies with EF tasks unable to be classified under Miyake et al.'s model or simple span tasks tapping only short-term memory capacity.

2.5. Data collection process

First, studies are collected based on the classification of cross-sectional and longitudinal designs. There are three cross-sectional comparisons: 1) interpreters vs non-interpreters (e.g. balanced or unbalanced bilinguals, monolinguals, multilinguals and translators); 2) professional interpreters vs novices; and 3) advanced trainees vs beginners. In addition, there are also cross-sectional correlational studies where the relationship of interpreting experience and EFs are investigated within the group. Longitudinal studies compare interpreter trainees' performance at the start and end of training.

Second the present study classifies data on tasks for Updating, Shifting or Inhibition based on the "unity and diversity" model of Miyake et al. (2000). Each task is categorized as verbal, number, letter or (visual-) spatial (Dong and Zhong, 2019). Tasks measuring each EF are collected and presented in **Table 1**.

2.6. Data analysis

The systematic review was conducted through synthesizing the T, F, or P values as well as group means, standard deviation (SD), standard error (SE), effect size, eta-squared (η^2) and other statistical measures in the original articles, be it longitudinal or cross-sectional comparative or correlational.

The meta-analysis was performed in the Review Manager software (RevMan 5.4). RevMan was also used to assess bias and check heterogeneity in systematic review.

Hedges' g was computed as a standardized mean difference (SMD) (Higgins and Green, 2008; Borenstein et al., 2011; Card, 2012). With $g = 0.20$ representing a small effect, $g = 0.50$ representing a medium effect, and $g = 0.80$ representing a large effect (Card, 2012). Chi-squared (χ^2, or Chi^2), tau-squared (Tau^2, or τ^2) and I^2 were the statistical indicators for heterogeneity. The larger the I^2 is, the more considerable heterogeneity is detected, ranging from 0% to 100% (Higgins and Green, 2008). In continuous variables, Z represents p-value results. The results of bias risk are presented as risk of bias graphs, and the results of meta-analyses as forest plots.

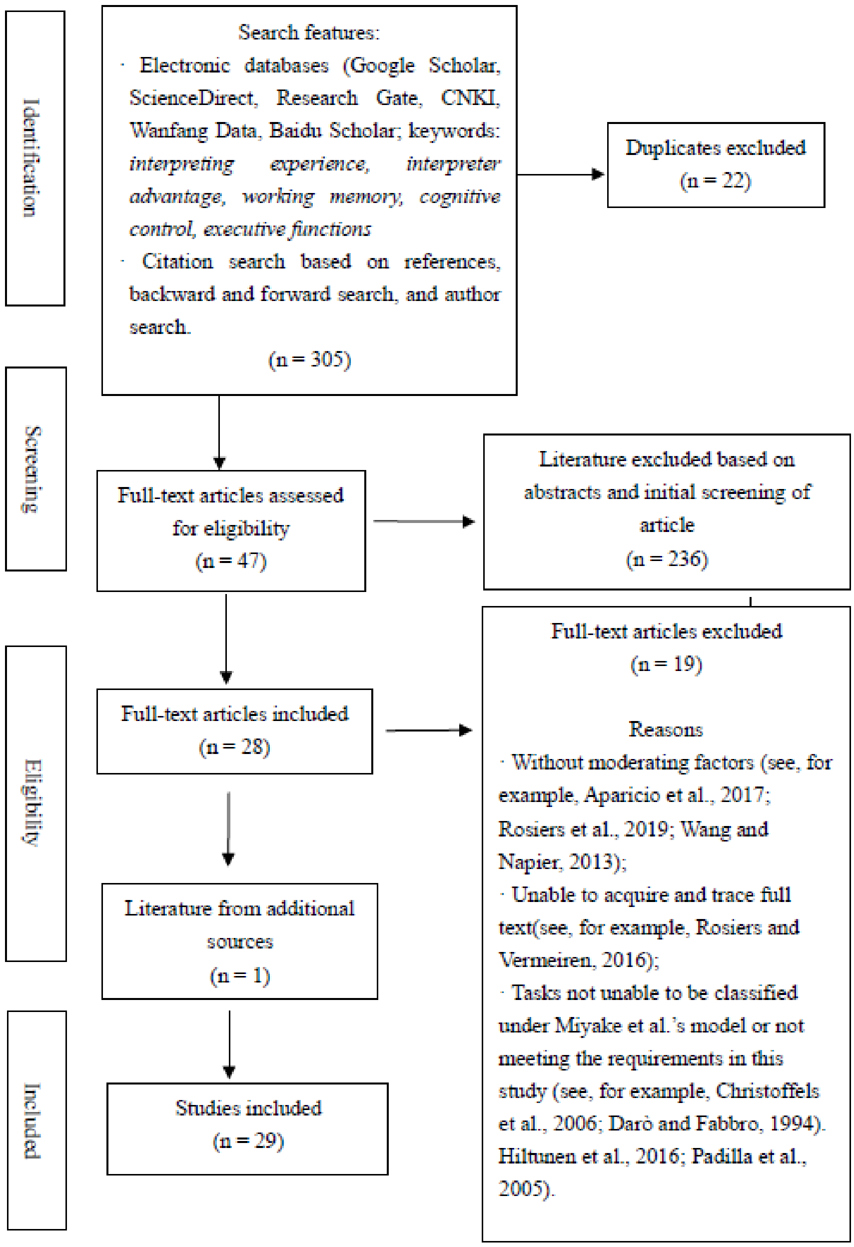


Figure 1. PRISMA flow diagram of selecting studies.

Table 1. EF tasks under the "unity and diversity" framework

EFs	EF Tasks	Task Types	Study ID.
Inhibition	Flanker	Spatial	Dong and Xie, 2014
		Spatial	Van der Linden et al., 2018
	Advanced flanker	Spatial	Van der Linden et al., 2018
	Arrow flanker	Spatial	Timarová et al., 2014
	Simon	Spatial	De Smedt, 2016
		Non-verbal; Spatial	Van der Linden et al., 2018
		Non-verbal; Spatial	Woumans et al., 2015
		Non-verbal; Spatial	Yudes et al., 2011
	ANT	Spatial	Babcock et al., 2017
		Spatial	Babcock and Vallesi, 2017
		Spatial	De Smedt, 2016
		Spatial	Woumans et al., 2015
	ANTI-V	Spatial	Morales et al., 2015
	Number Stroop	Number	Dong and Liu, 2016
		Number	Liu and Dong, 2017
		Number	Zou, 2016
	Color-word Stroop	Verbal	Babcock and Vallesi, 2017
		Verbal	Köpke and Nespoulous, 2006
		Verbal	Tian, 2016
	Antisaccade task	Spatial	Henrard and Van Daele, 2017
		Spatial	Timarová et al., 2014
	Brown-Peterson	Non-verbal	Henrard and Van Daele, 2017
Shifting	Task-switching	Non-verbal; Spatial	Babcock et al., 2017
		Non-verbal; Spatial	Babcock and Vallesi, 2017
		Non-verbal	Zou, 2016
	Number-letter task	Number/letter	Timarová et al., 2014
		Number/letter	Macnamara and Conway, 2014
		Number/letter	Macnamara and Conway, 2015
	Color-shape switch task	Spatial	De Smedt, 2016
		Spatial	Dong and Liu, 2016
		Spatial	Liu, 2018
	WCST	Spatial	Dong and Xie, 2014
		Spatial	Liu and Dong, 2017
		Spatial	Macnamara and Conway, 2014
		Spatial	Liu, 2018
		Spatial	Macnamara et al., 2011
		Non-verbal; Spatial	Wei, 2017

Table 1 (*continued*)

EFs	EF Tasks	Task Types	Study ID.
		Non-verbal; Spatial	Macnamara and Conway, 2015
		Non-verbal; Spatial	Yudes et al., 2011
	Plus-minus	Non-verbal	Henrard and Van Daele, 2017
	Semantic fluency	Verbal	Woumans et al., 2015
Updating	Complex span (listening span)	Verbal	Liu and Dong, 2020
		Verbal	Attanak et al., 2019
		Verbal	Chmiel, 2018
		Verbal	Dong et al., 2018
		Verbal	Liu et al., 2004
		Verbal	Stavrakaki et al., 2012
		Verbal	Tian, 2016
	Complex span (reading span)	Verbal	Chmiel, 2018
		Verbal	Zou, 2016
	Complex span (the automated operation, reading and automated symmetry span)	Spatial	Babcock et al., 2017
		Spatial	Babcock and Vallesi, 2017
		Spatial	Macnamara and Conway, 2014
		Non-verbal	Macnamara et al., 2011
		Non-verbal	Stead and Tripier, 2016
		Non-verbal	Macnamara and Conway, 2015
	Span task with articulatory suppression	Verbal	Injoque-Ricle et al., 2015
	N-back	Spatial	Van der Linden et al., 2018
		Spatial	Timarová et al., 2014
	Single and dual n-back	Spatial	Morales et al., 2015
	Visuo-spatial 2-back	Spatial	Liu and Dong, 2020
		Spatial	De Smedt, 2016
		Spatial	Dong and Liu, 2016
		Spatial	Dong et al., 2018
		Spatial	Dong and Liu, 2017
	Dual n-back	Spatial	Attanak et al., 2019
		Spatial	Stead and Tripier, 2016
	Letter-memory	Letter	Henrard and Van Daele, 2017
		Number	Zou, 2016
	Free call with suppression	Verbal	Köpke and Nespoulous, 2006
	Category/rhyme probe	Verbal	Köpke and Nespoulous, 2006
	Cued recall	Verbal	Signorelli et al., 2012

3. Results of the systematic review

3.1. Data extraction

A total of 305 studies was included based on relevance with the present systematic review, of which 215 was sourced from Google Scholar, 75 from Baidu Scholar and 15 from ScienceDirect. Then, 22 duplicates were removed. After initial screening of the abstracts and full texts, those not meeting the inclusion criteria were excluded, leaving 47 full texts for in-depth comprehensive reading. In the end, a total of 28 studies were included in the abovementioned snowballing procedure, which were conducted and completed in May-August 2020, covering the years from 1980 to 2020. Then, a second round of literature search was conducted with a cut-off date of October 1, 2020. One additional study by Liu and Dong in 2020 was added to the literature, bringing the total number of reviewed studies to 29 (see **Figure 1**).

3.2. Research design of reviewed studies

Among these 29 studies, 10 was longitudinal, 17 cross-sectional and 2 correlational. In addition, 8 additional post-test comparisons from longitudinal studies were extracted between the interpreter group and the controls, as presented in **Table 2**.

Table 2. Research designs of included studies

Number of time points	Group characteristics	Study number
One: Cross-sectional design	Between-group comparisons containing one group of interpreters	19
	Between-group comparisons containing more than one interpreter groups	6
	Correlational analysis	2
Two or more: Longitudinal design		10

3.3. Included EF tasks

Among these 29 studies, a total of 2, 034 subjects participated in 129 reported tasks, from which only 75 were included for analysis under the "unity and diversity" model (Miyake et al., 2000). Tasks were counted more than once when between-group comparison results were provided in longitudinal studies, bringing the total number of tasks to 87, as presented in **Table 3**.

Table 3. Included tasks taxing different EF components

Executive Functions	Tasks Included	Frequency of Use
Response-Distractor	ANT; ANTI-V; Antisaccade; Brown-Peterson; Flanker; Simon; Stroop	26
Shifting	Color-shape switch; Number-letter; Plus-minus; Task-switching (in switching cost); WCST	24
Updating	Category and rhyme probe task; Complex-span (listening span; automated operation, symmetry, or reading span; free call; cued recall); letter-memory; number switch; N-back	48

3.4. Results of the included studies

The present study used an effect below 0.05 in p value or above 0.5 in Cohen's d value as the "advantage" criteria for interpreters over bilinguals, multilinguals, translators or monolinguals, or for interpreters with more experience or training compared to those with less or no training or experience. The authors' analysis and conclusion were also checked to confirm the results.

3.5. Response-distractor inhibition

Cross-sectional or correlational studies investigating interpreters' possible inhibitory advantage were conducted on 21 tasks. Among them, five tasks (24%) exhibited the interpreter advantage, while fourteen tasks (76%) didn't.

Among the five longitudinal tasks included, three (60%) didn't reveal an advantage from interpreting training and experience, while the other two (40%) indicated minimal training effects. Specifically, De Smedt (2016) hinted a minimal improvement on the Simon incongruent trials and significant enhancement on the ANT incongruent trials. Detailed information of the included inhibition tasks is presented in **Table 4**. The histogram in **Figure 2** is a visualized presentation of the included results.

Table 4. Results of the included inhibition tasks

Article	Research design	Reason	Task(s)	Results
Babcock and Vallesi, 2017	Cros. 1 group I.	experience	Stroop	ns.
	Cros. 1 group I.	experience	ANT	ns.
Babcock et al., 2017	Longitudinal IT.	training	ANT	ns.
	Cros. Design 1 group IT.	training	ANT	ns.
De Smedt, 2016	Longitudinal IT.	training	Simon	TE
	Longitudinal IT.	training	ANT	TE
Dong and Liu, 2016	Longitudinal IT.	training	Stroop	ns.
	Cros. Design 1 group IT.	training	Stroop	ns.
Dong and Xie, 2014	Cros. 1 group IT.	training	Flanker	ns.
	Cros. >1 group IT.	training	Flanker	ns.
Dong and Liu, 2017	Longitudinal IT.	training	Stroop	ns.
	Cros. Design 1 group IT.	training	Stroop	ns.
Henrard and Van Daele, 2017	Cros. 1 group I.	experience	Antisaccade	I+
	Cros. 1 group I.	experience	Brown-Peterson	I+
Köpke and Nespoulous, 2006	Cros. >1 group I.	experience	Stroop	ns.
	Cros. >1 group I.	experience	Stroop	ns.
Morales et al., 2015	Cros. 1 group I.	experience	ANTI-V	ns.
Tian, 2016	Cros. 1 group I.	experience	Stroop	ns.
Van der Linden et al., 2018	Cros. 1 group I.	experience	Flanker	ns.
			Simon	ns.
	Cros. 1 group I.	experience	Advanced flanker	ns.
Woumans et al., 2015	Cros. 1 group IT.	training	Simon	I+
	Cros. 1 group IT.	training	ANT	I+
Yudes et al., 2011	Cros. 1 group I.	experience	Simon	ns.
Timarová et al., 2014	Corre. 1 group I.	experience	Flanker	I+
			Antisaccade	ns.

Note: Cros. = Cross-sectional study. Cros. Design = Cross-sectional design in longitudinal study. 1 group I.= There is merely one group of interpreters included in a certain task. >1 group I./IT. = There are more than one group of interpreters or interpreter trainees. 1 group IT. = There is only one group of interpreter trainees included in a task. TE = Training effect due to minimal improvement, but no significant difference is reached (usually in the longitudinal study). I+ = interpreter advantage. ns. = no significant difference or correlation.

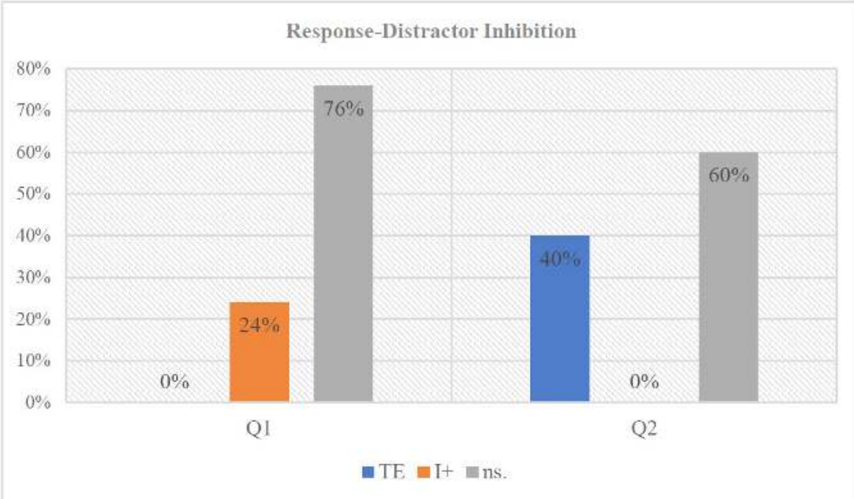


Figure 2. Results of inhibition advantages among interpreters.

Note: Q1: Do interpreters exhibit inhibitory advantages over non-interpreters? Q2: Do interpreters enhance inhibition with more training or accumulated experience?

3.6. Shifting

In the 18 cross-sectional studies on the shifting function, four (22%) reported no significant advantage for interpreters over non-interpreters or significant correlation between interpreting training or experience and shifting response. Two (11%) found only minimal advantages. However, twelve (67%) discovered significant advantages. Among the six longitudinal studies included, all (100%) found supporting evidence for the training or practice effect. See **Table 5** for the detailed results of the shifting tasks included. See a more visual presentation in the bar chart of **Figure 3**.

Table 5. Results of the included shifting tasks

Article	Research design	Reason	Task (s)	Results
Babcock and Vallesi, 2017	Cros. 1 group I.	Experience	Task-switching	ns.
Babcock et al., 2017	Longitudinal IT.	training	Task-switching	TE+
	Cros. Design 1 group IT.	training	Task-switching	ns.
De Smedt, 2016	Longitudinal IT.	training	Color-shape switch	TE
Dong and Liu, 2016	Longitudinal IT.	training	Color-shape switch	TE+
	Cros. Design 1 group IT.	training	Color-shape switch	I+
Dong and Xie, 2014	Cros. 1 group IT.	training	WCST	I+
	Cros. >1 group IT.	training	WCST	MI
Dong and Liu, 2017	Longitudinal IT.	training	WCST	TE+
	Cros. Design 1 group IT.	training	WCST	I+
Henrard and Van Daele, 2017	Cros. 1 group I.	experience	Plus-minus	ns.
	Cros. 1 group I.	experience	Plus-minus	I+
Macnamara and Conway, 2014	Cros. >1 group IT.	training	Task-switching	I+
	Cros. >1 group IT.	training	WCST	I+

Table 5 (*continued*)

Article	Research design	Reason	Task (s)	Results
Liu, 2020	Cros. 1 group IT.	training	WCST	I+
	Cros. 1 group IT.	training	Color-shape switch	I+
Macnamara et al., 2011	Cros. >1 group I.	experience	WCST	MI
Timarová et al., 2014	Corre. 1 group I.	experience	Number-letter	ns.
Wei, 2017	Cros. 1 group IT.	training	WCST	I+
	Cros. > 1 group IT.	training	WCST	I+
Macnamara and Conway, 2015	Longitudinal IT.	training	WCST	TE+
	Longitudinal IT.	training	Task-switching	TE+
Woumans et al., 2015	Cros. 1 group IT.	training	Semantic verbal fluency	I+
Yudes et al., 2011	Cros. 1 group I.	experience	WCST	I+

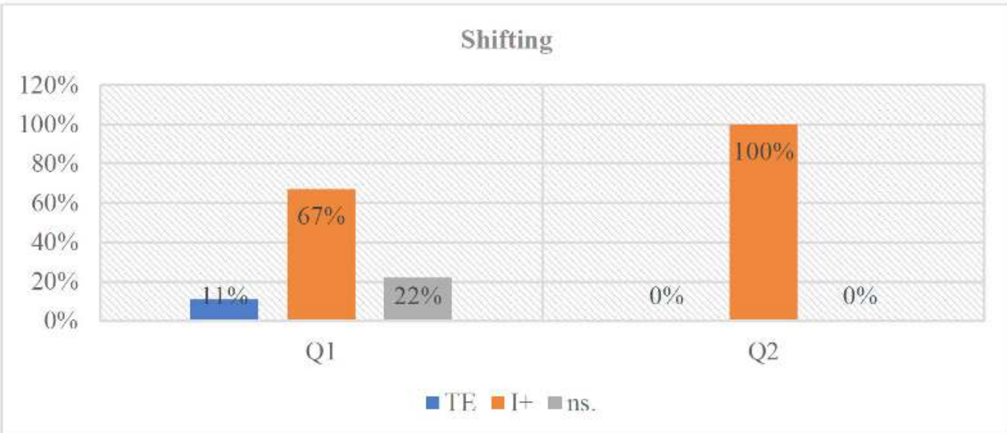


Figure 3. Results of shifting advantages among interpreters.

Note: Q1: Do interpreters exhibit shifting advantages over non-interpreters? Q2: Do interpreters enhance shifting with more training or accumulated experience?

3.7. Updating

On the 34 cross-sectional tasks, 23 (68%) revealed no significant group difference in updating between interpreters and controls or between experts and novices, while 11 (32%) did. On the 14 longitudinal tasks, seven studies (50%) didn't report a significant training effect, while six tasks (43%) did, with one more study (7%) revealing minimal improvement. See **Table 6** for the detailed results of the shifting tasks included. See a more visual presentation in the bar chart of **Figure 4**.

Table 6. Results of the included updating tasks

Article	Research design	Reason	Task (s)	Results
Liu and Dong, 2020	Longitudinal IT.	training	Complex span	ns.
	Cros. Design 1 group IT.	training	Complex span	ns.
	Longitudinal IT.	training	N-back	ns.
	Cros. Design 1 group IT.	training	N-back	ns.

Table 6 (*continued*)

Article	Research design	Reason	Task (s)	Results
Attanak et al., 2019	Longitudinal IT.	training	Complex span	TE+
	Cros. Design 1 group IT.	training	Complex span	I+
	Longitudinal IT.	training	N-back	TE+
	Cros. Design 1 group IT.	training	N-back	I+
Babcock and Vallesi, 2017	Cros. 1 group I.	experience	Complex span	I+
Babcock et al., 2017	Longitudinal IT.	training	Complex span	ns.
	Cros. Design 1 group IT.	training	Complex span	ns.
Chmiel, 2018	Longitudinal IT.	training	Complex span	TE+
	Cros. Design >1 group I.	experience	Complex span	I+
	Cros. Design 1 group I.	experience	Complex span	I+
De Smedt, 2016	Longitudinal IT.	training	N-back	TE+
Dong and Liu, 2016	Longitudinal IT.	training	N-back	TE+
	Cros. Design 1 group IT.	training	N-back	I+
Dong et al., 2018	Longitudinal IT.	training	N-back	TE+
	Longitudinal IT.	training	Complex span	TE
	Cros. Design 1 group IT.	training	N-back	IT+
	Cros. Design 1 group IT.	training	Complex span	ns.
Dong and Liu, 2017	Longitudinal IT.	training	N-back	ns.
	Cros. Design 1 group IT.	training	N-back	ns.
Henrard and Van Daele, 2017	Cros. 1 group I.	experience	Letter-memory	I+
	Cros. 1 group I.	experience	Letter-memory	I+
Macnamara and Conway, 2014	Cros. >1 group IT.	training	Complex span	ns.
Injoque-Ricle et al., 2015	Corre. 1 group I.	experience	Complex span	ns.
	Corre. 1 group I.	experience	Span with articulatory suppression	ns.
Köpke and Nespoulous, 2006	Cros. >1 group I.	experience	Free recall	I+
	Cros. >1 group I.	experience	Category and rhyme probe task	ns.
	Cros. >1 group I.	experience	Free recall	ns.
	Cros. >1 group I.	experience	Category and rhyme probe task	ns.
Liu et al., 2004	Cros. >1 group I.	experience	Complex span	ns.
	Cros. >1 group IT.	training	Complex span	ns.
Macnamara et al., 2011	Cros. >1 group I.	experience	Complex span	ns.
Morales et al., 2015	Cros. 1 group I.	experience	N-back	I+

Table 6 (*continued*)

Article	Research design	Reason	Results	Results
Macnamara and Conway, 2015	Longitudinal IT.	training	Complex span	ns.
Signorelli et al., 2012	Cros. >1 group I.	experience	Complex span	ns.
Stavrakaki et al., 2012	Cros. 1 group I.	experience	Complex span	ns.
Stead and Tripier, 2016	Longitudinal IT.	training	Complex span	ns.
	Cros. Design 1 group IT.	training	Complex span	ns.
	Longitudinal IT.	training	N-back	ns.
	Cros. Design 1 group IT.	training	N-back	ns.
Tian, 2016	Cros. >1 group I.	experience	Complex span	ns.
Timarová et al., 2014	Corre. 1 group I.	experience	N-back	ns.
Van der Linden et al., 2018	Cros. 1 group I.	experience	N-back	ns.
Zou, 2016	Cros. >1 group IT.	training	Complex span	ns.
	Cros. >1 group IT.	training	Number switch	ns.

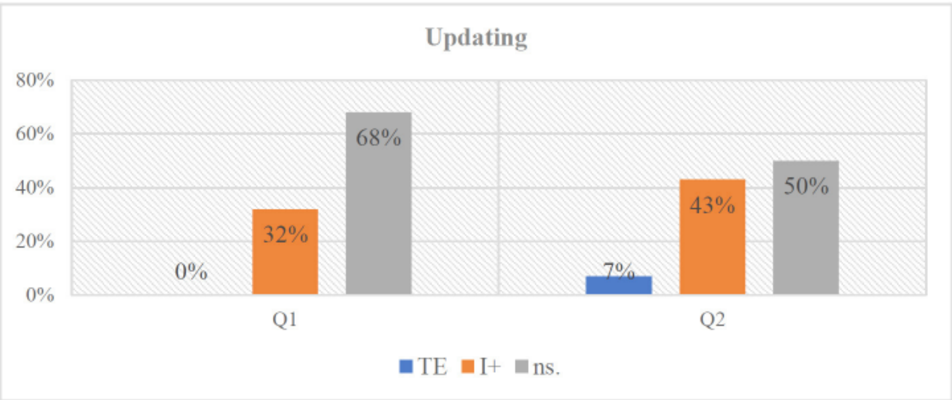


Figure 4. Results of updating advantages among interpreters.

Note: Q1: Do interpreters exhibit updating advantages over non-interpreters? Q2: Do interpreters enhance updating with more training or accumulated experience?

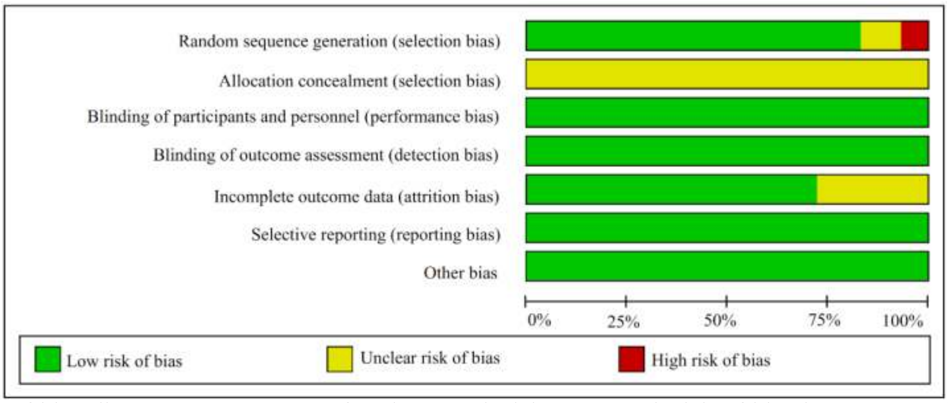


Figure 5. Risk of bias diagram: percentages of reviewers' decisions on each risk of bias item across all the 29 studies included.

Figure 6. Bias risk summary: review authors' assessments on each type of risk of bias for each included study.

Table 7. Detailed information of 29 studies

Study ID.	Author	Country of the author	Department	Project	Journal Titles	Journal Level
Attanak et al., 2019	Attapol Attanak	Thailand	Khon Kaen University Language Institute	/	Journal of Community Development Research (Humanities and Social Sciences)	
	Sirikran Juntapremjit	Thailand	Center of Excellence in Cognitive Science, Burapha University	/		
	Arnon Chaisuriya	Thailand	Language Institute, Chulalongkorn University	/		
Babcock et al., 2017	Laura Babcock	Sweden	Department of Neuroscience	Hemispheric Asymmetries for Executive Functions	Journal of Cognitive Enhancement	
	Mariagrazia Capizzi	Italy	Department of Neuroscience			
	Sandra Arbula	Italy	Cognitive Neuroscience Group, Scuola Internazionale Superiore di Studi Avanzati di Trieste			
	Antonino Vallesi	Italy	Department of Neuroscience			
Babcock and Vallesi, 2017	Laura Babcock	Sweden	Department of Neuroscience,Karolinska Institutet	The neurobiology of musical expertise; Hemispheric Asymmetries for Executive Functions	Bilingualism: Language and Cognition	SSCI
	Antonino Vallesi	Italy	Department of Neuroscience, University of Padova	Hemispheric Asymmetries for Executive Functions		
Chmiel, 2018	Agnieszka Chmiel	Poland	Department of Translation Studies, Adam Mickiewicz University	Respeaking - process, competences, quality; (PINC) Extreme language control: activation and inhibition as bilingual control mechanisms in conference interpreting; ADLAB PRO	International Journal of Bilingualism	A&HCI; SSCI
De Smedt, 2016	Sarah De Smedt	Belgium	University Gent	/	/	
Dong and Liu, 2016	Dong Yanping	China	School of International Studies, Zhejiang University	Neurocognitive Studies of Interpreter Training; Working memory in second language acquisition and processing	Frontiers in Psychology	SSCI
	Liu Yuhua	China	College of Foreign Studies, South China Agricultural University	Bilingual advantage		
Dong et al., 2018	Dong Yanping	China	School of International Studies, Zhejiang University	Neurocognitive Studies of Interpreter Training; Working memory in second language acquisition and processing	Frontiers in Psychology	SSCI
	Liu Yuhua	China	College of Foreign Studies, South China Agricultural University	Bilingual advantage		
	Cai Rendong	China	Bilingual Cognition and Education Lab, Guangdong University of Foreign Studies	The effect of using a foreign language on decision-making; Phonetic convergence in nonnative speech; Working memory in second language acquisition and processing		

Table 7 (*continued*)

Dong and Xie, 2014	Dong Yanping	China	School of International Studies, Zhejiang University	Neurocognitive Studies of Interpreter Training; Working memory in second language acquisition and processing	Journal of Cognitive Psychology	SSCI
	Xie Zhilong	China	Foreign Languages College, Jiangxi Normal University	Different bilingual experiences and cognitive control		
Henrard and Van Daele, 2017	Sébastien Henrard	Belgium	Department of Occupational Psychology, Université de Mons	Cognitive characteristics of work activities	Frontiers in Psychology	SSCI
	Agnès Van Daele	Belgium	Department of Occupational Psychology, Université de Mons	PROJET EXPERT'CRISE		
Injoque-Ricle et al., 2015	Irene Injoque-Ricle	Argentina	Department of Cognitive Processes, Psychology Research Institute, Universidad de Buenos Aires	Procesos cognitivos en niños; Procesos cognitivos en adultos; Memoria de trabajo e interpretación simultánea	Advances in Cognitive Psychology	SSCI
	Juan Pablo Barreyro	Argentina	Faculty of Psychology, UBA			
	Jesica Formoso	Argentina	Department of Research in Basic Cognitive Processes, National Scientific and Technical Research Council			
	Virginia I. Jaichenco	Argentina	Institute of Linguistics, UBA	Memoria de trabajo e interpretación simultánea		
Köpke and Nespoulous, 2006	Barbara Köpke	France	University of Toulouse, Toulouse, France		Interpreting	A&HCI; SSCI
	Jean-Luc Nespoulous	France	University of Toulouse	COGNIPROS. Linguistic and cognitive evaluation of prosodic production and		
Liu and Dong, 2017	Liu Yuhua	China	College of Foreign Studies, South China Agricultural University	Bilingual advantage	Foreign Language Research	Chinese Core Journals
	Dong Yanping	China	School of International Studies, Zhejiang University	Neurocognitive Studies of Interpreter Training; Working memory in second language		
Liu and Dong, 2020	Liu Yuhua	China	College of Foreign Studies, South China Agricultural University	Bilingual advantage	Journal of Foreign Languages	CSSCI
	Dong Yanping	China	School of International Studies, Zhejiang University	Neurocognitive Studies of Interpreter Training; Working memory in second language acquisition and processing		
Liu et al., 2004	Liu Minhua	Hong Kong, China	Centre for Translation, Hong Kong Baptist University	/	Interpreting	A&HCI; SSCI
	Diane L. Schallert	U.S.	Department of Educational Psychology, University of Texas at Austin	Written and oral language; Constructive Criticism Team		
	Patrick J. Carroll	/	/	/		
Liu, 2018	Liu Yuhua	China	College of Foreign Studies, South China Agricultural University	Bilingual advantage	Translation Research and Teaching	
Macnamara and Conway, 2014	Brooke N. Macnamara	U.S.	Department of Psychological Sciences, Case Western Reserve University	/	Psychonomic Bulletin & Review	SSCI
	Andrew R. A. Conway	U.S.	Department of Applied Cognitive Psychology, Claremont Graduate University	Virtues in Communities of Trust		
Macnamara and Conway, 2015	Brooke N. Macnamara	U.S.	Department of Psychological Sciences, Case Western Reserve University	/	Journal of Applied Research in Memory and Cognition	SSCI
	Andrew R. A. Conway	U.S.	Department of Applied Cognitive Psychology, Claremont Graduate University	Virtues in Communities of Trust		

Table 7 (*continued*)

Macnamara et al., 2011	Brooke N. Macnamara	U.S.	Department of Psychological Sciences, Case Western Reserve University	/	Interpreting	A&HCI; SSCI
	Adam B. Moore	/	/	/		
	Judy Kegl	U.S.	Department of Linguistics, University of Southern Maine	ASL Linguistics Project		
	Andrew R. A. Conway	U.S.	Department of Applied Cognitive Psychology, Claremont Graduate University	Virtues in Communities of Trust		
Morales et al., 2015	Julia Morales	Spain	Universidad Loyola Andalucia	Modulating cognition with tDCS	Acta Psychologica	SSCI
	Francisca Padilla	Spain	Department of Experimental Psychology and Physiology of Behaviour, University of Granada	Cognitive processes in translation or interpreting; Working memory and executive functions		
	Carlos J. Gomez-Ariza	Spain	Department of Psychology, Universidad de Jaén	Executive Control Training; Inhibitory control as a mechanism of memory regulation		
	Maria Teresa Bajo	Spain	University of Granada			
Signorelli et al., 2011	Teresa Signorelli Pisano	U.S.	Program in Speech–Language–Hearing Sciences, CUNY Graduate Center	Semantic working memory	International Journal of Bilingualism	A&HCI; SSCI
	Henk J Haarmann	U.S.	Center for Advanced Study of Language (CASL), University of Maryland, College Park	Semantic working memory; Creative cognition		
	Loraine K Obler	U.S.	Program in Speech–Language–Hearing Sciences, CUNY Graduate Center	The influence of second-language grammar on native language sentence comprehension; Semantic working memory		
Stavrakaki et al., 2012	Stavroula Stavrakaki	Greece	School of Italian Language and Literature, Aristotle University of Thessaloniki	Relationship between Developmental Language Disorder and Developmental Dyslexia	Journal of Clinical and Experimental Neuropsychology	SCIE; SSCI
	Kalliopi Megari	Greece	Department of Cardiothoracic Surgery I, Aristotle University of Thessaloniki	Call for papers for "American Journal of Psychiatry and Neuroscience"		
	Mary Helen Kosmidis	Greece	School of Psychology, Aristotle University of Thessaloniki	Neuropsychological Assessment of Greek Australian Migrants		
	Maria Apostolidou	/	/	/		
	Eleni Takou					
Stead and Tripier, 2016	Andrew Stead	Switzerland	University of Geneva	/	/	
	Coralie Triper					
Tian, 2016	Tian Jia	China	Guangdong University of Foreign Studies	/	/	
Timarová et al., 2014	Šárka Timarová	/	KU Leuven	/	Interpreting	A&HCI; SSCI
	Ivana Čeňková	Czechia	Institute of Translation Studies, Charles University in Prague	ImPLI (Improving police and legal interpreting) 2011-2012		
	Reine Meylaerts	Belgium	Faculty of the Arts, KU Leuven	translation policy		
	Erik Hertog	/	KU Leuven	/		
	Arnaud Szmalec	Belgium	Psychological Sciences Research Institute; Université Catholique de Louvain - UCLouvain	Assessing Content and Language Integrated Learning (CLIL): Linguistic, cognitive and educational perspectives		
	Wouter Duyck	Belgium	Department of Experimental Psychology, Ghent University	LEMMA project		
Van der Linden et al., 2018	Lize Van der Linden	Belgium	Psychological Sciences Research Institute, Université Catholique de Louvain - UCLouvain	Cognitive control of language in the bilingual brain: Behavioral and brain correlates in (a)typical populations	Frontiers in Psychology	SSCI
	Eowyn Van de putte	Belgium	Department of Experimental Psychology, Ghent University	/		
	Evy Woumans			LEMMA project		
	Wouter Duyck	Belgium	Department of Experimental Psychology, Ghent University	LEMMA project		
	Arnaud Szmalec	Belgium	Psychological Sciences Research Institute, Université Catholique de Louvain - UCLouvain	Assessing Content and Language Integrated Learning (CLIL): Linguistic, cognitive and educational perspectives		

Table 7 (*continued*)

Wei, 2017	Wei Yue	China	University of Electronic Science and Technology of China	/	/	
Woumans et al., 2015	Evy Woumans	Belgium	Department of Experimental Psychology, Ghent University	LEMMA project	Journal of Experimental Psychology: Learning, Memory, and Cognition	
	Evy Ceuleers	/	/	/		
	Lize Van der Linden	Belgium	Psychological Sciences Research Institute, Université Catholique de Louvain - UCLouvain	Cognitive control of language in the bilingual brain: Behavioral and brain correlates in (a)typical populations		
	Arnaud Szmalec	Belgium		Assessing Content and Language Integrated Learning (CLIL): Linguistic, cognitive and educational perspectives		
	Wouter Duyck	Belgium	Department of Experimental Psychology, Ghent University	LEMMA project		
Yudes et al., 2011	Carolina Yudes	Spain	Departamento de Psicología Evolutiva y de la Educación, University of Malaga	/	Frontiers in psychology	SSCI
	Pedro Macizo	Spain	Department of Experimental Psychology and Physiology of Behaviour, University of Granada	Second language learning		
	Maria Teresa Bajo	Spain	University of Granada	Executive Control Training; Inhibitory control as a mechanism of memory regulation		
Zou, 2015	Zou Deyan	China	Shanghai International Studies University	/	/	

3.8. Assessing risk of bias

Risk of bias was assessed in RevMan (see **Figure 5** and **Figure 6**). There was a 17% risk of a selection bias and a 28 % chance of incomplete data bias after ignoring concealment- and blinding-caused biases. Taken together, the bias risk was relatively low for the present systematic review, thus confirming validity of the included literature.

Table 7 shows author names, countries, departments and projects and journal titles. Nine out of the 29 included studies are from China, mainly led by Dong Yanping and Liu Yuhua. Other studies are led by authors from Thailand, Sweden, Italy, Poland, Argentina, France, U.S., Spain, Greece, Czechia, and Belgium. In addition to peer-reviewed journal articles, there are three Chinese dissertations downloaded from CNKI and Baidu Scholar and two international dissertations downloaded from school libraries. Journal ranking was checked according to the Shang Jiao Tong University Core Journal Finding System (http://corejournal.lib.sjtu.edu.cn/findcoreej.htm) and ISSN of the journal both on the article and the website to avoid mistakes. A total of 20 studies included are of high quality as they were published by SSCI, A & HCI, SCIE, CSSCI or Chinese Core Journals.

4. Meta-analyses

4.1. Reproducibility and replicability

One of the ways by which the scientific community confirms the validity of scientific discovery is by repeating the research that produces it. Popper stresses the importance of repeatedly testing and reproducing results before acknowledging the conclusions and their empirical validity (Popper, 2005: 23). In our case, executive functions are not a single mechanism measured by a singular task.

In fact, different EFs (even the same EF) are measured by different tasks under a variety of cross-sectional and longitudinal designs. However, we managed to synthesize data on some commonly used EF tasks for meta-analysis to see if the interpreter advantages can be replicated. These tasks included WCST, task-switching, Stroop, 2-back and some of the complex span tasks. To ensure validity, replicated evidence must be available from at least two primary studies for a task to be included for meta-analysis in RevMan (Card, 2012).

4.2. Results of meta-analyses

4.2.1 Inhibition: Stroop

For a Stroop study to be included, it must meet the following requirements: 1) The moderating factor should be interpreting training or experience; 3) The Stroop effect of the original study should be provided; 3) The comparison should be made between groups with (more) interpreting training and those with (less) or no training; 4) Task moderations are accepted, meaning it can be number Stroop or color-word Stroop. Six publications included the Stroop task. However, only three met all the inclusion criteria, with seven datasets. But only three met all the inclusion criteria. **Figure 7** showed no interpreter advantage on the Stroop task (g = 0.13; 95% CI, -0.07, 0.33; Z = 1.27, p = .20; I^2 = 0%).

	(More) Training			No (Less) Training				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV,Random,95% CI
Babcock and Vallesi, 2017	47	29	21	53	36	21	10.70%	-0.18 [-0.79, 0.43]
Dong and Liu, 2016	31.14	32.28	44	26.31	30.71	35	19.90%	0.15 [-0.29, 0.60]
Dong and Liu, 2016	31.14	32.28	44	28.03	28.42	37	20.60%	0.10 [-0.34, 0.54]
Dong and Liu, 2016	31.14	32.28	44	18.48	38.35	44	22.20%	0.35 [-0.07, 0.78]
Liu and Dong, 2017	33.01	30.83	26	25.85	41.11	26	13.30%	0.19 [-0.35, 0.74]
Liu and Dong, 2017	33.01	30.83	26	35.18	48.06	26	13.30%	-0.05 [-0.60, 0.49]
Total(95% CI)			**205**			**189**	**100.00%**	**0.13 [-0.07, 0.33]**
Heterogeneity: Tau²=0.00; Chi²=2.61, df= 5 (P=0.76); I² = 0%								
Test for overall effect: Z = 1.27 (P=0.20)								

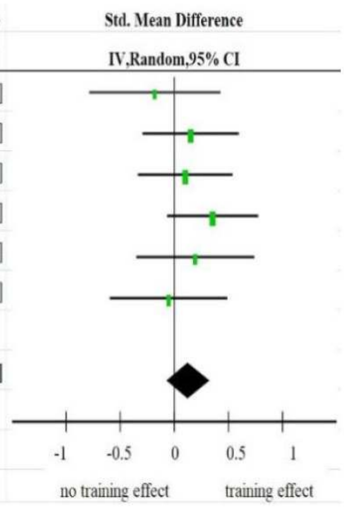


Figure 7. Forest plot on Stroop task, comparing Stroop effect between interpreters and non-interpreters or advanced trainees or less skilled trainees.

4.2.2 Updating: 2-back

The inclusion criteria for the 2-back task were similar to those for the Stroop task, except for the dependent variable being the 2-back mean accuracy score rather than the Stroop effect. Accuracy measured in other manners was converted according to De Smedt (2016). Ten studies included the 2-back task. However, only five studies with ten datasets met all the inclusion criteria.

Forest plot results in **Figure 8** showed a small effect size in favor of interpreter advantage on 2-back (g = 0.23; 95%CI, -0.02, 0.48; Z = 1.82, P = 0.07; I^2 = 61%). Due to the substantial heterogeneity, a sub-group analysis was conducted to identify the cause.

Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
De Smedt, 2016	0.858	0.082	15	0.843	0.094	16	7.10%	0.17 [-0.54, 0.87]
Dong et al., 2018	0.9	0.069	48	0.84	0.087	48	11.40%	0.76 [0.34, 1.17]
Dong et al., 2018	0.9	0.069	48	0.88	0.087	43	11.50%	0.25 [-0.16, 0.67]
Dong and Liu, 2016	0.9	0.068	44	0.84	0.091	44	11.10%	0.74 [0.31, 1.17]
Dong and Liu, 2016	0.9	0.068	44	0.91	0.067	35	10.90%	-0.15 [-0.59, 0.30]
Dong and Liu, 2016	0.9	0.068	44	0.89	0.084	37	11.00%	0.13 [-0.31, 0.57]
Liu and Dong, 2017	0.92	0.051	26	0.93	0.052	26	9.30%	-0.19 [-0.74, 0.35]
Liu and Dong, 2017	0.92	0.051	26	0.9	0.065	26	9.20%	0.34 [-0.21, 0.88]
Liu and Dong, 2020	0.92	0.05	26	0.89	0.07	26	9.20%	0.49 [-0.07, 1.04]
Liu and Dong, 2020	0.92	0.05	26	0.94	0.05	27	9.30%	-0.39 [-0.94, 0.15]
Total(95% CI)			**347**			**328**	**100.00%**	**0.23 [-0.02, 0.48]**
Heterogeneity: Tau^2=0.09; Chi^2=22.79, df= 9 (P=0.007); I^2 = 61%								
Test for overall effect: Z = 1.82 (P=0.07)								

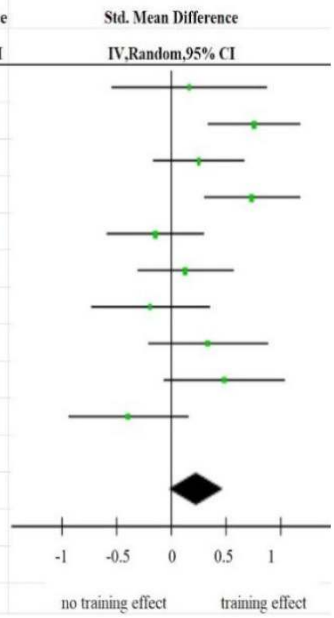


Figure 8. Forest plot of training effect on 2-back task, comparing 2-back accuracy rate between interpreters and controls or trainees' pre-post performance.

Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
1.1.1 Interpreters vs. Non-interpreters								
Dong et al., 2018	0.9	0.069	48	0.88	0.087	43	11.50%	0.25 [-0.16, 0.67]
Dong and Liu, 2016	0.9	0.068	44	0.91	0.067	35	10.90%	-0.15 [-0.59, 0.30]
Dong and Liu, 2016	0.9	0.068	44	0.89	0.084	37	11.00%	0.13 [-0.31, 0.57]
Liu and Dong, 2017	0.92	0.051	26	0.93	0.052	26	9.30%	-0.19 [-0.74, 0.35]
Liu and Dong, 2020	0.92	0.05	26	0.94	0.05	27	9.30%	-0.39 [-0.94, 0.15]
Subtotal (95% CI)			**188**			**168**	**52.00%**	**-0.03 [-0.26, 0.20]**
Heterogeneity: Tau^2=0.01; Chi^2=4.65, df= 4 (P=0.32); I^2 = 14%								
Test for overall effect: Z = 0.26 (P=0.80)								
1.1.2 Pre-traning vs. Post-training								
De Smedt, 2016	0.858	0.082	15	0.843	0.094	16	7.10%	0.17 [-0.54, 0.87]
Dong et al., 2018	0.9	0.069	48	0.84	0.087	48	11.40%	0.76 [0.34, 1.17]
Dong and Liu, 2016	0.9	0.068	44	0.84	0.091	44	11.10%	0.74 [0.31, 1.17]
Liu and Dong, 2017	0.92	0.051	26	0.9	0.065	26	9.20%	0.34 [-0.21, 0.88]
Liu and Dong, 2020	0.92	0.05	26	0.89	0.07	26	9.20%	0.49 [-0.07, 1.04]
Subtotal (95% CI)			**159**			**160**	**48.00%**	**0.58 [0.35, 0.80]**
Heterogeneity: Tau^2=0.00; Chi^2=3.43, df= 4 (P=0.49); I^2 = 0%								
Test for overall effect: Z = 5.03 (P<0.0001)								
Total(95% CI)			**347**			**328**	**100.00%**	**0.23 [-0.02, 0.48]**
Heterogeneity: Tau^2=0.09; Chi^2=22.79, df= 9 (P=0.007); I^2 = 61%								
Test for overall effect: Z = 1.82 (P=0.07)								
Test for subgroup difference: Chi^2 =13.90, df=1 (P=0.0002), I^2=92.8%								

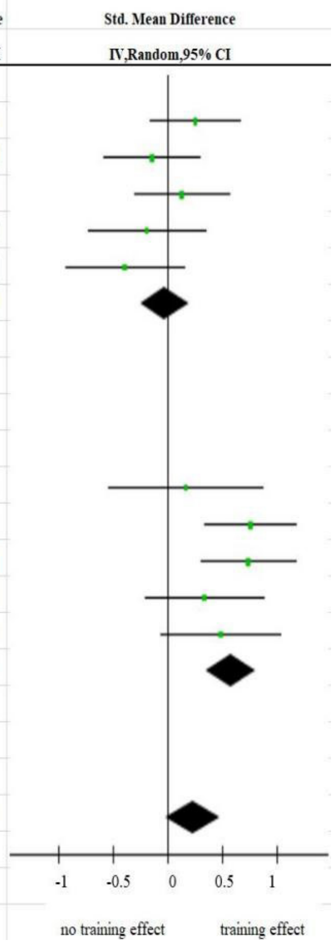


Figure 9. Forest plot of sub-group analysis on 2-back training effect.

In **Figure 9**, the interpreter vs non-interpreter subgroup showed no interpreter advantage on

2-back (g = -0.03; 95% CI from -0.26 to 0.20; Z = 0.26, p = 0.80; I^2 = 14%). However, the pre-training vs post-training subgroup exhibited significant interpreter advantage on 2-back (g = 0.58, 95% CI from 0.35 to 0.80; Z = 5.03, p < 0.00001; I^2 = 0%). The sub-group analysis indicated that that heterogeneity could arise when between- and within-group results were synthesized.

4.2.3 Updating: L2 listening span

L2 listening span is complex span task that requires the participants to recall the last word of a set of each sentence after listening in their second language (L2) and judging if the sentences make sense. The inclusion criteria are the same except for the scoring method, which can be the total number of correctly recalled words (Nour et al., 2020) or the highest number of recalled words for more than two out of five sentence set (truncated span) (Liu et al., 2004). Six studies included the L2 listen span task, but only four met all the selection criteria, with seven datasets. In Liu et al. (2004: 32), the means and standard deviations of the L2 listening span results are provided for professional interpreters, advanced students and beginning students. However, results of the whole student groups are not clearly provided. Equations (1) and (2) presented below are used to calculate the separate means and SDs of the advanced (x) and beginning students (y). After calculation, the mean of the whole student group is 3.295, with SD being 1.597.

$$\bar{z} = \frac{n\bar{x}+m\bar{y}}{m+n} \quad (1)$$

$$\sigma = \sqrt{\frac{n\sigma_x^2+m\sigma_y^2+\frac{mn(\bar{x}-\bar{y})^2}{m+n}}{m+n}} \quad (2)$$

	(More) Training			No (Less) Training				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV,Random,95% CI
Dong et al., 2018	29.75	6.29	48	26.1	5.49	48	23.40%	0.61 [0.20, 1.02]
Dong et al., 2018	29.75	6.29	48	27.97	7.68	43	23.20%	0.25 [-0.16, 0.67]
Liu et al., 2004	3.64	1.23	11	3.295	1.597	22	10.20%	0.23 [-0.50, 0.95]
Liu et al., 2004	3.23	1.19	11	3.36	1.33	11	8.00%	-0.10 [-0.94, 0.74]
Liu and Dong, 2020	36.46	7.76	26	30.27	5.41	26	14.80%	0.91 [0.34, 1.48]
Liu and Dong, 2020	36.46	7.76	26	34.7	5.83	27	16.10%	0.25 [-0.29, 0.79]
Tian, 2016	38.75	6.946	4	40.9	5.99	10	4.40%	-0.32 [-1.49, 0.85]
Total(95% CI)			**174**			**187**	**100.00%**	**0.38 [0.12, 0.63]**
Heterogeneity: Tau^2=0.03; Chi^2=7.92, df= 6 (P=0.24); I^2=24%								
Test for overall effect: Z = 2.93 (P=0.0003)								

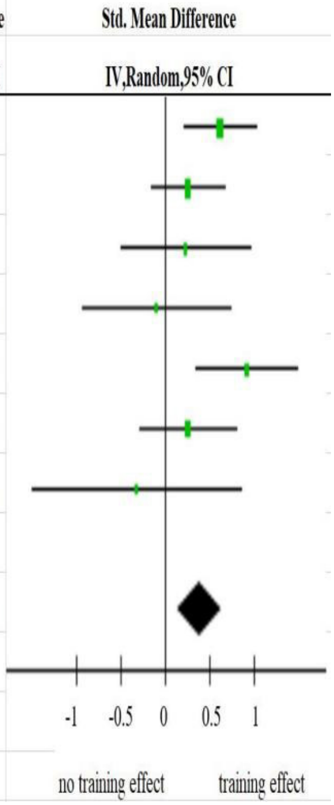


Figure 10. Forest plot of L2 listening span.

The results in **Figure 10** showed that more interpreting training or experience significantly enhanced L2 listening span (g = 0.38; 95% CI, 0.12, 0.63; Z = 2.93, p = 0.0003; I^2 = 24%). Although there was only a small heterogeneity in the datasets (Higgins and Green 2008: 278), we conducted a sub-group analysis to address it.

Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
1.2.1 Interpreters vs. Non-interpreters								
Dong et al., 2018	29.75	6.29	48	27.97	7.68	43	23.20%	0.25 [-0.16, 0.67]
Liu et al., 2004	3.64	1.23	11	3.295	1.597	22	10.20%	0.23 [-0.50, 0.95]
Liu et al., 2004	3.23	1.19	11	3.36	1.33	11	8.00%	-0.10 [-0.94, 0.74]
Liu and Dong, 2020	36.46	7.76	26	34.7	5.83	27	16.10%	0.25 [-0.29, 0.79]
Tian, 2016	38.75	6.946	4	40.9	5.99	10	4.40%	-0.32 [-1.49, 0.85]
Subtotal (95% CI)			**100**			**113**	**61.80%**	**0.18 [-0.09, 0.45]**
Heterogeneity: Tau^2=0.00; Chi^2=1.34, df= 4 (P=0.85); I^2=0%								
Test for overall effect: Z = 1.29 (P=0.20)								
1.2.2 Pre-training vs. Post-training								
Dong et al., 2018	29.75	6.29	48	26.1	5.49	48	23.40%	0.61 [0.20, 1.02]
Liu and Dong, 2020	36.46	7.76	26	30.27	5.41	26	14.80%	0.91 [0.34, 1.48]
Subtotal (95% CI)			**74**			**74**	**38.20%**	**0.71 [0.38, 1.05]**
Heterogeneity: Tau^2=0.00; Chi^2=0.69, df= 1 (P=0.41); I^2=0%								
Test for overall effect: Z = 4.20 (P<0.0001)								
Total(95% CI)			**174**			**187**	**100.00%**	**0.38 [0.12, 0.63]**
Heterogeneity: Tau^2=0.03; Chi^2=7.92, df= 6 (P=0.24); I^2=24%								
Test for overall effect: Z = 2.93 (P=0.003)								
Test for Subgroup differences: Chi^2=5.89, df= 1 (P=0.02); I^2=83.0%								

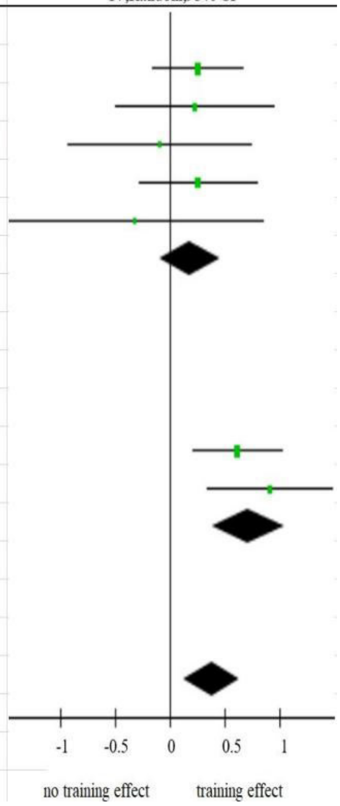


Figure 11. Forest plot of sub-group analysis on L2 listening span.

As presented in **Figure 11**, the 24% heterogeneity was due to the combination of between-group and within-group data. The sub-group analysis showed that interpreters possessed no advantage over non-interpreters in L2 listening span (g = 0.18; 95% CI, -0.09, 0.45; Z = 1.29, p = 0.20; I^2 = 0%). However, post-training interpreters performed significantly better in L2 listening span than before training, with a high effect size (g = 0.71, 95%CI, 0.38, 1.05; Z = 4.20, p < 0.0001; I^2 = 0%).

4.2.4 Shifting: WCST

The inclusion criteria for WCST are the same except for the dependent variable, which is the number of completed categories. Eight studies included WCST, but only four met all the selection criteria, with seven datasets. In Dong and Xie (2014) there are two groups of interpreters and two groups of non-interpreters. Equations (1) and (2) are used again to convert the means and SDs of the separate groups into those of the whole group.

The results in **Figure 12** showed a highly significant interpreter advantage with a medium-to-high effect size and no heterogeneity in the datasets (g = 0.68; 95%CI, 0.48, 0.87; Z = 6.86, p <

0.00001; I^2 = 0%)

Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
Dong and Xie, 2014	12.8	3.4	20	11	3.3	46	13.20%	0.53 [0.00, 1.07]
Dong and Xie, 2014	11.545	3.431	66	9.098	3.453	88	34.60%	0.71 [0.38, 1.04]
Liu, 2018	9.87	2.09	31	8.93	2.03	30	14.50%	0.45 [-0.06, 0.96]
Liu and Dong, 2017	10.27	1.97	26	8.27	2.61	26	11.50%	0.85 [0.28, 1.42]
Liu and Dong, 2017	10.27	1.97	26	8.62	1.92	26	11.60%	0.84 [0.27, 1.40]
Yudes et al., 2011	5.5	1.31	16	4.37	1.66	16	7.20%	0.74 [0.02, 1.46]
Yudes et al., 2011	5.5	1.31	16	4.56	1.45	16	7.30%	0.66 [-0.05, 1.38]
Total(95% CI)			**201**			**248**	**100.00%**	**0.68 [0.48, 0.87]**
Heterogeneity: Tau^2=0.00; Chi^2=1.76, df= 6 (P=0.94); I^2=0%								
Test for overall effect: Z = 6.86 (P<0.00001)								

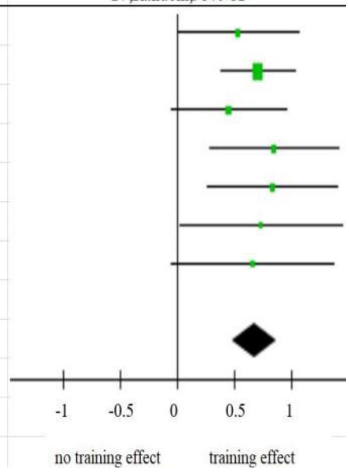


Figure 12. Forest plot on WCST.

4.2.5 Shifting: Task-switching

Five out of the eight task-switching studies meet the inclusion criteria, with nine data sets. The included dependent variable is the switching cost, i.e. difference in response time between repeat trials and switch trials. The higher the switching cost is, the weaker the shifting ability is (Liu, 2018). Results presented in Figure 13 showed that more interpreting training or experience significantly reduced switching cost (g = -0.23; 95%CI, 0.03, 0.43; Z = 2.22; P = 0.03; I^2 = 32%).

Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
Babcock and Vallesi, 2017	192	147	23	234	172	21	8.60%	-0.26 [-0.85, 0.34]
Babcock et al., 2017	131	99	47	155	101	47	14.50%	-0.24 [-0.64, 0.17]
Babcock et al., 2017	131	99	47	147	98	10	6.90%	-0.16 [-0.84, 0.52]
Babcock et al., 2017	131	99	47	105	68	35	13.10%	0.30 [-0.14, 0.74]
De Smedt, 2016	144.64	111.41	15	143.07	128.6	16	6.60%	0.01 [-0.69, 0.72]
Dong and Liu, 2016	94.2	58.19	44	139.23	90.19	44	13.60%	-0.59 [-1.02, -0.16]
Dong and Liu, 2016	94.2	58.19	44	118.71	70.88	35	12.80%	-0.38 [-0.83, 0.07]
Dong and Liu, 2016	94.2	58.19	44	132.73	82.93	37	12.90%	-0.54 [-0.99, -0.10]
Liu, 2018	134.44	8338	31	186.03	107.12	30	11.00%	-0.01 [-0.51, 0.49]
Total(95% CI)			**342**			**275**	**100.00%**	**-0.23 [-0.43, -0.03]**
Heterogeneity: Tau^2=0.03; Chi^2=11.72, df= 8 (P=0.16); I^2=32%								
Test for overall effect: Z = 2.22 (P=0.03)								

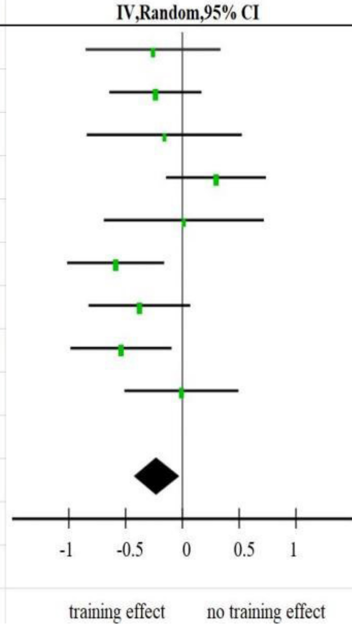
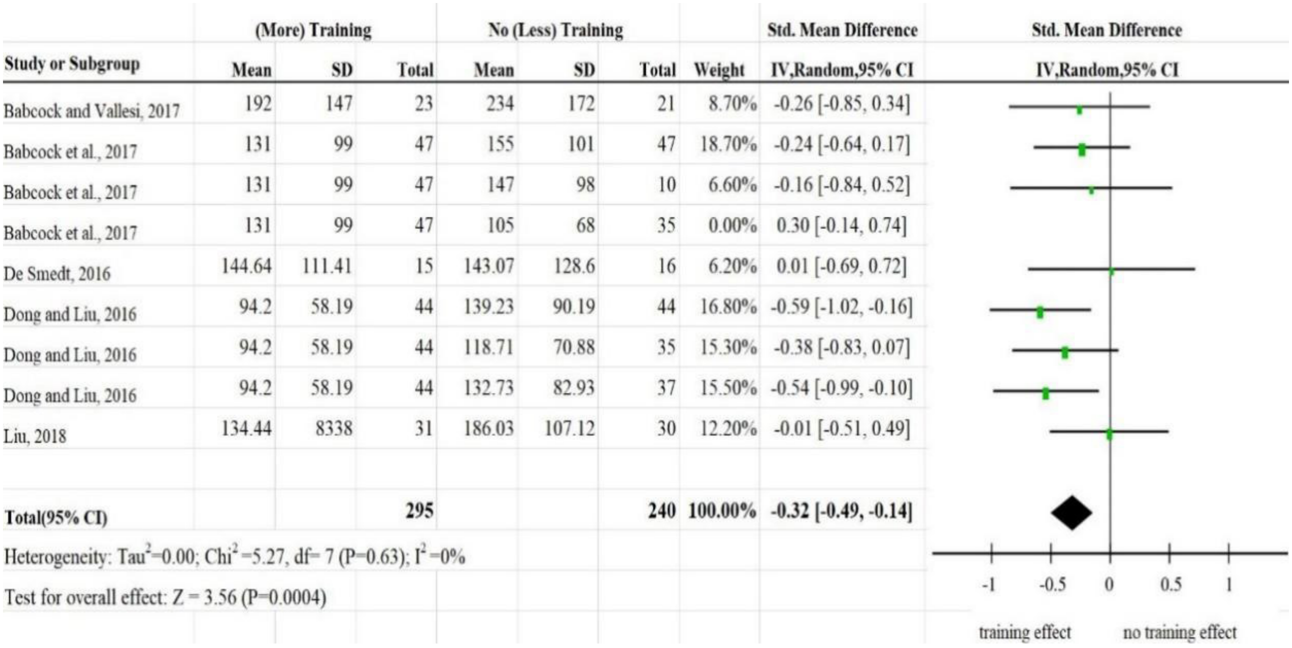


Figure 13. Forest plot on switching cost.

Although the 32% heterogeneity was not too significant to affect the results, a sensitivity analysis was conducted. It was found that the third dataset extracted from Babcock et al. (2017) included interpreters and non-language controls while the other eight datasets compared interpreters with controls with language training. After excluding the heterogeneity, a new forest plot in **Figure 14** showed a bigger effect estimate, with more interpreting training or experience significantly lowering switching cost (g = -0.32; 95% CI, 0.14, 0.49; Z = 3.56, P = 0.0004; I^2 = 0%).



Study or Subgroup	(More) Training Mean	SD	Total	No (Less) Training Mean	SD	Total	Weight	Std. Mean Difference IV,Random,95% CI
Babcock and Vallesi, 2017	192	147	23	234	172	21	8.70%	-0.26 [-0.85, 0.34]
Babcock et al., 2017	131	99	47	155	101	47	18.70%	-0.24 [-0.64, 0.17]
Babcock et al., 2017	131	99	47	147	98	10	6.60%	-0.16 [-0.84, 0.52]
Babcock et al., 2017	131	99	47	105	68	35	0.00%	0.30 [-0.14, 0.74]
De Smedt, 2016	144.64	111.41	15	143.07	128.6	16	6.20%	0.01 [-0.69, 0.72]
Dong and Liu, 2016	94.2	58.19	44	139.23	90.19	44	16.80%	-0.59 [-1.02, -0.16]
Dong and Liu, 2016	94.2	58.19	44	118.71	70.88	35	15.30%	-0.38 [-0.83, 0.07]
Dong and Liu, 2016	94.2	58.19	44	132.73	82.93	37	15.50%	-0.54 [-0.99, -0.10]
Liu, 2018	134.44	8338	31	186.03	107.12	30	12.20%	-0.01 [-0.51, 0.49]
Total(95% CI)			**295**			**240**	**100.00%**	**-0.32 [-0.49, -0.14]**

Figure 14. Forest plot on switching cost after sensitivity analysis.

5. Discussion

5.1. Summary of major findings

Despite growing interest in the cognitive processes of interpreters, prior research on the presumed interpreter advantage in executive functions has produced inconsistent results. To find patterns in these mixed results, the systematic review and meta-analysis synthesized data from 98 tasks of 29 highly relevant studies. As shown in **Table 8**, a shifting advantage was confirmed whereas an inhibitory advantage was rejected. For updating, findings were mixed in the systematic review and the meta-analysis.

Table 8. Summary of results

	EFs	Q1: Cross-sectional	Q2: Longitudinal
Systematic Reviewv	Inhibition	×	×
	Shifting	√	√
	Updating	×	mixed
Meta-analyses	Inhibition (Stroop)	×	×
	Shifting (WCST; Task-switching)	√	√
	Updating (2-back; L2 Listening span)	×	√

5.1.1 Shifting

In cross-sectional between-group comparisons, 69% of publications on shifting supported an interpreter advantage, while all longitudinal studies (100%) presented positive evidence. The results aligned with the conclusion in Nour et al. (2020). Besides, the interpreter advantage in shifting was also shown in the meta-analysis of two shifting tasks (SMD = 0.68 for WCST and SMD = -0.32 for Task-switching). According to the Adaptive Control Hypothesis (Green and Abutalebi, 2013: 17–518), cognitive advantages are modulated by the interactional context being the dual-language, single-language or dense code-switching. As interpreters routinely switch between two languages at work or during training (Aparicio et al., 2017; Babcock and Vallesi, 2017), their abilities in shifting might be significantly enhanced, as proven by most existing literature.

5.1.2 Inhibition

Only five out of the 21 (24%) cross-sectional or correlational tasks supported an inhibitory advantage for interpreters (Timarová et al., 2014; Woumans et al., 2015; Henrard and Van Daele, 2017). Three of the five (60%) longitudinal tasks did not report significant improvement after the training (Babcock et al., 2017; Dong and Liu, 2016; Liu and Dong, 2017). These accord with the review results of Nour et al., (2020). Besides, the meta-analysis of Stroop (SMD = 0.13) exhibited no inhibition advantage of interpreters.

5.1.3 Updating

On the 34 cross-sectional tasks of updating, 23 (68%) revealed no significant difference between interpreters and controls or between experts and novices. On the 14 longitudinal tasks, 50% didn't report a significant training effect. The pooled effect estimates only suggested significant impact of interpreting training on 2-back (SMD = 0.58) and L2 listening span (SMD = 0.71), but not in cross-sectional comparisons. This is consistent with findings of Wen and Dong (2019), but at variance with those of Nour et al. (2020). The discrepancy mainly comes from the 29 (this review: 48 vs Nour et al. 2020: 19) newly-added effects.

5.2. Moderating factors: PICOS

5.2.1 Participant

Participant differences affect the results. Difference in demographics (age, social economic status etc.), language experience and expertise, interactional context (dual, single, or dense code-switching are crucial variables in the development of executive control in bilinguals (Yudes et al., 2011; Green and Abutalebi, 2013; Verreyt et al., 2017; Kidd et al., 2018).

There is a trade-off between age and EFs. With more years, interpreters get to build more cognitive reserves. However, growing older means decline in working memory and executive functioning. (Zhang et al., 2020). For inhibition, Henrard and Van Daele (2017) found superior inhibitory performance of interpreters over translators. while Dong and Liu (2016) didn't reveal such superiority. The much older participants in Henrard and Van Daele (2017) than those in Dong and Liu (2016) (M = 19.85) and (M = 44.98) could be the explanation, meaning an interpreter's inhibitory advantage may emerge in older age. For shifting, Babcock and Vallesi (2017), Macnamara et al. (2011) and Timarová et al. (2014) failed to prove the interpreter advantage in shifting with mean participant ages of 34.1, 42 and 37.1, older than those of other studies. Could this mean that an interpreter's shifting advantage is likely to emerge at a younger age? In the same vein, supporting evidence for an updating advantage mostly came from student interpreters between 19 and 22 years old. With older students at 26.68 and 28.87, Macnamara and Conway (2014; 2015) and Stead and Tripier (2016) did not find an interpreting training effect.

Not only age, but L2 proficiency and switching frequency have also been shown to moderate EFs (e.g. Woumans et al., 2015; Verreyt et al., 2016). In Woumans et al. (2015), student interpreters exhibited higher inhibition accuracy over unbalanced bilinguals but not the balanced bilinguals, possibly due to the moderating effect of L2 proficiency. On the other hand, Verreyt et al. (2016) found inhibitory advantages in balanced switching bilinguals over unbalanced and balanced non-switching bilinguals, indicating that language switching might be a key determinant.

5.2.2 Intervention

Interpreting experience is not categorical but continuous, like bilingualism (Luk and Bialystok, 2013). At different levels of expertise, the interpreting experience or training intervenes differently. According to Chein and Schneider (2012), during the three stages of skill acquisition, *formation, controlled execution and automatic execution*, there is a shift from metacognition to cognitive control to representation. As the cognitive control network is heavily recruited during the stage of controlled execution, the most likely period to see superior cognitive control could be during intense training. When interpreters start training, the metacognitive system plays a dominant role, with participants not engaging in the necessary code-switching practice and being insufficiently exposed to the cognitive processes of interpreting. On the other hand, professional interpreters can find interpreting effortless if they have automated language and processing control (Dong and Li, 2020). In Hervais-Adelman et al. (2015) recruitment of the right caudate nucleus was reduced in simultaneous interpreting after 15 months of intense training. Hervais-Adelman et al. (2017) revealed that cortical thickness increased after simultaneous interpreting training. Such structural change decreases demand on cognitive control as the task becomes more automatized. This is in line with the adaptive control hypothesis (Green and Abutalebi, 2013)

So far, the exact amount of training that brings on the EF advantages has not been confirmed. In Dong and Liu (2016), and Liu and Dong (2017; 2020), with other factors being similar, results diverged due to different duration of training (2016: 1 semester and 32 class hours; 2017: 1 year and 144 class hours; 2020: 1 year and 144 class hours) and students being English or non-English majors. Dong and Liu (2020) believed that participants at the beginning stage or at lower levels endure more interpreting pressure and therefore may need WM and EF more.

5.2.3 Control

The lack of differences between interpreters and other well-matched linguistic groups may be attributable to the fact that learning is such a fundamental human behavior that it is constantly pursued in multiple ways. The control group may not engage in the heavy code-switching necessary for developing interpreting skill, but they may pursue a myriad of other goals and interests intensely. Other acquired skills, such as being a professional musician (Bialystok and DePape, 2009), playing American football (Wylie et al., 2018) and aerobic exercise (see a review by Heijnen et al., 2016), have all been shown to produce discernible effects on cognition.

5.2.4 Outcome

Task impurity has a negative impact on the outcome. Since EFs are three independently single mechanism but unitary to some degree (Miyake et al., 2000), one task could test more than one aspect of EFs. For example, task switching requires inhibition as well as the shifting function, which explains why the color-shape switching in Dong and Liu (2016) and the color-word switching Babcock and Vallesi (2017) and Babcock et al. (2017) produced different results. With more complex stimuli, the latter two studies did not show an interpreter advantage in shifting. For complex span (e.g. operation span; listening span), tasks requiring only the ability to understand and judge plausibility led to positive findings (Chmiel, 2018; Attanak et al., 2019). However, if the task requires more focus on information details, or conducting an arithmetic calculation, researchers don't find the advantage (Stead and Tripier, 2016; Babcock et al., 2017; Liu and Dong, 2020).

5.2.5 Study design

Most study designs in the existing literature were cross-sectional, correlational or longitudinal with a control group. Only a few studies followed the pre/post and experimental/control longitudinal design to explain the causal relationship between interpreting training/experience and EFs (e.g. Dong and Liu, 2016, Liu and Dong, 2017, 2020; Dong et al., 2018). More studies focused on how interpreting experience enhanced EFs, not how bigger a role EFs play in shaping interpreting performance, with a few exceptions (e.g. Liu et al., 2004; Dong and Xie, 2014; Timarová et al., 2014).

6. Conclusion

This systematic review and meta-analysis found significant evidence for the interpreter advantage in shifting, mixed findings in updating and little support in terms of inhibition. Inconsistency in previous studies is mainly caused by the heterogeneity of the demographic background, second language and interpreting experience, the universality of executive functioning, the diversity of experimental tasks and indicators used, and the mismatch between interpreting experience and experimental tasks. The findings of the present study can be replicated and extended. In decades ahead, increasing research on the role of executive functions in interpreting practice and vice versa will expand current knowledge of this growing field.

Conflict of interest

No conflict of interest was reported by all authors.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities of China (Grant number: 2019skzx-pt211).

References

Aparicio X, Heidlmayr K and Isel F (2017) Inhibition efficiency in highly proficient bilinguals and simultaneous interpreters: Evidence from language switching and Stroop tasks. *Journal of Psycholinguistic Research* 46: 1427–1451.

Atkinson LZ and Cipriani A (2018) How to carry out a literature search for a systematic review: A practical guide. *BJPsych Advances* 24(2): 74–82.

Attanak A, Juntapremjit S and Chaisuriya A (2019) The impact on working memory of the 4-D multiple resource model as embedded in interpretation training. *Journal of Community Development Research (Humanities and Social Sciences)* 12(4): 125–135.

Babcock L and Vallesi A (2017) Are simultaneous interpreters expert bilinguals, unique bilinguals, or both. *Bilingualism: Language and Cognition* 20(2): 403–417.

Babcock L, Capizzi M, Arbula S, et al. (2017) Short-term memory improvement after simultaneous interpretation training. *Journal of Cognitive Enhancement* 1(3): 254–267.

Baddeley A (1996) Exploring the central executive. *The Quarterly Journal of Experimental Psychology* 49A(1): 5–28.

Baddeley A (2012) Working memory: Theories, models, and controversies. *The Annual Review of Psychology* 63(1): 1–29.

Borenstein M, Hedges L, Higgins J, et al. (2011) *Introduction to Meta-analysis*. United Kingdom: John Wiley & Sons.

Bialystok E and DePape A (2009) Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance* 35(2): 565–574.

Card N (2012) *Applied Meta-Analysis for Social Science Research*. New York: The Guilford Press.

Chein JM and Schneider W (2012) The brain's learning and control architecture. *Current Directions in Psychological Science* 21: 78–84.

Chen S (2017) The construct of cognitive load in interpreting and its measurement. *Perspectives* 25(4): 640–657.

Chmiel A (2018) In search of the working memory advantage in conference interpreting-training, experience and task effects. *International Journal of Bilingualism* 22 (3): 371–384.

Christoffels IK, de Groot AMB and Kroll JF (2006) Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language* 54(3): 324–345.

Conway A, Kane M, Bunting M, et al. (2005) Working memory span task: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12(5): 769–786.

Darò V and Fabbro F (1994) Verbal memory during simultaneous interpretation: Effects of phonological interference. *Applied Linguistics* 15(4): 365–381.

De Smedt S (2016) *The effect of interpreter training on working memory executive control*. Master's Thesis, University of Gent.

Diamond A (2013) Executive functions. *Annual Review of Psychology* 64: 135–168.

Dong Y and Li P (2020) Attentional control in interpreting: A model of language control and processing control. *Bilingualism: Language and Cognition* 23(4): 716–728.

Dong Y, Li Y and Zhao N (2019) Acquisition of interpreting strategies by student interpreters. *Interpreter and Translator Trainer* 13(4): 408–425.

Dong Y and Liu Y (2016) Classes in translating and interpreting produce differential gains in switching and updating. *Frontiers in Psychology* 7: 1297.

Dong Y, Liu Y and Cai R (2018) How does consecutive interpreting training influence working memory: A longitudinal study of potential links between the two. *Frontiers in Psychology* 9: 875.

Dong Y and Xie Z (2014) Contributions of L2 proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *Journal of Cognitive Psychology* 26 (5): 506–519.

Dong Y and Zhong F (2019) The intense bilingual experience of interpreting and its neurocognitive consequences. In: Schwieter J (ed) *The Handbook of the Neuroscience of Multilingualism*. Chichester: Wiley-Blackwell, pp.685–700.

Fantinuoli C (2018) Interpreting and technology: the upcoming technological turn. In: Fantinuoli C (ed) *Interpreting and Technology*. Berlin: Language Science Press, pp.1–12.

Friedman NP and Miyake A (2017) Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex* 86: 186–204.

García AM (2014) The interpreter advantage hypothesis: Preliminary data patterns and empirically motivated questions. *The Journal of the American Translation and Interpreting Studies Association* 9: 219–238.

Gerver D (1975) A psychological approach to simultaneous interpretation. *Meta* 20(2): 119–128.

Gilbert S and Burgess P (2008) Executive function. *Current Biology* 18(3): R110–R114.

Gile D (2009) *Basic Concepts and Models for Interpreter and Translator Training (2nd ed.)*. Amsterdam: John Benjamins.

Gile D (2011) Errors, omissions and infelicities in broadcast interpreting. Preliminary findings from a case study. In: Alvstad C, Hild A and Tiselius E (eds) *Methods and Strategies of Process Research. Integrative Approaches in Translation Studies*. Amsterdam/Philadelphia: John Benjamins, pp.201–218.

Gile D (2008) Local cognitive load in simultaneous interpreting and its implications for empirical research. *Forum* 6: 59–77.

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

Heijnen S, Hommel B, Kibele A, et al. (2016) Neuromodulation of aerobic exercise—A review. *Frontiers in Psychology* 6: 1890. DOI: 10.3389/fpsyg.2015.01890.

Henrard S and Van Daele A (2017) Different bilingual experiences might modulate executive tasks advantages: Comparative analysis between monolinguals, translators, and interpreters. *Frontiers in psychology* 8: 1–11. DOI: 10.3389/ fpsyg.2017.01870.

Hervais-Adelman A, Moser-Mercer B and Golestani N (2015) Brain functional plasticity associated with the emergence of expertise in extreme language control. *Neuroimage* 114: 264–274.

Hervais-Adelman A, Moser-Mercer B, Murray MM, et al. (2017) Cortical thickness increases after simultaneous interpretation training. *Neuropsychologia* 98: 212–219.

Higgins J and Green S (2008) *Cochrane Handbook for Systematic Reviews of Interventions*. England: John Wiley & Sons Ltd.

Hiltunen S, Paakkonen R, Vik GV, et al. (2016) On interpreters' working memory and executive control. *International Journal of Bilingualism* 20(3): 297–314.

Injoque-ricle I, Barreyro JP, Formoso J, et al. (2015) Expertise, working memory and articulatory suppression effect: Their relation with simultaneous interpreting performance. *Advances in Cognitive Psychology* 11(2): 56–63.

Kane M and Engle R (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review* 9(4): 637–671.

Kidd E, Donnelly S and Christiansen M (2018) Individual differences in language acquisition and processing. *Trends in Cognitive Sciences* 22(2): 154–169.

Köpke B and Nespoulous J (2006) Working memory performance in expert and novice interpreters. *Interpreting* 8(1): 1–23.

Lehtonen M, Soveri A, Laine A, et al. (2018) Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin* 144: 394–425.

Lezak MD (1982) The problem of assessing executive functions. *International Journal of Psychology* 17: 281–297.

Li X (2013) Are interpreting strategies teachable? Correlating trainees' strategy use with trainers' training in the consecutive interpreting classroom. *Interpreters Newsletter* 18(9): 105–128.

Liberati A, Altman D, Tetzlaff J, et al. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Journal of Clinical Epidemiology* 62: e1–e34.

Liu M, Schallert DL and Carroll PJ (2004) Working memory and expertise in simultaneous interpreting. *Interpreting* 6: 19–42.

Liu Y (2018) The influence of interpreting training on young adults' switching ability. *Translation Research and Teaching* 1: 102–108.

Liu Y and Dong Y (2020) A longitudinal study of the relationship between early-stage interpreting and working memory. *Journal of Foreign Languages* 43(1): 112–121.

Liu Y and Dong Y (2017) The influence of early-stage interpreting training on young adults' cognitive control abilities. *Foreign Language Research* 4: 73–78.

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

Macnamara BN and Conway ARA (2014) Novel evidence in support of the bilingual advantage: Influences of task demands and experience on cognitive control and working memory. *Psychonomic Bulletin & Review* 21(2): 520–525.

Macnamara BN and Conway ARA (2015) Working memory capacity as a predictor of simultaneous language interpreting performance. *Journal of Applied Research in Memory and Cognition* 1–9.

Macnamara BN, Moore A, Kegl J, et al. (2011) Domain-general cognitive abilities and simultaneous interpreting skill. *Interpreting* 13(1): 121–142.

Miyake A, Friedman NP, Emerson MJ, et al. (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.

Mellinger CD and Hanson TA (2019) Meta-analyses of simultaneous interpreting and working memory. *Interpreting* 21(2): 165–195.

Morales J, Padilla F, Gómez-Ariza CJ, et al. (2015) Simultaneous interpretation selectively influences working memory and attentional networks. *Acta Psychologica* 155: 82–91.

Nour S, Struys E, Woumans E, et al. (2020) An interpreter advantage in executive functions? A systematic review. *Interpreting* 22(2): 163–186.

Padilla F, Bajo MT and Macizo P (2005) Articulatory suppression in language interpretation: Working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition* 8(3): 207–219.

Pöchhacker F (2016) *Introducing Interpreting Studies (2nd ed.)*. Amsterdam/Philadelphia: John Benjamins.

Popper K (2005) *The Logic of Scientific Discovery*. London and New York: Routledge.

Rosiers A and Vermeiren H (2016) Do Interpreters Have Better Developed Executive Control Skills Compared to Other Language Professionals? *AAAL 2016.* Presented at the AAAL 2016.

Rosiers A, Woumans E, Duyck W, et al. (2019) Is it all in the mind? Investigating the presumed cognitive advantage of aspiring interpreters. *Interpreting* 21(1): 115–134.

Seeber KG (2011) Cognitive load in simultaneous interpreting: Existing theories—New models. *Interpreting* 13(2): 176–204.

Signorelli TM, Henk JH and Loraine KO (2012) Working memory in simultaneous interpreters: effects of task and age. *International Journal of Bilingualism* 16(2): 198–212.

Stachowiak-Szymczak K (2019) *Eye Movements and Gestures in Simultaneous and Consecutive Interpreting*. Cham: Springer International Publishing.

Stavrakaki S, Megari K, Kosmidis M, et al. (2012) Working memory and verbal fluency in simultaneous interpreters. *Journal of Clinical and Experimental Neuropsychology* 34(6): 624–633.

Stead A and Tripier C (2016) *A longitudinal study of the evolution of working memory amongst interpreting and translation students at the FTI*. Thesis, University of Geneva, Geneva.

Tian J (2016) *Working memory performance of simultaneous interpreters—A comparative study of expert and novice interpreters*. Master's Thesis, Guangdong University of Foreign Studies, Guangzhou.

Timarová Š, Čeňková1 I, Meylaerts R, et al. (2014) Simultaneous interpreting and working memory executive control. *Interpreting* 16(2): 39–168.

Verreyt N, Woumans E, Vandelanotte D, et al. (2016) The influence of language switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition* 19(1): 181–190.

Van der Linden L, Van de Putte E, Woumans E, et al. (2018) Does extreme language control training improve cognitive control? A comparison of professional interpreters, L2 teachers and monolinguals. *Frontiers in Psychology* 9: 1998. DOI: 10.3389/fpsyg.2018.01998.

Wang J and Napier J (2013) Signed language working memory capacity of signed language interpreters and deaf signers. *Journal of Deaf Studies and Deaf Education* 18(2): 271–286.

Wei Y (2017) *Contributions of interpreting experience to mental set shifting among young adult interpretation-learners*. Master's Thesis, University of Electronic Science and Technology of China, Chengdu.

Wen H and Dong Y (2019) How does interpreting experience enhance working memory and short-term memory: A meta-analysis. *Journal of Cognitive Psychology* 31(8): 769–784.

Woumans E, Ceuleers E, Van der Linden L, et al. (2015) Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41(5): 1579–1586.

Wylie SA, Bashore TR, Van Wouwe NC, et al. (2018) Exposing an "intangible" cognitive skill among collegiate football players: Enhanced interference control. *Frontiers in Psychology* 9: 49. DOI: 10.3389/fpsyg.2018.00049.

Yudes C, Macizo P and Bajo T (2011) The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology* 2: 309. DOI: 10.3389/fpsyg.2011.00309.

Zhang H, Wu Y and Thierry G (2020) Bilingualism and aging: A focused neuroscientific review. *Journal of Neurolinguistics* 54: 100890. DOI: 10.1016/j.jneuroling.2020.100890.

Zhao H and Dong Y (2020) The early presence and developmental trend of interpreter advantages in cognitive flexibility: Effects from task differences and L2 proficiency. *Translation, Cognition and Behavior* 3(2): 241–262.

Zou D (2016) *Differences of working memory competence in consecutive interpreting between beginners and advanced learners*. PhD Thesis, Shanghai International Studies University, Shanghai, pp.1–177.