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


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 ($\eta^2$) 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 (\( \chi^2 \), or Chi\(^2\)), tau-squared (\( \tau^2 \), or \( \tau^2 \)) 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>
<thead>
<tr>
<th>EFs</th>
<th>EF Tasks</th>
<th>Task Types</th>
<th>Study ID.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluorine</td>
<td>Spatial</td>
<td>Dong and Xie, 2014</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Van der Linden et al., 2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced flanker</td>
<td>Spatial</td>
<td>Van der Linden et al., 2018</td>
<td></td>
</tr>
<tr>
<td>Arrow flanker</td>
<td>Spatial</td>
<td>Timarová et al., 2014</td>
<td></td>
</tr>
<tr>
<td>Simon</td>
<td>Spatial</td>
<td>De Smedt, 2016</td>
<td></td>
</tr>
<tr>
<td>ANT</td>
<td>Spatial</td>
<td>Babcock et al., 2017</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Babcock and Vallesi, 2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>De Smedt, 2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Woumans et al., 2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANTI-V</td>
<td>Spatial</td>
<td>Morales et al., 2015</td>
<td></td>
</tr>
<tr>
<td>Number Stroop</td>
<td>Number</td>
<td>Dong and Liu, 2016</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Liu and Dong, 2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number-letter task</td>
<td>Number/letter</td>
<td>Timarová et al., 2014</td>
<td></td>
</tr>
<tr>
<td>Color-shape switch task</td>
<td>Spatial</td>
<td>De Smedt, 2016</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Dong and Liu, 2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Liu, 2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WCST</td>
<td>Spatial</td>
<td>Dong and Xie, 2014</td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Liu and Dong, 2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Macnamara and Conway, 2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Liu, 2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial</td>
<td>Macnamara et al., 2011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-verbal; Spatial</td>
<td>Wei, 2017</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 1 (continued)

<table>
<thead>
<tr>
<th>EFs</th>
<th>EF Tasks</th>
<th>Task Types</th>
<th>Study ID.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Non-verbal; Spatial</td>
<td>Macnamara and Conway, 2015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-verbal; Spatial</td>
<td>Yudes et al., 2011</td>
</tr>
<tr>
<td>Plus-minus</td>
<td></td>
<td>Non-verbal</td>
<td>Henrard and Van Daele, 2017</td>
</tr>
<tr>
<td>Semantic fluency</td>
<td></td>
<td>Verbal</td>
<td>Woumans et al., 2015</td>
</tr>
<tr>
<td>Complex span (listening span)</td>
<td></td>
<td>Verbal</td>
<td>Liu and Dong, 2020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Attanak et al., 2019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Chmiel, 2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Dong et al., 2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Liu et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Stavrakaki et al., 2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Tian, 2016</td>
</tr>
<tr>
<td>Complex span (reading span)</td>
<td></td>
<td>Verbal</td>
<td>Chmiel, 2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal</td>
<td>Zou, 2016</td>
</tr>
<tr>
<td>Complex span (the automated operation, reading and automated symmetry span)</td>
<td></td>
<td>Spatial</td>
<td>Babcock et al., 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Babcock and Vallesi, 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Macnamara and Conway, 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-verbal</td>
<td>Macnamara et al., 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-verbal</td>
<td>Stead and Tripier, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-verbal</td>
<td>Macnamara and Conway, 2015</td>
</tr>
<tr>
<td>Updating</td>
<td>Span task with articulatory suppression</td>
<td>Verbal</td>
<td>Injoque-Ricle et al., 2015</td>
</tr>
<tr>
<td></td>
<td>N-back</td>
<td>Spatial</td>
<td>Van der Linden et al., 2018</td>
</tr>
<tr>
<td></td>
<td>Single and dual n-back</td>
<td>Spatial</td>
<td>Morales et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Visuo-spatial 2-back</td>
<td>Spatial</td>
<td>Liu and Dong, 2020</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>De Smedt, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Dong and Liu, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Dong et al., 2018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Dong and Liu, 2017</td>
</tr>
<tr>
<td></td>
<td>Dual n-back</td>
<td>Spatial</td>
<td>Attanak et al., 2019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial</td>
<td>Stead and Tripier, 2016</td>
</tr>
<tr>
<td></td>
<td>Letter-memory</td>
<td>Letter</td>
<td>Henrard and Van Daele, 2017</td>
</tr>
<tr>
<td></td>
<td>Free call with suppression</td>
<td>Verbal</td>
<td>Köpke and Nespoulous, 2006</td>
</tr>
<tr>
<td></td>
<td>Category/rhyme probe</td>
<td>Verbal</td>
<td>Köpke and Nespoulous, 2006</td>
</tr>
<tr>
<td></td>
<td>Cued recall</td>
<td>Verbal</td>
<td>Signorelli et al., 2012</td>
</tr>
</tbody>
</table>
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>
<thead>
<tr>
<th>Number of time points</th>
<th>Group characteristics</th>
<th>Study number</th>
</tr>
</thead>
<tbody>
<tr>
<td>One: Cross-sectional design</td>
<td>Between-group comparisons containing one group of interpreters</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Between-group comparisons containing more than one interpreter groups</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Correlational analysis</td>
<td>2</td>
</tr>
</tbody>
</table>

| 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>
<thead>
<tr>
<th>Executive Functions</th>
<th>Tasks Included</th>
<th>Frequency of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response-Distractor</td>
<td>ANT; ANTI-V; Antisaccade; Brown-Peterson; Flanker; Simon; Stroop</td>
<td>26</td>
</tr>
<tr>
<td>Shifting</td>
<td>Color-shape switch; Number-letter; Plus-minus; Task-switching (in switching cost); WCST</td>
<td>24</td>
</tr>
<tr>
<td>Updating</td>
<td>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</td>
<td>48</td>
</tr>
</tbody>
</table>
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

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

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.
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

<table>
<thead>
<tr>
<th>Article</th>
<th>Research design</th>
<th>Reason</th>
<th>Task (s)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babcock et al., 2017</td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>Task-switching</td>
<td>TE+</td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT.</td>
<td>training</td>
<td>Task-switching</td>
<td>ns.</td>
</tr>
<tr>
<td>De Smedt, 2016</td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>Color-shape switch</td>
<td>TE</td>
</tr>
<tr>
<td>Dong and Liu, 2016</td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>Color-shape switch</td>
<td>TE+</td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT.</td>
<td>training</td>
<td>Color-shape switch</td>
<td>I+</td>
</tr>
<tr>
<td>Dong and Xie, 2014</td>
<td>Cros. 1 group IT.</td>
<td>training</td>
<td>WCST</td>
<td>I+</td>
</tr>
<tr>
<td></td>
<td>Cros. &gt;1 group IT.</td>
<td>training</td>
<td>WCST</td>
<td>MI</td>
</tr>
<tr>
<td>Dong and Liu, 2017</td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>WCST</td>
<td>TE+</td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT.</td>
<td>training</td>
<td>WCST</td>
<td>I+</td>
</tr>
<tr>
<td>Henrard and Van Daele, 2017</td>
<td>Cros. 1 group I.</td>
<td>experience</td>
<td>Plus-minus</td>
<td>ns.</td>
</tr>
<tr>
<td></td>
<td>Cros. 1 group I.</td>
<td>experience</td>
<td>Plus-minus</td>
<td>I+</td>
</tr>
<tr>
<td>Macnamara and Conway, 2014</td>
<td>Cros. &gt;1 group IT.</td>
<td>training</td>
<td>Task-switching</td>
<td>I+</td>
</tr>
<tr>
<td></td>
<td>Cros. &gt;1 group IT.</td>
<td>training</td>
<td>WCST</td>
<td>I+</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Article</th>
<th>Research design</th>
<th>Reason</th>
<th>Task(s)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liu and Dong, 2020</td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>Complex span</td>
<td>ns.</td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT.</td>
<td>training</td>
<td>Complex span</td>
<td>ns.</td>
</tr>
<tr>
<td></td>
<td>Longitudinal IT.</td>
<td>training</td>
<td>N-back</td>
<td>ns.</td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT.</td>
<td>training</td>
<td>N-back</td>
<td>ns.</td>
</tr>
<tr>
<td>Article</td>
<td>Research design</td>
<td>Reason</td>
<td>Task(s)</td>
<td>Results</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Attanak et al., 2019</td>
<td>Longitudinal IT. training</td>
<td>Complex span</td>
<td>TE+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>Complex span</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Longitudinal IT. training</td>
<td>N-back</td>
<td>TE+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>N-back</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td>Babcock and Vallesi, 2017</td>
<td>Cros. 1 group I. experience</td>
<td>Complex span</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td>Babcock et al., 2017</td>
<td>Longitudinal IT. training</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Chmiel, 2018</td>
<td>Longitudinal IT. training</td>
<td>Complex span</td>
<td>TE+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design &gt;1 group I. experience</td>
<td>Complex span</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group I. experience</td>
<td>Complex span</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td>De Smedt, 2016</td>
<td>Longitudinal IT. training</td>
<td>N-back</td>
<td>TE+</td>
<td></td>
</tr>
<tr>
<td>Dong and Liu, 2016</td>
<td>Longitudinal IT. training</td>
<td>N-back</td>
<td>TE+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>N-back</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td>Dong et al., 2018</td>
<td>Longitudinal IT. training</td>
<td>N-back</td>
<td>TE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>N-back</td>
<td>IT+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Dong and Liu, 2017</td>
<td>Longitudinal IT. training</td>
<td>N-back</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. Design 1 group IT. training</td>
<td>N-back</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Henrard and Van Daele, 2017</td>
<td>Cros. 1 group I. experience</td>
<td>Letter-memory</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td>Macnamara and Conway, 2014</td>
<td>Cros. &gt;1 group IT. training</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Injoque-Ricle et al., 2015</td>
<td>Corre. 1 group I. experience</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corre. 1 group I. experience</td>
<td>Span with articulatory suppression</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Köpke and Nespoulous, 2006</td>
<td>Cros. &gt;1 group I. experience</td>
<td>Free recall</td>
<td>I+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. &gt;1 group I. experience</td>
<td>Category and rhyme probe task</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. &gt;1 group I. experience</td>
<td>Free recall</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cros. &gt;1 group I. experience</td>
<td>Category and rhyme probe task</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Liu et al., 2004</td>
<td>Cros. &gt;1 group I. experience</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Macnamara et al., 2011</td>
<td>Cros. &gt;1 group I. experience</td>
<td>Complex span</td>
<td>ns.</td>
<td></td>
</tr>
<tr>
<td>Morales et al., 2015</td>
<td>Cros. 1 group I. experience</td>
<td>N-back</td>
<td>I+</td>
<td></td>
</tr>
</tbody>
</table>
Table 6 (continued)

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

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

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Author</th>
<th>Country of the author</th>
<th>Department</th>
<th>Project</th>
<th>Journal Titles</th>
<th>Journal Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attanak et al., 2019</td>
<td>Attapol Attanak</td>
<td>Thailand</td>
<td>Khon Kaen University Language Institute</td>
<td>/</td>
<td>Journal of Community Development Research (Humanities and Social Sciences)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sirikorn Juntasiri</td>
<td>Thailand</td>
<td>Center of Excellence in Cognitive Science, Burapha University</td>
<td>/</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Arnon Chaisuriya</td>
<td>Thailand</td>
<td>Language Institute, Chulalongkorn University</td>
<td>/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babcock et al., 2017</td>
<td>Laura Babcock</td>
<td>Sweden</td>
<td>Department of Neuroscience</td>
<td></td>
<td>Journal of Cognitive Enhancement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mariagrazia Capizzi</td>
<td>Italy</td>
<td>Department of Neuroscience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sandra Arbula</td>
<td>Italy</td>
<td>Cognitive Neuroscience Group, Scuola Internazionale Superiore di Studi Avanzati di Trieste</td>
<td>Hemispheric Asymmetries for Executive Functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Antonino Vallesi</td>
<td>Italy</td>
<td>Department of Neuroscience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babcock and Vallesi, 2017</td>
<td>Laura Babcock</td>
<td>Sweden</td>
<td>Department of Neuroscience, Karolinska Institutet</td>
<td>The neurobiology of musical expertise; Hemispheric Asymmetries for Executive Functions</td>
<td>Bilingualism: Language and Cognition</td>
<td>SSCI</td>
</tr>
<tr>
<td></td>
<td>Antonino Vallesi</td>
<td>Italy</td>
<td>Department of Neuroscience, University of Padova</td>
<td>Hemispheric Asymmetries for Executive Functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chmiel, 2018</td>
<td>Agnieszka Chmiel</td>
<td>Poland</td>
<td>Department of Translation Studies, Adam Mickiewicz University</td>
<td>Respeaking - process, competences, quality, (PINC) Extreme language control: activation and inhibition as bilingual control mechanisms in confrontation interpreting; ADI-LAB PRO</td>
<td>International Journal of Bilingualism</td>
<td>A&amp;HCI; SSCI</td>
</tr>
<tr>
<td>De Smelt, 2016</td>
<td>Sarah De Smelt</td>
<td>Belgium</td>
<td>University Gent</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Dong and Liu, 2016</td>
<td>Dong Yangping</td>
<td>China</td>
<td>School of International Studies, Zhejiang University</td>
<td>Neurocognitive Studies of Interpreter Training, Working memory in second language acquisition and processing</td>
<td>Frontiers in Psychology</td>
<td>SSCI</td>
</tr>
<tr>
<td></td>
<td>Liu Yuhua</td>
<td>China</td>
<td>College of Foreign Studies, South China Agricultural University</td>
<td>Bilingual advantage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dong et al., 2018</td>
<td>Dong Yangping</td>
<td>China</td>
<td>School of International Studies, Zhejiang University</td>
<td>Neurocognitive Studies of Interpreter Training, Working memory in second language acquisition and processing</td>
<td>Frontiers in Psychology</td>
<td>SSCI</td>
</tr>
<tr>
<td></td>
<td>Liu Yuhua</td>
<td>China</td>
<td>College of Foreign Studies, South China Agricultural University</td>
<td>Bilingual advantage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cai Rendong</td>
<td>China</td>
<td>Bilingual Cognition and Education Lab, Guangdong University of Foreign Studies</td>
<td>The effect of using a foreign language on decision-making; Phonetic convergence in nonnative speech; Working memory in second language acquisition and processing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7 (continued)

<table>
<thead>
<tr>
<th>Study Authors</th>
<th>Country</th>
<th>Institution/Department</th>
<th>Research Focus</th>
<th>Journal/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dong and Xie, 2014</td>
<td>China</td>
<td>School of International Studies, Zhejiang University</td>
<td>Neurocognitive Studies of Interpreter Training, Working memory in second language acquisition and processing</td>
<td>Journal of Cognitive Psychology SSCI</td>
</tr>
<tr>
<td>Xie Zhilong, China</td>
<td>China</td>
<td>Foreign Languages College, Jiangxi Normal University</td>
<td>Different bilingual experiences and cognitive control</td>
<td></td>
</tr>
<tr>
<td>Henard and Van Daele, 2017</td>
<td>Belgium</td>
<td>Department of Occupational Psychology, Université de Mons</td>
<td>Cognitive characteristics of work activities</td>
<td>Frontiers in Psychology SSCI</td>
</tr>
<tr>
<td>Agnès Van Daele, Belgium</td>
<td>Belgium</td>
<td>Department of Occupational Psychology, Université de Mons</td>
<td>PROJET EXPERT*CRISE</td>
<td></td>
</tr>
<tr>
<td>Irene Injoque-Ricle, Argentina</td>
<td>Argentina</td>
<td>Department of Cognitive Processes, Psychology Research Institute, Universidad de Buenos Aires</td>
<td>Procesos cognitivos en niños; Procesos cognitivos en adultos; Memoria de trabajo e interpretación simultánea</td>
<td>Advances in Cognitive Psychology SSCI</td>
</tr>
<tr>
<td>Juan Pablo Barreto, Argentina</td>
<td>Argentina</td>
<td>Faculty of Psychology, UBA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jesica Formoso, Argentina</td>
<td>Argentina</td>
<td>Department of Research in Basic Cognitive Processes, National Scientific and Technical Research Council</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virginia I. Jaichenco, Argentina</td>
<td>Argentina</td>
<td>Institute of Linguistics, UBA</td>
<td>Memoria de trabajo e interpretación simultánea</td>
<td></td>
</tr>
<tr>
<td>Köpke and Nespoulsous, 2006</td>
<td>France</td>
<td>University of Toulouse, Toulouse, France</td>
<td>Interpreting</td>
<td>A&amp;HCI; SSCI</td>
</tr>
<tr>
<td>Barbara Köpke, France</td>
<td>France</td>
<td>University of Toulouse</td>
<td>COGNIPROS. Linguistic and cognitive evaluation of prosodic production and</td>
<td></td>
</tr>
<tr>
<td>Jean-Luc Nespoulsous, France</td>
<td>France</td>
<td>University of Toulouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu and Dong, 2017</td>
<td>China</td>
<td>College of Foreign Studies, South China Agricultural University</td>
<td>Bilingual advantage</td>
<td>Foreign Language Research Journals</td>
</tr>
<tr>
<td>Dong Yanping, China</td>
<td>China</td>
<td>School of International Studies, Zhejiang University</td>
<td>Neurocognitive Studies of Interpreter Training, Working memory in second language acquisition and processing</td>
<td></td>
</tr>
<tr>
<td>Liu and Dong, 2020</td>
<td>China</td>
<td>College of Foreign Studies, South China Agricultural University</td>
<td>Bilingual advantage</td>
<td>Journal of Foreign Languages CSSCI</td>
</tr>
<tr>
<td>Dong Yanping, China</td>
<td>China</td>
<td>School of International Studies, Zhejiang University</td>
<td>Neurocognitive Studies of Interpreter Training, Working memory in second language acquisition and processing</td>
<td></td>
</tr>
<tr>
<td>Liu et al., 2004</td>
<td>Hong Kong, China</td>
<td>Centre for Translation, Hong Kong Baptist University</td>
<td>/</td>
<td>Interpreting</td>
</tr>
<tr>
<td>Diane L. Schallert, U.S.</td>
<td>Department of Educational Psychology, University of Texas at Austin</td>
<td>Written and oral language; Constructive Criticism Team</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patrick J. Carroll /</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Liu, 2018</td>
<td>China</td>
<td>College of Foreign Studies, South China Agricultural University</td>
<td>Bilingual advantage</td>
<td>Translation Research and Teaching SSCI</td>
</tr>
<tr>
<td>Macnamara and Conway, 2014</td>
<td>U.S.</td>
<td>Department of Psychological Sciences, Case Western Reserve University</td>
<td>/</td>
<td>Psychonomic Bulletin &amp; Review SSCI</td>
</tr>
<tr>
<td>Brooke N. Macnamara, U.S.</td>
<td>Department of Applied Cognitive Psychology, Claremont Graduate University</td>
<td>Virtues in Communities of Trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andrew R. A. Conway, U.S.</td>
<td>Department of Applied Cognitive Psychology, Claremont Graduate University</td>
<td>/</td>
<td>Journal of Applied Research in Memory and Cognition SSCI</td>
<td></td>
</tr>
<tr>
<td>Macnamara and Conway, 2015</td>
<td>U.S.</td>
<td>Department of Psychological Sciences, Case Western Reserve University</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Brooke N. Macnamara, U.S.</td>
<td>Department of Applied Cognitive Psychology, Claremont Graduate University</td>
<td>Virtues in Communities of Trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andrew R. A. Conway, U.S.</td>
<td>Department of Applied Cognitive Psychology, Claremont Graduate University</td>
<td>/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 7 (continued)

<table>
<thead>
<tr>
<th>Author(s) (Year)</th>
<th>Location</th>
<th>Institution</th>
<th>Research Area</th>
<th>Journal</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macnamara et al., 2011</td>
<td>U.S.</td>
<td>Department of Psychological Sciences, Case Western Reserve University</td>
<td>/</td>
<td>/</td>
<td>Interpreting A&amp;HCI; SSCI</td>
</tr>
<tr>
<td>Adam B. Moore</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Judy Keal</td>
<td>U.S.</td>
<td>Department of Linguistics, University of Southern Maine</td>
<td>ASL Linguistics Project</td>
<td>Acta Psychologica SSCI</td>
<td></td>
</tr>
<tr>
<td>Andrew R. A. Conway</td>
<td>U.S.</td>
<td>Department of Applied Cognitive Psychology, Claremont Graduate University</td>
<td>Virtues in Communities of Trust</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Julia Morales</td>
<td>Spain</td>
<td>Universidad Loyola Andalucia</td>
<td>Modulating cognition with DCS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Francisca Padilla</td>
<td>Spain</td>
<td>Department of Experimental Psychology and Physiology of Behaviour, University of Granada</td>
<td>Cognitive processes in translation or interpreting; Working memory and executive functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carlos J. Gómez-Ariza</td>
<td>Spain</td>
<td>Department of Psychology, Universidad de Jaén</td>
<td>Executive Control Training; Inhibitory control as a mechanism of memory regulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maria Teresa Bajo</td>
<td>Spain</td>
<td>University of Granada</td>
<td>Semantic working memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signorelli et al., 2011</td>
<td>U.S.</td>
<td>Program in Speech-Language--Hearing Sciences, CUNY Graduate Center</td>
<td>The influence of second-language grammar on native language sentence comprehension; Semantic working memory</td>
<td>International Journal of Bilingualism A&amp;HCI; SSCI</td>
<td></td>
</tr>
<tr>
<td>Honk Haarman</td>
<td>U.S.</td>
<td>Center for Advanced Study of Language (CASL), University of Maryland, College Park</td>
<td>Semantic working memory; Creative cognition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorraine K. Oldfield</td>
<td>U.S.</td>
<td>Program in Speech-Language--Hearing Sciences, CUNY Graduate Center</td>
<td>Relationship between Developmental Language Disorder and Developmental Dyslexia</td>
<td>Journal of Clinical and Experimental Neuropsychology SCIE; SSCI</td>
<td></td>
</tr>
<tr>
<td>Stavroula Stavratekaki</td>
<td>Greece</td>
<td>School of Italian Language and Literature, Aristotle University of Thessaloniki</td>
<td>Call for papers for &quot;American Journal of Psychiatry and Neuroscience&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kalliopi Megari</td>
<td>Greece</td>
<td>Department of Cardiothoracic Surgery I, Aristotle University of Thessaloniki</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mary Helen Kosmidis</td>
<td>Greece</td>
<td>School of Psychology, Aristotle University of Thessaloniki</td>
<td>Neuropsychological Assessment of Greek Australian Migrants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maria Apostolidou</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Stead and Tripier, 2016</td>
<td>Swiss</td>
<td>University of Geneva</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Tim, 2016</td>
<td>Switzerland</td>
<td>University of Geneva</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Šárka Timarová</td>
<td>/</td>
<td>Institute of Translation Studies, Charles University in Prague</td>
<td>Improving police and legal interpreting (2011-2012)</td>
<td>Interpreting A&amp;HCI; SSCI</td>
<td></td>
</tr>
<tr>
<td>Ivana Češković</td>
<td>Czechia</td>
<td>Institute of Translation Studies, Charles University in Prague</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Reine Mooykerts</td>
<td>Belgium</td>
<td>Faculty of the Arts, KU Leuven</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Erik Hertog</td>
<td>/</td>
<td>KU Leuven</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Arnaud Szmaloec</td>
<td>Belgium</td>
<td>Psychological Sciences Research Institute; Université Catholique de Louvain - UCLouvain</td>
<td>Assessing Content and Language Integrated Learning (CLIL): Linguistic, cognitive and educational perspectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wouter Duyck</td>
<td>Belgium</td>
<td>Department of Experimental Psychology, Ghent University</td>
<td>LEMMA project</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timarović et al., 2014</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Lize Van der Linden</td>
<td>Belgium</td>
<td>Psychological Sciences Research Institute, Université Catholique de Louvain - UCLouvain</td>
<td>Cognitive control of language in the bilingual brain: Behavioral and brain correlates in (a)typical populations</td>
<td>Frontiers in Psychology SSCI</td>
<td></td>
</tr>
<tr>
<td>Eowyn Van de Hutte</td>
<td>Belgium</td>
<td>Department of Experimental Psychology, Ghent University</td>
<td>/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ewy Wosmans</td>
<td>Belgium</td>
<td>Department of Experimental Psychology, Ghent University</td>
<td>LEMMA project</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wouter Duyck</td>
<td>Belgium</td>
<td>Department of Experimental Psychology, Ghent University</td>
<td>LEMMA project</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arnaud Szmaloec</td>
<td>Belgium</td>
<td>Psychological Sciences Research Institute, Université Catholique de Louvain - UCLouvain</td>
<td>Assessing Content and Language Integrated Learning (CLIL): Linguistic, cognitive and educational perspectives</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 \((g = 0.13; 95\% \ CI, -0.07, 0.33; Z = 1.27, p = .20; I^2 = 0\%\).  

![Figure 7](image_url)

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.
In Figure 9, the interpreter vs non-interpreter subgroup showed no interpreter advantage on
2-back ($g = -0.03; 95\% \text{ CI from } -0.26 \text{ to } 0.20; Z = 0.26, p = 0.80; I^2 = 14\%$). However, the pre-training vs post-training subgroup exhibited significant interpreter advantage on 2-back ($g = 0.58, 95\% \text{ CI from } 0.35 \text{ to } 0.80; Z = 5.03, p < 0.00001; I^2 = 0\%$). 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.

$$Z = \frac{n \bar{x} + m \bar{y}}{m+n}$$

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

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dong et al., 2018</td>
<td>29.75</td>
<td>6.29</td>
<td>48</td>
<td>26.1</td>
<td>5.49</td>
<td>48</td>
<td>23.40%</td>
<td>0.61 [0.20, 1.02]</td>
<td></td>
</tr>
<tr>
<td>Dong et al., 2018</td>
<td>29.75</td>
<td>6.29</td>
<td>48</td>
<td>27.97</td>
<td>7.68</td>
<td>43</td>
<td>23.20%</td>
<td>0.25 [0.16, 0.67]</td>
<td></td>
</tr>
<tr>
<td>Liu et al., 2004</td>
<td>3.64</td>
<td>1.23</td>
<td>11</td>
<td>3.295</td>
<td>1.597</td>
<td>22</td>
<td>10.20%</td>
<td>0.23 [0.50, 0.95]</td>
<td></td>
</tr>
<tr>
<td>Liu et al., 2004</td>
<td>3.23</td>
<td>1.19</td>
<td>11</td>
<td>3.36</td>
<td>1.33</td>
<td>11</td>
<td>8.00%</td>
<td>-0.10 [-0.94, 0.74]</td>
<td></td>
</tr>
<tr>
<td>Liu and Dong, 2020</td>
<td>38.46</td>
<td>7.76</td>
<td>26</td>
<td>30.27</td>
<td>5.41</td>
<td>26</td>
<td>14.80%</td>
<td>0.91 [0.34, 1.48]</td>
<td></td>
</tr>
<tr>
<td>Liu and Dong, 2020</td>
<td>38.46</td>
<td>7.76</td>
<td>26</td>
<td>34.7</td>
<td>5.83</td>
<td>27</td>
<td>16.10%</td>
<td>0.25 [0.29, 0.79]</td>
<td></td>
</tr>
<tr>
<td>Tim, 2016</td>
<td>38.75</td>
<td>6.946</td>
<td>4</td>
<td>40.9</td>
<td>5.99</td>
<td>10</td>
<td>4.40%</td>
<td>-0.32 [-1.49, 0.85]</td>
<td></td>
</tr>
</tbody>
</table>

| Total (95% CI)        | 174  | 187  | 100.00% | 0.38 [0.12, 0.63] |

Heterogeneity: $T^2 = 0.05; \chi^2 = 7.92, df = 6 (P = 0.24); I^2 = 24\%$

Test for overall effect: $Z = 2.93 (P = 0.0003)$

**Figure 10.** Forest plot of L2 listening span.
The interpreter advantage in executive functions—A systematic review and meta-analysis

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.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>(More) Training</th>
<th>No (Less) Training</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Total</td>
<td>Mean</td>
</tr>
<tr>
<td>1.2.1 Interpreters vs. Non-interpreters</td>
<td>Dong et al., 2018</td>
<td>29.75</td>
<td>6.29</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Liu et al., 2004</td>
<td>3.64</td>
<td>1.23</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Liu et al., 2004</td>
<td>3.23</td>
<td>1.19</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Liu and Dong, 2020</td>
<td>36.46</td>
<td>7.76</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Tim, 2016</td>
<td>38.75</td>
<td>6.946</td>
<td>4</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>100</td>
<td>4</td>
<td>113</td>
<td>61.80%</td>
</tr>
<tr>
<td>Heterogeneity: (\chi^2=2.34, df= 4 (P=0.85); I^2=0%</td>
<td>Test for overall effect: (Z = 1.29 (P=0.20))</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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 <
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 ($g = -0.23; 95\% CI, 0.03, 0.43; Z = 2.22; P = 0.03; I^2 = 32\%$).

![Forest plot on switching cost](image)

**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 ($g = -0.32; 95\% \text{ CI}, 0.14, 0.49; Z = 3.56, P = 0.0004; I^2 = 0\%$).

![Figure 14. Forest plot on switching cost after sensitivity analysis.](image)

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>
<thead>
<tr>
<th>EFs</th>
<th>Q1: Cross-sectional</th>
<th>Q2: Longitudinal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic Review</td>
<td>Inhibition</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Shifting</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Updating</td>
<td>×</td>
</tr>
<tr>
<td>Meta-analyses</td>
<td>Inhibition (Stroop)</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Shifting (WCST; Task-switching)</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Updating (2-back; L2 Listening span)</td>
<td>×</td>
</tr>
</tbody>
</table>
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; Timaróvá 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.

Acknowledgements

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

References


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*


