Guay, S., Landry, M., Rigoulot, S., & Boller, B. (2025). A Systematic Review and a Meta-Analysis of Age-Related Differences in Inhibitory Control on the Flanker Task. Experimental Aging Research, 1–36.
https://doi.org/10.1080/0361073X.2025.2548740

A systematic review and a meta-analysis of age-related differences in inhibitory control on the flanker task

Sandryne Guay [a, b]*, Mathieu Landry [a], Simon Rigoulot [a] & Benjamin Boller [a, b]

[a] *Department of Psychology, Université du Québec à Trois-Rivières, QC, Canada;*

[b] *Research Centre of the Institut universitaire de gériatrie de Montréal, QC, Canada*

*Corresponding author: Sandryne.guay@uqtr.ca

A systematic review and meta-analysis of age-related differences in inhibitory control on the flanker task

Aging is associated with declines in cognitive functions, particularly inhibitory control. The flanker task is widely used to assess this function; however, research findings on age-related differences remain inconsistent. This systematic review and meta-analysis synthesize findings from 22 studies comparing young and older adults across different versions of the flanker task. The results confirm that older adults exhibit slower RTs, particularly on incongruent trials, indicating greater difficulty in suppressing interference. However, differences in accuracy between both groups remain inconsistent, suggesting that older adults sometimes adopt a speed-accuracy trade-off to compensate for processing speed declines.

Our systematic review indicates that variability across studies likely stems from differences in participant demographics, cognitive screening protocols, task design, and statistical approaches. Task variations, such as stimulus type (arrows, letters, or moving stimuli), cueing conditions, and spatial arrangements, significantly influence interference effects. Furthermore, methodological differences in the computation of inhibition cost scoring also contribute to discrepancies in findings.

Meta-analytical results reveal that age-related effects on the flanker task depend on the specific version used. The arrow flanker task produced the most consistent age-related differences, likely due to its standardized implementation across studies. In contrast, letter-based and cued versions exhibited greater variability, potentially due to additional cognitive demands and task complexity. Notably, when controlling for age-related slowing using transformed RTs, some effects were no longer significant, supporting the hypothesis that processing speed rather than inhibitory deficits may explain some of the observed differences.

Future research should standardize task protocols, refine statistical methods, and explore novel adaptations such as dynamic stimuli to better understand inhibitory control changes in aging. Addressing these inconsistencies will enhance our ability to identify age-related inhibitory difficulties and develop targeted interventions to mitigate cognitive decline.

Keywords: cognitive aging; inhibitory control; flanker task; age-related differences; executive function

Introduction

Cognitive aging is a complex and dynamic process that varies between individuals and is characterized by a decline in several cognitive functions. Along with a general slowing of information processing (Salthouse, 1994), diminished working memory capacity (Grady & Craik, 2000; Salthouse, 1994), and decreased cognitive flexibility (Kramer et al., 1999) older adults often show difficulties in inhibitory processing (Hasher & Zacks, 1988; Kramer et al., 1994). For example, older adults often find it more difficult to listen to an animated radio show while driving in increasingly heavy traffic. Cognitive control encompasses multiple processes, including inhibitory control and interference control. The first refers to the ability to suppress predominant responses, whereas the latter involves resisting irrelevant distractions to maintain focus on a task (Diamond, 2013; Egner, 2008). These processes play a crucial role in everyday activities, including the ability to remain engaged and attentive while driving or during a conversation.

In their seminal review, Hasher and Zacks (1988) proposed that the decline of inhibitory mental processes contributes to age-related impairments observed in several other tasks, such as visual search and working memory tasks (e.g. Salthouse & Babcock, 1991). Similarly, Gazzaley et al. (2005) proposed that age-related difficulties in inhibition are linked to reduced working memory performance, potentially due to difficulties in suppressing task-irrelevant information while maintaining the ability to focus on relevant material. Hence, the age-related decline in inhibitory processes pervades several aspects of cognition.

Despite overwhelming evidence showing that aging is marked by declining inhibitory capacities, findings from research based on conflict tasks show some level of inconsistency. Specifically, many studies have explored the effect of age on inhibitory control using paradigms such as the Stroop color-naming task (Kok, 1999) and the Simon task (Proctor et al., 2005), which consistently demonstrate that older adults have a decreased ability to resist and suppress task-irrelevant stimuli. However, in contrast to this body of work, research using the flanker task has produced inconsistent findings. While some studies report that older adults show reduced inhibitory control compared to younger adults (e.g. Endrass et al., 2012; Erb et al., 2021; Kouwenhoven & Machado, 2024), others reveal a larger interference effect in younger adults (e.g. Di Chiaro & Holmes, 2024) or no difference between young and older adults (e.g. Salthouse, 2010). Several factors may explain these inconsistencies, including variations in task design (Wild-Wall et al., 2008).

The flanker task is commonly used to assess interference control, a subcomponent of cognitive control that involves resisting the influence of distracting information (Egner, 2008; Eriksen & Eriksen, 1974; Ridderinkhof et al., 2011). In the original version of the task, a central target letter is presented to participants alongside irrelevant distractors, called flankers. The flankers can either be the same as the target (compatible) or different from the target (incompatible). This experimental approach yields the *flanker effect*, wherein response times (RTs) are slower and error rates are higher for incompatible trials compared to compatible ones. In one study using the letter flanker task, Zeef and Kok (1993) found that older adults exhibit a greater *flanker effect*, which suggests greater difficulty in suppressing task-irrelevant information. This outcome is consistent with the theory of an age-related decline in interference control, which refers to difficulties in filtering out irrelevant information rather than simply suppressing a dominant response (Ridderinkhof et al., 2021).

Building on these findings, the flanker task provides a particularly relevant framework for investigating age-related changes in interference control. This paradigm assesses inhibitory control by requiring individuals to focus on a central target while suppressing the influence of surrounding distractors that can be congruent, incongruent, or neutral. Importantly, it targets the ability to filter out stimuli located outside the attentional spotlight, distinguishing it from other conflict tasks such as the Stroop and Simon paradigms, which involve interference from internal or spatially aligned sources. This spatial separation of relevant and irrelevant information makes the flanker task especially suited to examining how older adults handle distraction originating from the periphery of attention. Given the well-documented age-related decline in attentional control, increased susceptibility to interference from flankers among older adults is frequently observed, although not systematically. These findings suggest that aging may selectively impair the ability to suppress external distractors, a core component of efficient interference control.

Since its inception, the flanker task has been adapted in various forms, including the arrow flanker task. In this version, a central arrow is presented alongside arrows pointing in the same direction (congruent) or the opposite direction (incongruent). A neutral condition is also sometimes used, where flankers are neither congruent nor incongruent arrows, but rather neutral symbols. Other types of stimuli, such as colored squares and moving dots, have also been used in different versions of the task. Hence, variability in task designs across studies may contribute to divergent findings regarding inhibitory control in aging.

Differences in findings may also stem from several other factors. First, it is possible that studies showing no reduced inhibitory control in older adults rest on a sampling bias, recruited participants with largely preserved cognitive functioning (Kramer et al., 1994). Second, different studies employed different tasks to measure inhibition, and the experimental procedures varied across studies. Participants completed multiple inhibition tasks, such as a Stroop or a Simon task, in addition to the Flanker task (e.g. Ludwig et al., 2010). Studies using multiple tasks before administering the Flanker task may yield different results compared to those using only a Flanker task. Engaging in prior cognitively demanding activities can induce mental fatigue, which has been shown to impair selective attention and increase reaction times (RT) in subsequent tasks. This fatigue can also alter response strategies, leading to more conservative decision-making and reduced perceptual certainty (Wylie et al., 2020). Third, a confounding variable related to processing speed skew the results. Older adults typically exhibit an overall slowing of processing speed (Salthouse, 1996), which likely influences the outcome in the flanker task. While most studies controlled for speed of processing between younger and older adults, not all of them did. Moreover, when speed differences were considered, adjustments in statistical tests were made using different methods, such as proportional scores (e.g. Di Chiaro & Holmes, 2024), natural logarithm transformations (van der Lubbe & Verleger, 2002), or interference scoring by subtracting baseline conditions from experimental conditions (e.g. de Bruin & Sala, 2018; Kawai et al., 2012; Lemire et al., 2024). This methodological variability complicates comparisons across studies and may contribute to the mixed results regarding age-related declines in interference control.

Previous work on age-related decline in inhibitory processing already includes some reviews and meta-analyses (e.g. Rey-Mermet & Gade, 2018). Some have provided valuable insights into a wider range of inhibitory control or attentional tasks (Rey-Mermet & Gade, 2018; Verhaeghen, 2011; Verhaeghen & Cerella, 2002). Yet, none has focused specifically on the flanker task and the inconsistencies observed in the context of aging populations. The present study aims to address this gap by conducting a systematic review and meta-analysis to provide an up-to-date and comprehensive overview of age-related changes across different versions of the flanker task.

Methods

Eligibility Criteria of the Selected Studies

Our approach included studies based on the following criteria: 1) included a group of young adults and a group of cognitively healthy older adults, 2) used a flanker task, 3) compared age groups, 4) used RTs as the dependent variable, and 5) were written in English (both the abstract and the full article).

Outcome and analysis

Included studies had to report behavioral data and compare the performance of younger and older adults.

Search strategy and information source

We searched for studies published between January 1st, 2004, and June 30th, 2024. The search was conducted in PubMed, PsycInfo, and PsycNet, using the keywords "flanker task*", “flanker test*”, “flanker inhibit*”, “flanking effect*”, “flanker paradigm*”, “aging", “ageing”, “older*", and “elder*” using the following syntax ("flanker" OR “flanker task*” OR “flanker test*” OR “flanker inhibit*” OR “flanking effect*” OR “flanker paradigm*”) AND (“aging” OR “ageing” OR "older*" OR “elder*”). Potential articles were identified and screened based on their titles and abstracts. In addition, the snowballing method was used by examining the reference list of eligible studies included in this paper, with three studies selected using this method (Erb et al., 2021; Reuter et al., 2017; Rey-Mermet & Gade, 2020). After excluding duplicates, we reviewed 612 articles. Following the screening process, we excluded qualitative studies, reviews, meta-analyses, and off-topic papers (e.g., models studies, clinical trials, studies without age group comparisons). This resulted in 22 potentially eligible studies (Figure 1).

Selection Process and Risk of Bias

The risk of bias of selected studies was assessed using the Newcastle-Ottawa Scale (NOS) (Wells et al., 2014), a widely used tool for assessing the quality of non-randomized studies. The NOS evaluates studies based on three criteria: 1) selection of participants, 2) comparability of study groups, and 3) quality of the outcome assessment procedure. Both S.G. and B.B. evaluated each eligible study based on the inclusion criteria. In case of disagreement emerged, the most conservative result was selected. A consensus was reached, and no major disagreements occurred.

Analysis

We used Meta-Essentials (Suurmond et al., 2017) to conduct the meta-analysis. To obtain a standardized effect size for each comparison, Hedges' g was calculated using the F-statistic of the interaction between flanker type and age group on raw RTs, along with group sizes. When a study did not provide these data, we calculated it using the means, standard deviations, and group sizes, employing the following formula:

$$F_{Interaction} = \frac{MS_{Interaction}}{MS_{Error}}$$

Where

$$MS_{Interaction} = \frac{SS_{Interaction}}{df_{Interaction}}$$

And

$$MS_{Error} = \frac{SS_{Error}}{df_{Error}}$$

Where

$$SS_{Interaction} = \sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij}\left(M_{ij} - M_i - M_j + M_{..}\right)^2$$

df = Degrees of freedom

n_{ij} = Sample size in group *i,* condition *j*

MS = Mean Square

M_{ij} = Mean of group *i*, condition *j*

M_i= Mean of group *i* across all conditions

M_j= mean of condition *j* across all groups

$M_{..}$= grand mean (overall mean across all groups and conditions)

And

$$SS_{Error} = \sum_{i=1}^{k}\sum_{j=1}^{m}\left(n_{ij} - 1\right) \times SD_{ij}^2$$

n_{ij} = Number of participants in group *i,* condition *j*

SD_{ij} = Standard deviation for group *i,* condition *j*

k = Number of groups

m = Number of conditions

SS = Sum of squares

If the means and standard deviation were not included, we sent an e-mail to the corresponding authors to request the missing data. Two researchers did not respond, and their studies were not included in the meta-analysis (Endrass et al., 2012; Fu et al., 2021).

<u>Risk of bias across studies</u>

The presence of heterogeneity was assessed using Cochrane's Q-statistic and the I^2 was used to quantify the heterogeneity among effect sizes (Higgins et al., 2024).

Results

Study Selection and Characteristics

A total of 22 studies met the inclusion criteria, covering the period from 2007 to 2024.

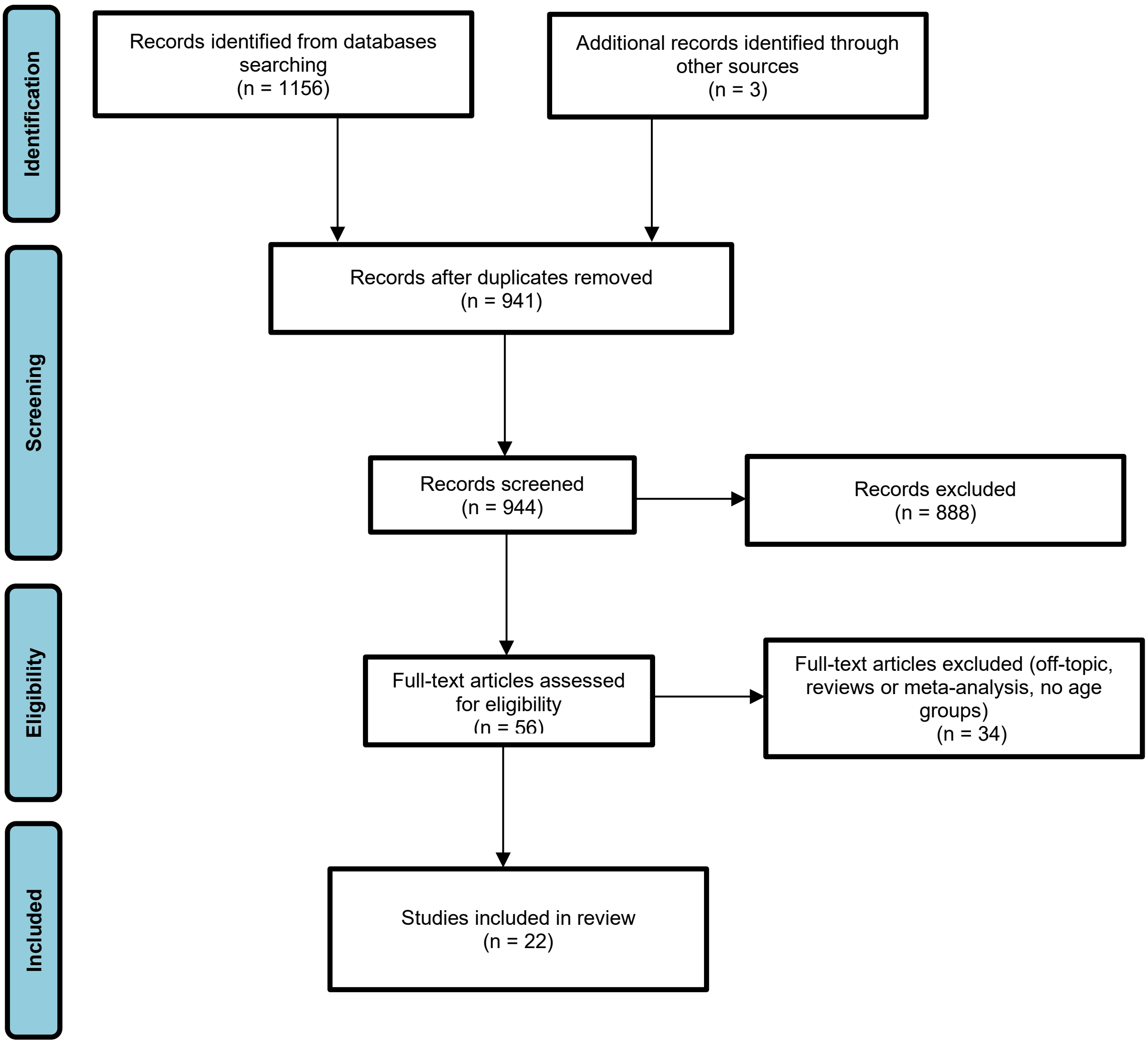

Figure 1. PRISMA flowchart illustrating the selection of studies

Out of the 22 studies selected, 10 used only behavioral data and 12 included physiological data, such as electroencephalography and functional magnetic resonance imaging. Most studies focused on the effect of aging on inhibitory control during the flanker task. These studies were conducted in multiple countries, with the majority originating from North America and Europe. The studies varied in terms of task conditions, with some using additional cognitive tasks (e.g. the Simon task) to assess inhibitory control beyond the flanker task.

Sample sizes ranged from 26 to 302 participants. Participants' ages ranged from 6 to 87 years, with the majority of the studies focusing on healthy older adults. One study also included children (see **Table 1** for more detailed demographic information). Years of education varied between studies, ranging from 10.5 to 17.4 years. Gender distribution was relatively balanced, though some studies had slightly more female than male participants. All studies included cognitively healthy older adults, and the majority screened participants for any signs of dementia or cognitive impairment using a variety of tests, such as the Mini-Mental State Examination (Folstein et al., 1975) and the Montreal Cognitive Assessment (Nasreddine et al., 2005). None of the studies included individuals with self-reported history of neurological or psychiatric disorders.

Table 1. Study demographic characteristics

Study	Groups	N & Sex	Age (mean age ± SD)	Age range	Education (mean years ± SD)	Country	Exclusions
Bowie et al., 2021	Young adults Older adults	N = 24 (11M/13F) N = 24 (10M/14F)	21.67 ± 3.07 71.38 ± 4.19	18 – 29 65 – 80	15.13 ± 1.83 17.4 ± 2.78	United States	psychiatric or neurological illness, depression, scores ≥ 51 on the modified MMSE (old adults), exceedingly low task performance on practice trials
de Bruin & Sala, 2018 **Study 1**	Young adults Older adults	N = 20 (9M/11F) N = 20 (9M/11F)	21.45 ± 2.84 66.35 ± 3.92	18-27 60 – 74	15.60 ± 1.85 16.45 ± 2.67	United Kingdom	Bad vision, bad hearing, neurological disorders, any sign of dementia (older adults; ACE-III < 88)
de Bruin & Sala, 2018 **Study 2**	Young adults Older adults	N = 30 (4M/26F) N = 28 (5M/23F)	20.50 ± 2.60 68.57 ± 6.97	18 – 25 60 – 86	15.37 ± 1.97 16.36 ± 3.68	United Kingdom	Bad vision, bad hearing, neurological disorders, any sign of dementia (older adults; ACE-III < 88)
Di Chiaro & Holmes, 2024	Children Young adults Older adults	N = 92 (45M/47F) N = 25 (9M/16F) N = 33 (14M/19F)	8.8 ± 2.1 28.3 ± 5.1 70.2 ± 6.5	6 – 14 20 – 43 60 – 83	NR	United Kingdom	neurological or psychological disorders, dementia (MoCA; older adults), bad visual acuity (older adults)
Endrass et al., 2012	Young adults Older adults	N = 22 (11M/11F) N = 22 (11M/11F)	22 69.1	19 – 28 62 – 80	12.7± 0.5 11.6 ± 1.7	Germany	Bad vision, neurological and psychiatric diseases, language disorder, any sign of dementia (Older adults; MMST)
Erb et al., 2021	Young adults Older adults	N = 45 (16M/29F) N = 45 (13M/32F)	19 ± 2.5 69 ± 2.8	18 – 34 65 – 75	NR	United States	cognitive or motoric impairment (older adults), language disorder (older adults)

Study	Groups	N & Sex	Age (mean age ± SD)	Age range	Education (mean years ± SD)	Country	Exclusions
Fu et al., 2021	Young adults Older adults	N = 38 (20M/18F) N = 36 (18M/18F)	21.35 ± 1.8 72.17 ± 2.9	19 – 25 68 – 80	NR	China	Psychiatric or neurological illness, taking psychoactive pharmaceutical treatments, bad vision and hearing, any sign of dementia (MMSE < 26; older adults)
Gamboz et al., 2010	Young adults Older adults	N = 70 N = 65	25.8 ± 4.0 67.9 ± 5.6	NR	15 ± 2.1 10.8 ± 4.0	Italy	Psychiatric or neurological illness, taking psychoactive pharmaceutical treatments, bad vision and hearing, any sign of dementia (MMSE < 26; older adults)
Hsieh & Fang, 2012 **study 1**	Young adults Older adults	N = 16 (6M/10F) N = 16 (9M/7F)	20.44 ± 1.71 64.63 ± 4.13	18 – 24 60 – 72	14.25 ± 1.24 14 ± 1.93	Taiwan	Any sign of dementia (All; MMSE)
Hsieh & Fang, 2012 **study 2**	Young adults Older adults	N = 16 (7M/9F) N = 16 (9M/7F)	21.06 ± 1.61 64.13 ± 2.47	19 – 24 60 – 69	14.81 ± 1.05 13.81 ± 1.80	Taiwan	Any sign of dementia (All; MMSE)
Hsieh & Fang, 2012 **study 3**	Young adults Older adults	N = 16 (9M/7F) N = 16 (8M/8F)	21.19 ± 2.20 64.19 ± 5.72	19 – 25 60 – 73	15.188 ± 1.40 13 ± 1.26	Taiwan	Any sign of dementia (All; MMSE)
Hsieh et al., 2012	Young adults Older adults	N = 16 (6M/10F) N = 16 (9M/7F)	20.44 ± 1.71 64.63 ± 4.13	18 – 24 60 – 72	14.25 ± 1.25 14 ± 1.93	Taiwan	Depression (BDI-II), any signs of dementia (All; MMSE)
Jennings et al., 2007	Young adults Older adults	N = 60 (25M/35F) N = 63 (35M/28F)	19.2 ± 0.12 69.14 ± 0.83	18 – 21 61 – 87	14.30 ± 0.13 17.19 ± 0.83	United States	Any sign of dementia (Older adults; MMSE)
Kaufman et al., 2016	Young adults Older adults	N = 19 (7M/12F) N = 16 (8M/8F)	22.9 ± 4.0 64.8 ± 8.0	NR	NR	United States	Any sign of dementia or global cognitive impairment (MMSE < 25; older adults), history of psychiatric illness, learning disabilities
Kawai et al., 2012	Young adults Older adults	N = 13 (8M/5F) N = 15 (13M/2F)	22.6 70	20 – 31 65 – 75	NR	Japan	Bad vision, neurological disorders, medication that would influence performance, any sign of dementia (MMSE; Older adults)

Study	Groups	N & Sex	Age (mean age ± SD)	Age range	Education (mean years ± SD)	Country	Exclusions
Korsch et al., 2016	Young adults Older adults	N = 19 (10M/9F) N = 23 (11M/12F)	23.05 ± 2.76 70.32 ± 3.24	NR	NR	Germany	Bad vision, neurological or psychiatric disorders, cognitive deficits
Kouwenhoven & Machado, 2024	Young adults Older adults	N = 60 (22M/38F) N = 60 (22M/38F)	21 ± 3 71 ± 6	18 – 30 60 – 88	15 14	New Zealand	Neurological disorders, any sign of dementia (Older adults; MMSE)
Lemire et al., 2024	5 age groups	N = 69 (20M/49F) N = 68 (31M/37F) N = 55 (30M/25F) N = 68 (22M/46F) N = 42 (18M/24F)	24 ± 3 39 ± 4 49 ± 4 60 ± 3 69 ± 3	18 – 30 31 – 44 45 – 54 55 – 64 65 – 78	NR	Canada	neurological or psychiatric disorders
Reuter et al., 2017	Young adults Older adults	N = 14 (5M/9F) N = 26 (12M/14F)	22.14 ± 1.83 76.58 ± 1.58	20 – 27 75 – 82	13. 54 ± 1.28 13.92 ± 3.3	Germany	neurological or psychiatric problems, bad vision, colorblind, MMSE <27 (older adults)
Rey-Mermet & Gade, 2020	Young adults Older adults	N = 110 (31M/79F) N = 131 (77M/54F)	22.54 ± 2.63 69.61 ± 2.81	18 – 28 65 – 75	NR	Germany	Bad vision, Colorblindness, MMSE < 27 (older adults), BDI-II > 19 (younger adults) or GDS > 5 (older adults)
Salthouse, 2010	Young adults Adults Older adults	N = 62 (21M/41F)* N = 89 (23M/66F)* N = 114 (39M/75F)*	27 ± 6 51.2 ± 5 72.6 ± 8.9	18 – 39 40 – 59 60 +	15.2 ± 2.3 15.6 ± 2.8 15.6 ± 2.6	United States	MMSE < 24
Wild-Wall et al., 2008	Young adults Older adults	N = 13 (6M/7F) N = 13 (6M/7F)	23.7 ± 3.7 60.9 ± 6.5	NR	NR	Germany	Bad vision, neurological or psychiatric disorders, any drugs affecting the central nervous system
Williams et al., 2016	Young adults Older adults	N = 24 N = 26	21.6 ± 3.0 65.1 ± 5.1	19 – 29 60 – 76	15 ± 1.6 17 ± 3.1	Canada	History of psychiatric or neurological illness, taking psychoactive pharmaceutical treatments, any sign of dementia (MMSE < 26; older adults)

Study	Groups	N & Sex	Age (mean age ± SD)	Age range	Education (mean years ± SD)	Country	Exclusions
Zhou et al., 2011	Young adults Middle aged adults Older adults	N = 30 (15M/15F) N= 30 (15M/15F) N = 30 (15M/15F)	27.8 ± 5.63 51.2 ± 5.85 70.9 ± 5.86	19 – 38 40 – 58 61 – 83	12.8 ± 3.14 11.2 ± 2.98 10.5 ± 4.88	China	no history of significant health problems such as hypertension, diabetes, and atherosclerotic cardiovascular disease, of alcohol or drug abuse and of psychiatric or neurological disorders Participants in the “old adults” group was free of significant abnormalities beyond the expected age-related incidence of atrophy, ventricular dilation, and white matter hyperintensities (seen by IRM)
Zhu et al., 2010	Young adults Older adults	N = 22 (11M/11F) N = 22 (9M/13F)	20 ± 3 74 ± 6	NR	NR	United States	Bad vision, neurological disorders

Note: * The ratio of women in the group was reported, the exact number was calculated by the authors

Flanker Task variations

The flanker task was used with various stimulus presentations, including the arrow flanker task (e.g. >><>>), the letter flanker task (e.g. HHGHH) and the colored arrow flanker task. Stimulus presentation varied across studies: 10 studies presented the target and the distractors simultaneously, while 7 presented the distractors before the target. Some articles included multiple experiments. Eight studies used cues before trials (see **Table 2** for a description of the task variants).

The number of experimental trials varied across studies, ranging from 80 to 1200 trials. The number of practice trials also differed. Four studies altered the proportion of congruent and incongruent trials, while 18 studies maintained equiprobable conditions. Interestingly, one study introduced an original variant of the task, using moving stimuli instead of conventional static arrows.

Most of the studies were conducted in a laboratory setting. However, one (Lemire et al., 2024) was conducted at home without supervision, and another one (Di Chiaro & Holmes, 2024) was carried out during a public engagement event (see **Table 3** for a summary of key characteristics of all studies).

Table 2. Task descriptions in all studies and conclusions on the behavioural age-related inhibition effect

Study	Task description	Behavioural age-related inhibition effect?
Arrows		
Bowie et al., 2021	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent or incongruent. Each trial was cued by an image indicating the probability that the trial is congruent.	No
de Bruin & Sala, 2018 study 2	Arrow version: Participants were asked to indicate the direction of the target arrow, surrounded by arrows pointing in the same direction (congruent), in the opposite direction (incongruent) or by black squares (neutral)	No (RTs and proportional inhibition costs)
Endrass et al., 2012	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent or incongruent and were presented on a vertical axis. Two different instructions were given, one where the participants were instructed to answer as fast as they can, an another one where they were instructed to answer the most precisely possible.	Yes (RTs)
Erb et al., 2020	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) by touching one of two gray squares.	Yes (RTs and Log-Transformed RTs)
Hsieh & Fang, 2012	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral with square flankers. When the target was red, participants had to indicate the opposite direction of the target, whereas when the target was green, they had to indicate direction of the arrow. There were three different studies with different probabilities of the appearance of the two conditions.	No
Hsieh et al., 2012	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral with square flankers. When the target was red, participants had to indicate the opposite direction of the target, whereas when the target was green, they had to indicate direction of the arrow.	No
Kawai et al., 2012	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>).	No

Study	**Task description**	**Behavioural age-related inhibition effect?**
Korsch et al., 2016	Participants were asked to indicate as fast as possible the color of the central arrow, surrounded by flanker arrows. Three rows of three arrows were presented). The arrows surrounding the target could be the same color as the target (congruent), different color (incongruent), they could point in the same direction or in the opposite direction	No
Lemire et al., 2024	Participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent or incongruent. They answered at home	Yes (interference scoring)
Rey-Mermet & Gade, 2020	Arrow version: Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--).	Yes (Log-transformed RTs)
Salthouse, 2010	Arrow version: Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>).	No
Wild-Wall et al., 2008	Two flanker task versions were presented in the study. In one task, participants were asked to indicate in which direction the central target arrow was pointing while ignoring the flanker arrows. The string of arrows could be congruent or incongruent and were presented on a vertical axis. Flankers arrived before the target. In the other version, participants had a similar instruction, but a neutral condition was added, and the flanker and the target arrows were presented simultaneously.	No
Zhu et al., 2010	Participants were asked to identify the direction of the central target arrow flanked by three arrows each side. The flankers could be either in the same direction as the target arrow (congruent), in the opposite direction (incongruent) or be squares (neutral)	Yes (RTs)
Colored		
Di Chiaro & Holmes, 2024	Participants were asked to indicate the color of the central target square flanked by 2 colored squares each side. The string of squares could be congruent (all the same colour), stimulus incongruent (target and flankers not the same color but associated to the same key) or stimulus-response incongruent (flankers and target in different colors not associated to the same key)	Yes (young adults more difficulties than old adults RTs)
Reuter et al., 2017	Participants were asked to indicate the color of the center target cercle and to ignore the flanking cercles. The array of cercles could be congruent (all the same color), incongruent (flanker and target are different colors) or neutral (the flankers where of a color not associated to any key).	Yes (accuracy)
Cued		

Study	Task description	Behavioural age-related inhibition effect?
Fu et al., 2021	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	No
Gamboz et al., 2010	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	Yes (Raw RTs) No (Proportion scores)
Jennings et al., 2007	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	Yes (Raw RTs) No (Proportion scores)
Kaufman et al., 2016	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	Yes (RTs) No (z-transformed RTs)
Williams et al., 2016	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	Yes (Raw RTs) No (RTs *z*-transformed)
Zhou et al., 2011	Participants were asked to indicate the direction of the target arrow, flanked by 2 arrows each side. The string of arrows could be congruent (p.ex >>>>>), incongruent (p.ex >><>>) or neutral (-->--). Each trial was cued by a "*" either replacing the fixation cross, above or below the fixation cross or one above and one below.	Yes (RTs and ratio score transformation)
Letters		
Kouwenhoven & Machado, 2024	Two consonants appeared on the vertical axis. Participants were asked to identify the central letter target. The flanker letter could be the same as the target (compatible), a letter assigned to another button (incompatible) or another neutral letter (neutral)	Yes (RTs and Log-Transformed RTs)
Rey-Mermet & Gade, 2020	Letter version: Participants were asked to identify the category of the central target character. The string could be congruent (p.ex EEUEE or UUUU), incongruent (p.ex HHUHH) or neutral (p.ex %%U%%)	No (Log-transformed RTs)
Salthouse, 2010	Letter version: Participants were asked to identify the central target character. The string could be congruent (p.ex HHHH, incongruent (p.ex HHGHH).	No

Study	Task description	Behavioural age-related inhibition effect?
Moving		
de Bruin & Sala, 2018 study 1	Participants were asked to indicate the motion direction of the central target dots group. The central group of dots was surrounded by two other groups of dots that moved randomly (neutral), in the same direction (congruent) or in the opposite direction (incongruent)	No (RTs and proportional inhibition costs)
de Bruin & Sala, 2018 study 2	Moving version: Participants were asked to indicate the motion direction of the central target dots group. The central group of dots was surrounded by two other groups of dots that moved randomly (neutral), in the same direction (congruent) or in the opposite direction (incongruent)	No (RTs and proportional inhibition costs)

Table 3. Experiment characteristics of the studies

Study	Total Practice Trials	Total Experimental Trials	Stimulus duration	Stimulus	Response	Congruency Probabilities	Block Feedback	Trial Feedback	Other measures	Software used	Place of experiment
Arrows											
Bowie et al., 2021	96 for young adults 48 slow + 96 normal for old adults	864	150 ms	(Semantic cued) Arrows	Two-handed	Cues are equiprobable	With	Without	Neuroimaging	E-Prime 2.0	Lab
de Bruin & Sala, 2018 study 2 (arrow version)	Min 12 trials 96 conflict condition trials	480	Presentation of flankers and target differed 100 ms flanker Max. 3000 ms	Arrows	Two-handed	Equiprobable	Without	Without		PsychoPy	Lab

Study	Total Practice Trials	Total Experimental Trials	Stimulus duration	Stimulus	Response	Congruency Probabilities	Block Feedback	Trial Feedback	Other measures	Software used	Place of experiment
Endrass et al., 2012	15	900	Presentation of flankers and target differed 100 ms flanker 50 ms with target	Arrow (vertical axe)	Two-handed	50/50	Without	With		NR	Lab
Hsieh & Fang, 2012 Study 1	36	1200	Presentation of flanker and target differed 100 ms flanker 50 ms with target	(Colored) arrows	Two-handed	Compatibility (ANTI/PRO) 30/70	Without	Without	Neuroimaging	E-Prime 2.0	Lab
Hsieh & Fang, 2012 Study 2	36	1200	Presentation of flanker and target differed 100 ms flanker 50 ms with target	(Colored) arrows	Two-handed	Compatibility (ANTI/PRO) 50/50	Without	Without	Neuroimaging	E-Prime 2.0	Lab
Hsieh & Fang, 2012 Study 3	36	1200	Presentation of flanker and target differed 100 ms flanker 50 ms with target	(Colored) arrows	Two-handed	Compatibility (ANTI/PRO) 70/30	Without	Without	Neuroimaging	E-Prime 2.0	Lab
Hsieh et al., 2012	36	1200	Presentation of flanker and target differed 100 ms flanker 50 ms with target	(Colored) arrows	Two-handed	Compatibility (ANTI/PRO) 50/50	Without	Without	Neuroimaging	E-Prime 2.0	Lab
Kawai et al., 2012	NR	96	Maximum 2000 ms	Arrow	Two-handed	Equiprobable	Without	Without	Near-infrared spectroscopy	NR	Lab
Lemire et al., 2024	NR	80	500 ms	Arrows	One or two-handed	50/50	Without	Without		Psychopy via Pavlovia	Home

Study	Total Practice Trials	Total Experimental Trials	Stimulus duration	Stimulus	Response	Congruency Probabilities	Block Feedback	Trial Feedback	Other measures	Software used	Place of experiment
Rey-Mermet & Gade, 2020 (arow version)	NR	NR	Maximum 2000 ms	Arrow	Two-handed	Equiprobable neutral, incongruent and congruent	Without	With		NR	Lab
Salthouse, 2010 (arrow version)	20	1000	Maximum 1500 ms	Arrow	Two-handed	Equiprobable	Without	Without		NR	Lab
Wild-Wall et al., 2008 Study 1	Until stable performance	200	Flanker 100 ms before target	Arrow (vertical axe)	One-handed	Equiprobable	Without	With	Neuroimaging	NR	Lab
Wild-Wall et al., 2008 Study 2	Until stable performance	280	Simultaneous	Arrow (vertical axe)	One-handed	equiprobable for neutral, incongruent and congruent	Without	With	Neuroimaging	NR	Lab
Zhu et al., 2010	2-min practice	384	Maximum 2500 ms	Arrow	Two-handed	Equiprobable	Without	Without	Neuroimaging	NR	Lab
Colored											
Di Chiaro & Holmes, 2024	Children: 16 Young and old adults: 32	Children: 64 Young adults: 384 Old adults: 128	Presentation of flanker and target differed 200 ms flanker Maximum 2000 ms with target	Colored squares	Two-handed	Response imbalanced bloc: 50% C, 25% SI, 25% SRI Response balanced: 25% C, 25% SI, 50% SRI	Without	Without		MATLAB & PsychToolBox 3	Lab or public engagement event
Reuter et al., 2017	20	300	200 ms	Colored circles	One-handed	Equiprobable for neutral, incongruent and congruent	Without	Without	Neuroimaging	Presentation	Lab
Korsch et al., 2016	NR	360	250 ms	Colored arrows	One-handed	Equiprobable	Without	Without	Neuroimaging	Presentation	Lab

Study	Total Practice Trials	Total Experimental Trials	Stimulus duration	Stimulus	Response	Congruency Probabilities	Block Feedback	Trial Feedback	Other measures	Software used	Place of experiment
Cued											
Erb et al., 2020	10	192	Maximum 10 000 ms	Cued Arrows	One-handed	Equiprobable	Without	With		MATLAB	Lab
Fu et al., 2021	20	288	Maximum 1700 ms	Cued Arrows	One-handed	Equiprobable	Without	Without		NR	Lab
Gamboz et al., 2010	32	192	Maximum 1700 ms	Cued Arrows	Two-handed	Equiprobable	Without	Without		E-Prime	Lab
Jennings et al., 2007	20	192	Maximum 1700 ms	Cued arrows	Two-handed	Equiprobable	Without	Without		E-Prime	Lab
Kaufman et al., 2016	24	288	Maximum 1700 ms	Cued Arrows	NR	Equiprobable	Without	Only during practice	Neuroimaging	E-Prime 1.0	Lab
Williams et al., 2016	48	576	Maximum 1700 ms	Cued arrows	NR	Equiprobable	With (monetary rewards)	Without	Neuroimaging	Presentation Software v16.5	Lab
Zhou et al., 2011	24	288	Maximum 1700 ms	Cued arrows	Two-handed	Equiprobable	Without	Without		E-Prime 1.1	Lab
Letter											
Kouwenhoven & Machado, 2024	10	84	Max. 1500 ms	Letter	One-handed	Equiprobable	Without	With		MATLAB	Lab
Rey-Mermet & Gade, 2020 (letter version)	NR	NR	Maximum 2000 ms	Letter	Two-handed	Equiprobable neutral, incongruent and congruent	Without	With		NR	Lab
Salthouse, 2010 (letter version)	20	1000	Maximum 1500 ms	Letter	Two-handed	Equiprobable	Without	Without		NR	Lab

Study	Total Practice Trials	Total Experimental Trials	Stimulus duration	Stimulus	Response	Congruency Probabilities	Block Feedback	Trial Feedback	Other measures	Software used	Place of experiment
Moving											
de Bruin & Sala, 2018 study 1	Min. 8 trials, until accuracy = 80% 30 trials baseline Min. 24 trials for conflict condition	300	Max. 3000 ms	Moving dots	Two-handed	Equiprobable	Without	Without		PsychoPy	Lab
de Bruin & Sala, 2018 study 2 (moving version)	Min 12 trials 96 conflict condition trials	480	Presentation of flankers and target differed 100 ms flanker Max. 3000 ms	Moving	Two-handed	Equiprobable	Without	Without		PsychoPy	Lab

Analysis and Results

Table 4 presents the main dependent measures used in each study, as well as the methods used to calculate an inhibition cost to compare the inhibitory ability between groups. **Table 5** provides an overview of the statistical analyses performed on the measures of interest and their results. Analysis of variance (ANOVA) was the most commonly used statistical method to assess the interaction of the flanker effect, alongside analysis of covariance (ANCOVA) and multivariate analysis of covariance (MANCOVA). Some studies applied more advanced statistical methods, such as Bayesian analysis, to assess the strength of evidence for the alternative hypothesis (e.g. presence of an interaction) and the null hypothesis (e.g. absence of an interaction).

Table 4. Main dependant measures and inhibition scoring method

Study	Main dependant measure	Inhibition scoring
Arrow		
Bowie et al., 2021	RTs on correct answers Proportion of corrects answers IES	None
de Bruin & Sala, 2018 Study 2 (arrow version)	RTs Proportion of correct answers	(incongruent RTs – congruent RTs/congruent trials RTs)[2]
Endrass et al., 2012	RTs Error rates	None
Erb et al., 2020	RTs on correct answers	None
Hsieh & Fang, 2012	RTs Error rates	None
Hsieh et al., 2012	RTs Error rates	None
Kawai et al., 2012	RTs Inhibition scoring	Mean RTs in incongruent trials – Mean RTs in congruent trials
Korsch et al., 2016	RTs Error rates	None
Lemire et al., 2024	RTs*	Mean RTs in incongruent trials – Mean RTs in congruent trials
Rey-Mermet & Gade, 2020 (arrow version)	RTs Error rates	None
Salthouse, 2010 (arrow version)	RTs Error rates	None

Wild-Wall et al., 2008	RTs Error rates	None
Zhu et al., 2010	RTs on correct answers Proportion of correct answers	Mean RTs in incongruent trials – Mean RTs in congruent trials
Colored		
Di Chiaro & Holmes, 2024	RTs on correct answers Error rates IES	None
Reuter et al., 2017	RTs on correct answers Error rates	None
Cued		
Fu et al., 2021	RTs Accuracy	None
Gamboz et al., 2010	RTs Proportion scores: mean RTs in each condition/ overall RT (each participant) Error rates	None
Jennings et al., 2007	RTs Error rates	None
Kaufman et al., 2016	Median RTs on correct trials	Incongruent RT – congruent RT
Williams et al., 2016	RTs	(Overall mean - mean RT for an individual participant for a given cue x target condition) / Overall SD
Zhou et al., 2011	RTs	Mean RTs in incongruent trials – Mean RTs in congruent trials
Letters		
Kouwenhoven & Machado, 2024	RTs on correct answers	None

Rey-Mermet & Gade, 2020 (letter version)	RTs Error rates	None
Salthouse, 2010 (letter version)	RTs Error rates	None
Moving		
de Bruin & Sala, 2018 Study 1	RTs Proportion of correct answers	(incongruent RTs− congruent RTs/congruent trials RTs)^2
de Bruin & Sala, 2018 Study 2 (moving version)	RTs Proportion of correct answers	(incongruent RTs – congruent RTs /congruent trials RTs)^2

* RTs were used to calculate the inhibition scoring but were not reported, Studies with "None" did not use a specific method for the inhibition scoring

Table 5. Analysis and results for data of interest of the review

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
Arrow					
Bowie et al., 2021	Three-way mixed ANOVAs were used to analyze the RT and accuracy data, with cue (PC/PE/PI) and congruency (congruent/incongruent) as within-subject factors and age (young/old) as the between-subjects factor.	1: F (1,46) = 26.24, p < .001, η_p^2 = .363 2: F(1,46) = 198.384, p < .001, η_p^2 = .812 3: F(1,46) = 4.588, p = .038, η_p^2 = .091	1: F(1, 46) = 1.184, p = .282, η_p^2 = .025 2: F(1, 46) = 41.181, p < .001, η_p^2 = .472 3: F(1, 46) = 0.037, p = .849, η^2_p = .001	F(1, 46) = 2.802, p = 0.101, ηp^2 = .057	
de Bruin & Sala, 2018 Study 2 (arrow version)	Accuracy scores were analysed using a binary logistic regression analysis. RTs were analyzed by a two-way repeated ANOVA with trial type (congruent, neutral and incongruent) as a within-subject factor and age group	1: F(1, 56) = 70.56, MSE = 12,533.20, p < 0.001, η_p^2 = 0.56 2: F(2, 112) = 168.41, MSE = 362.39, p < 0.001, η_p^2 = 0.75 3: F(2, 112) = 0.59, MSE = 362.39, p = 0.555, η_p^2 = 0.01	1: $\chi^2(1)$ = 13.31, p < 0.001 2: $\chi 2(2)$ = 30.88, p < 0.001 3: $\chi^2(2)$ = 0.82 p = 0.664		$t(56)$ = 2.29, p = 0.026 (younger adults larger inhibition cost than older adults)

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
	(young, old) as a between-subject factor.				
Endrass et al., 2012	Repeated measures ANOVA were computed with the between subject factor age group (young vs. old). Error rates were analyzed with the within subject factor condition (accuracy vs. speed) and response type (correct vs. incorrect)	1: $F(1, 42) = 47.76$, $p = .001$, $\eta^2 = .532$ 2: NR 3: NR	1: $F(1, 42) = 1.69$, $p = .201$ 2: NR 3: NR		
Erb et al., 2020	ANOVA featuring previous congruency (c, i), current congruency (C, I), and response type (switch, repeat) as within-subjects factors, and age group (young adults vs. older adults) as a between-subjects factor.	1: $F(1, 88) =35.12$, $p < 0.001$, $\eta_p^2 = 0.29$ 2: $F(1, 88) =285. 67$, $p < 0.001$, $\eta_p^2 = 0.76$ 3: $F(1,88) = 7.98$, $p = 0.006$, $\eta_p^2 = 0.08$	1: $F(1, 88) = 0.05$, $p = .83$		
Hsieh & Fang, 2012	4-way mixed ANOVA with the between-subject factors of experiment (Experiment 1 = PRO-bias, Experiment 2 = non-bias, Experiment 3 = ANTI-bias) and age group (young, old) and the within-subject factors of trial condition (PRO, ANTI) and flanker type (congruent, neutral, incongruent)	1 : $F(1, 90) = 130.77$, $p < 0.001$ 2 : $F(2, 180) = 214.78$, $p < 0.001$ 3 : $F(2, 180) = 0.820$, $p =$ n.s.	1: $F(1,90) = 4.55$, $p < 0.05$ 2: $F(2, 180) = 68.08$, $p < 0.001$ 3: $F(2, 180) = 21.95$, $p < 0.001$		
Hsieh et al., 2012	ANOVA, with age as a between-subjects factor and condition (PRO, ANTI) and flanker type (congruent, neutral, incongruent) as within-subject factors	1: $F(1, 30) = 137.22$, $p < .001$ 2: $F(1, 30) = 116.82$, $p < .001$ 3: $F(2, 60) = 4.65$, $p < .05$	1: $F(1, 30) = 2.68$, $p = 0.13$ 2: $F(1, 30) = 29.37$, $p < .001$ 3: $F(1, 30) = 0.004$, $p = .95$		

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
Kawai et al., 2012	A 2 (age) x 2 (congruency) x 2 (task) ANOVA	1: $F(1, 26) = 11.08$, $p < 0.005$ 2: $F(1, 26) = 27.11$, $p < 0.001$ 3: $F(1, 26) = 3.43$, $p = 0.075$			$F(1, 26) = 0.08$, $p = 0.77$
Korsch et al., 2016	A 2 × 2 × 2 ANOVA with the within-subject factors flanker (congruent vs. incongruent) and SRC (congruent vs. incongruent), and the between-subject factor Age (young vs. elderly)	1: $F(1,40) = 18.60$, $p < 0.001$ 2: $F(1,40) = 146.60$, $p < 0.001$) 3: n.s.	1: $F(1,40) = 0.38$, $p = 0.540$ 2: $F(1,40) = 3.87$, $p = 0.056$ 3: n.s.		
Lemire et al., 2024	A multivariate analysis of covariance (MANCOVA) was conducted using a factorial model with sex and age as independent variables, the scores of inhibition as dependent variables and education as a control variable	1: Pillai's Trace (16,774) = 4.837, $p < 0.001$.			$F(4,186) = 6.405$, $p < 0.001$, $\eta_p^2 = 0.098$
Rey-Mermet & Gade, 2020 (arrow version)	Three-way ANOVA with congruency (incongruent, congruent) and previous congruency (incongruent, congruent) as within-subject factors and age group (young, older) as a between-subjects factor.	1: F (1, 239)= 236.11, $p < 0.001$, $\eta_p^2 = 0.48$ 2: F (1, 239)= 298.25, $p < 0.001$, $\eta_p^2 = 0.05$ 3: F (1, 239)= 8.61, $p = 0.004$, $\eta_p^2 = 0.001$	1: F (1, 239)= 29.25, $p < 0.001$, $\eta_p^2 = 0.05$ 2: F (1, 239)= 123.34, $p < 0.001$, $\eta_p^2 = 0.1$ 3: F (1, 239)= 16.05, $p < 0.001$, $\eta_p^2 = 0.01$		
Salthouse, 2010 (arrow version)	ANOVAs	1: F (2, 262) = 51.9, $p < 0.01$ 2: F (2, 262) = 292.9, $p < 0.01$ 3: F (2, 262) = 2.0, n.s.	1: F (2, 262) = 0.2, n.s. 2: F (2, 262) = 44.3, $p < 0.01$ 3: F (2, 262) = 6.1, $p < 0.01$		

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
Wild-Wall et al., 2008 Study 1	Mixed ANOVA with the between-factor age (older, young) and the within-factor stimulus (compatible, incompatible)	1: $F(1,28) = 63.6$, p < .001 2: $F(1,28) = 377$, $p < .001$ 3: n.s.	1: $F(1,28) = 11.3$, $p < .01$ 2: $F(1,28) = 110.0$, $p < .001$ 3: $F(1,28) = 8.5$, $p < .01$		
Wild-Wall et al., 2008 Study 2	Mixed ANOVA with the between-factor age (older, young) and the within-factor stimulus (compatible, incompatible; neutral; solo)	1: $F(1,28) = 29.8$, $p < .001$ 2: $F(3,84) = 101.3$, $p < .001$ 3: n.s.	1: $F(1,28) = 8.3$, $p < .001$ 2: $F(3,84) = 48.5$, $p < .001$ 3: $F(3,84) = 2.7$, $p = .096$		
Zhu et al., 2010	Mixed-model ANOVA in which flanker condition (Incongruent versus Congruent) was the repeated-measures factor and age group was the between-group factor	1: $F(1, 42) = 11.73$, $p = 0.001$ 2: $F(1, 42) = 83.105$, $p < 0.001$ 3: $F(1, 42) = 4.267$, $p = 0.045$	1: $F(1, 42) = 3.67$, $p = 0.061$ 2: $F(1, 42) = 9.105$, $p = 0.004$ 3: $F(1, 42) = 2.984$, $p = 0.091$		NR
Colored					
Di Chiaro & Holmes, 2024	A univariate ANOVA was performed, submitting the perceptual, response and general interference effects as dependent variables and log10(age) as a covariate. Significant effects of age were planned to be followed up with independent samples t tests to compare interference effects between the age groups.	1 (perceptual interference effect): $F(1,134) = 5.31$, $p = 0.023$, $\eta_p^2 = 0.0381$, MSE = 10,297 1 (response interference effect): $F(1,134) = 16.9$, $p < 0.001$, $\eta_p^2 = 0.112$, MSE = 11,317 1 (general interference effect): $F(1,134) = 19.2$, $p < 0.001$, $\eta_p^2 = 0.125$, MSE = 23,483		Perceptual interference: $F(1, 134) = 5.31$, $p = 0.023$, $\eta_p^2 = 0.0381$, MSE = 10,297 Response interference: $F(1,134) = 16.9$, $p < 0.001$, $\eta_p^2 = 0.112$, MSE = 11,317 General interference effect: $F(1,134) = 19.2$, $p < 0.001$, $\eta_p^2 = 0.125$, MSE = 23,48	General interference: $t(53) = 4.30$, $p < 0.001$, $d = 0.58$ (young vs old)
Reuter et al., 2017	2 (Age; young, old) x 3 (Condition; incongruent, congruent, neutral) ANOVAs on accuracies and RTs.	1: $F(1, 38) = 155{,}81$, $p < .001$, $\eta_p^2 = .80$ 2: $F(2, 76) = 14.921$, $p < .001$, $\eta_p^2 = .29$	1: $F(1, 38) = 22.46$, $p < .001$, $\eta_p^2 = .37$ 2: $F(2, 76) = 20.36$, $p < .001$, $\eta_p^2 = .35$		

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
		3: $F(2,76) = 2.29$, $p = 0.131$, $\eta_p^2 = 0.067$	3: $F(2, 76) = 3{,}960$, $p = .030$, $\eta_p^2 = .09$		
Cued					
Fu et al., 2021	Two-way mixed model analysis of variance with block (1 vs 2) as a within-subject factor and age group (YA vs OA) as the between-subjects factor for each attentional network	1: $p < 0.001$ 2: NR 3: $p = 0.548$	1: $p > 0.10$		
Gamboz et al., 2010	Mixed factors ANOVA 2 (age: young vs. old) x 4 (cue type: no cue vs. central cue vs. double cue vs. spatial cue) x 3 (flanker type: neutral vs. congruent vs. incongruent) Follow-up mixed factors analyses were conducted with age as a between-subjects factor and cue-flanker conditions as within-subjects factors.	1 : $F(1, 133) = 210.6$, $p < .0001$, $\eta_p^2 = 0.61$ 2 : $F(1.3, 170.4) = 488.9$, $p < .0001$, $\eta_p^2 = 0.78$ 3 : $F(1.3, 170.4) = 5.7$, $p < .0001$, $\eta_p^2 = 0.04$ Conflict effect: $F(1, 133) = 4.9$, $p < .05$, $\eta_p^2 = 0.04$			
Jennings et al., 2007	2 (group: old, young) × 4 (cue type: no cue, center cue, double cue, spatial cue) × 3 (flanker type: neutral, congruent, incongruent) mixed factors ANOVA on RT and accuracy	1: $F(1, 120) = 220.84$, $p < .001$ 2: $F(2, 240) = 646.52$, $p < .001$ 3: $F(2, 240) = 10.70$, $p < .001$	1: n.s. 2: $F(2, 240) = 26.25$, $p < .001$ 3: n.s.		
Kaufman et al., 2016	2-Group (young adults, older adults) × 3-Flanker type (incongruent, neutral,	1: $F(1,33) = 38.75$, $p < 0.0001$	1: $F(2,363) = 36.56$, $p < 0.0001$ 2: $F(1,33) = 0.30$, $p = 0.59$		conflict effect: $F(1,33) = 2.39$, *ps* > 0.13

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
	congruent) × 4-Cue type (no, spatial, double, and center) mixed-model REML ANOVAs.	2: $F(2,363) = 832.73$, $p < 0.0001$ 3: $F(2,363) = 4.59$, $p < 0.02$ Conflict effect: $F(1,33) = 4.32$, $p < 0.05$	3: $F(2,363) = 2.84$, $ps > 0.06$		
Williams et al., 2016	ANOVAs that included the factors age (young, old), cue (no, double, center, spatial), and target (congruent, incongruent)	1: $F(1, 46) = 63.70$, $p < .001$, $\eta_p^2 = 0.58$ 2: $F(1, 46) = 187.76$, $p < .001$, $\eta_p^2 = 0.58$ 3: $F(1, 46) = 8.81$, $p = .005$, $\eta_p^2 = 0.16$.	1: $F(1, 46) = 28.25$, $p = .002$, $\eta_p^2 = 0.11$ 2: $F(1, 46) = 2.45$, $p = .124$, $\eta_p^2 = 0.05$ 3: $F(1, 46) = 3.25$, $p = .078$, $\eta_p^2 = 0.07$		$F(1, 46) = .41$, $p = .527$, $\eta_p^2 = .01$
Zhou et al., 2011	4 (cue condition: center cue, double cue, none cue, spatial cue) x 3 (flanker type: congruent, incongruent, neutral) ANOVA. To determine differences between individual groups, a Student–Newman–Keuls (SNK) test was used	1: $F(2, 87) = 43.863$, $p > 0.05$, SNK: $p < 0.05$ 2: NR 3: $r = 0.54$, $p<0.001$			$F(2, 87) = 16.357$, $p < 0.001$
Letter					
Kouwenhoven & Machado, 2024	Mixed ANOVAs with trial type (compatible, incompatible) as the within-subjects variable and age group (young, old) as the between-subjects variable	1 (raw RT): $F(1, 118) = 87.46$, $p < 0.001$, $\eta_p^2 = 0.43$ 2 (raw RT): $F(1, 118) = 109.89$, $p < 0.001$, $\eta_p^2 = 0.49$ 3 (raw RT): $F(1, 118) = 10.60$, $p = 0.001$, $\eta_p^2 = 0.08$			

Study	Analysis	Results RTs	Results Accuracy	Results IES	Results flanker interference measures
Rey-Mermet & Gade, 2020 (letter version)	Three-way ANOVA with congruency (incongruent, congruent) and previous congruency (incongruent, congruent) as within-subject factors and age group (young, older) as a between-subjects factor.	1: F (1, 239)= 151.22, $p <$ 0.001, $\eta_p^2 = 0.33$ 2: F (1, 239)= 107.28, $p <$ 0.001, $\eta_p^2 = 0.04$ 3: F (1, 239)= 1.67, $p =$ 0.197, $\eta_p^2 < 0.001$	1: F (1, 239)= 15.57, $p <$ 0.001, $\eta_p^2= 0.02$ 2: F (1, 239)= 0.46, $p =$ 0.498, $\eta_p^2< 0.001$ 3: F (1, 239)= 1.06, $p =$ 0.304, $\eta_p^2= 0.001$		
Salthouse, 2010 (letter version)	ANOVAs	1: F (2, 262) = 39.9, p < 0.01 2: F (2, 262) = 137.2, $p <$ 0.01 3: F (2, 262) = 0.0, n.s.	1: F (2, 262) = 1.0, n.s. 2: F (2, 262) = 34.0, $p < 0.01$ 3: F (2, 262) = 0.4, n.s.		
Moving					
de Bruin & Sala, 2018 Study 1	Accuracy data were analysed using a binary logistic regression analysis RTs were analyzed by a two-way repeated ANOVA with trial type (congruent, neutral and incongruent) as a within-subject factor and age group (young, old) as a between-subject factor.	Main effect of trial type $F(2,76) = 15.72$, MSE = 427.04, $p < 0.001$, $\eta_p^2 = 0.29$ Difference of age groups in conflict condition: $F(1, 38) =$ 14.45, MSE = 73,666.75, $p =$ 0.001, $\eta_p^2 = 0.28$ Age x Trial type: $F(2, 76) =$ 1.00, MSE = 427.04, $p =$ 0.374, $\eta_p^2 = 0.03$	Effect of age on conflict condition: $\chi^2(1) = 3.44$, $p =$ 0.06		$t(38) = 1.18$, $p = 0.246$
de Bruin & Sala, 2018 Study 2 (moving version)	Accuracy scores were analysed using a binary logistic regression analysis. RTs were analyzed by a two-way repeated ANOVA with trial type (congruent, neutral and incongruent) as a within-subject factor and age group (young, old) as a between-subject factor.	1: $F(2, 112) = 14.86$, MSE = 695.44, $p < 0.001$, $\eta_p^2 = 0.21$ 2: $F(1, 56) = 24.13$, MSE = 33,117.09, $p < 0.001$, $\eta_p^2 =$ 0.30 3: $F(2, 112) = 0.36$, MSE = 695.44, $p = 0.696$, $\eta_p^2 = 0.01$	1: $\chi^2(1) = 28.77$, $p < 0.001$ 2: $\chi^2(1) = 49.95$, $p < 0.001$ 3: n.s.		$t(56) = 1.34$, $p = 0.186$

Note: 1: Main effect of age, 2: Main effect of condition (congruency) 3: Interaction of age x condition, NR: Not reported

Synthesis of Results

Table 1 summarizes the key characteristics of each study.

For the arrow version, 5 out of 13 studies reported a significant age-related inhibition effect on raw RTs. Three studies applied transformations, such as proportions or log-transformed RTs, to determine whether the effect was truly linked to inhibition, or simply due to age-related slowing in processing speed. Among these 3 studies, 2 found a significant age-related inhibition effect.

For the colored version, both studies found an age-related effect. Surprisingly, one study reported a reversed effect, where young adults exhibited an enhanced flanker effect compared to older adults.

For the cued version, results were more nuanced. Six out of seven studies found a significant age-related effect on raw RTs. Six studies also used transformed RTs, similar to those in the arrow version. However, only two found a significant difference between age groups, suggesting that age-related effects may stem from general slowing of processing speed rather than reduced inhibitory control.

For the letter version, three studies used this variation of the flanker task. Two out of three found a significant age-related effect. Only one study applied transformed RTs, and in that case, the effect remained significant.

Finally, for the moving version, two studies investigated this task variation. Neither found a significant age-related effect, either for raw RTs or transformed scores.

In summary, the results varied depending on the version of the Flanker task used. Some versions, like the arrow and cued tasks, often showed differences in RTs between younger and older adults. However, when researchers adjusted the RTs to account for general slowing with age, these differences were sometimes no longer significant. This suggests that slower processing speed, rather than reduced ability to ignore distractions, might explain some of the results. On the other hand, the letter and color versions showed more consistent differences between age groups, even after adjusting for speed. The moving version did not show any clear age-related differences. Overall, these findings show that the type of task and how the data are analyzed can strongly affect whether age differences are found.

Interference scoring

As we indicated previously, 10 studies used different methods to measure inhibition cost while controlling age-related slowing in processing speed.

One method computed the difference between incongruent and congruent untransformed RTs. This measure provides a direct estimate of the additional time required to respond in the presence of conflicting information, without adjusting for individual differences in processing speed (Townsend & Ashby, 1983).

Another method corrected for overall RT differences by dividing the difference between incongruent and congruent RTs by the mean RT of congruent trials. This approach accounts for individual differences in processing speed and isolates the specific cost of interference. In some cases, this measure was squared to further emphasize the individual differences and ensure all values remain positive.

One commonly used metric was the “Inverse efficiency score” (IES), calculated by dividing the mean RT of correct responses by the proportion of correct responses in each condition. This approach integrates both RT and accuracy, providing a more comprehensive measure of performance. By accounting for both speed and accuracy, IES helps isolate inhibitory control from general age-related slowing (Bruyer & Brysbaert, 2011; Townsend & Ashby, 1983).

Meta-analysis

We were interested in the age-related differences in the flanker task across different stimulus types. Subgroups were created by grouping studies that used the same type of stimuli for effect sizes analysis.

The motion and the colored versions were excluded due to insufficient sample sizes for meaningful analysis, the study of Lemire et al. (2024) was also excluded from the meta-analysis because it was conducted online and at home, while all others studies were conducted in a laboratory setting.

A total of 20 studies divided into three groups were included. Some papers contained multiple studies and each study was identified by a number in parentheses to avoid confusion.

For the arrow version, a small to moderate effect size was found, indicating that older adults exhibited an enhanced flanker effect compared to young adults (g = 0.36, 95% CI [0.25, 0.46]). The results for the arrow group were not significantly heterogeneous (Q = 7.63, pQ = 0.746) which was confirmed using an additional heterogeneity measure (I^2 < 0.00%).

For the cued version, a large effect size was found (g = 0.99, 95% CI [0.35, 1.64]). This group showed significant heterogeneity (Q = 27.25, pQ < 0.001, I^2 = *85.32%).*

For the letter version, a small effect size was found (g = 0.22, 95% CI [-0.11, .55]). This group was significantly heterogeneous (Q = 7.06, pQ = 0.029, I^2= 71.67%*).*

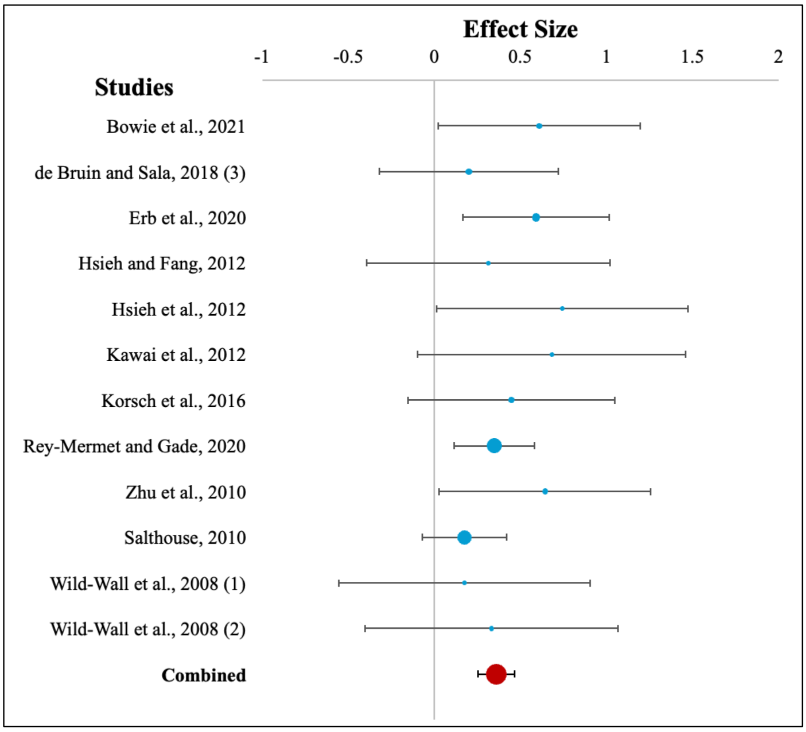

Figure 2. Forest plot of effect sizes for interactions between age and flanker effect for the arrow versions

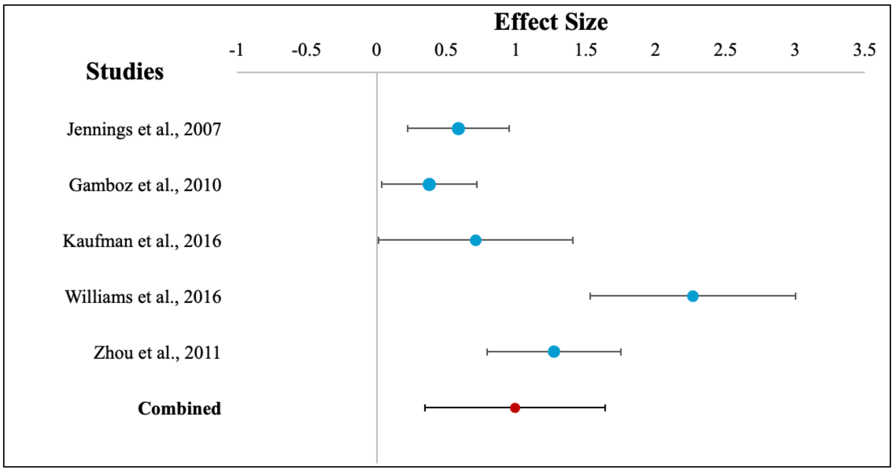

Figure 3. Forst plot of effect sizes for interactions between age and flanker effect in the cued versions

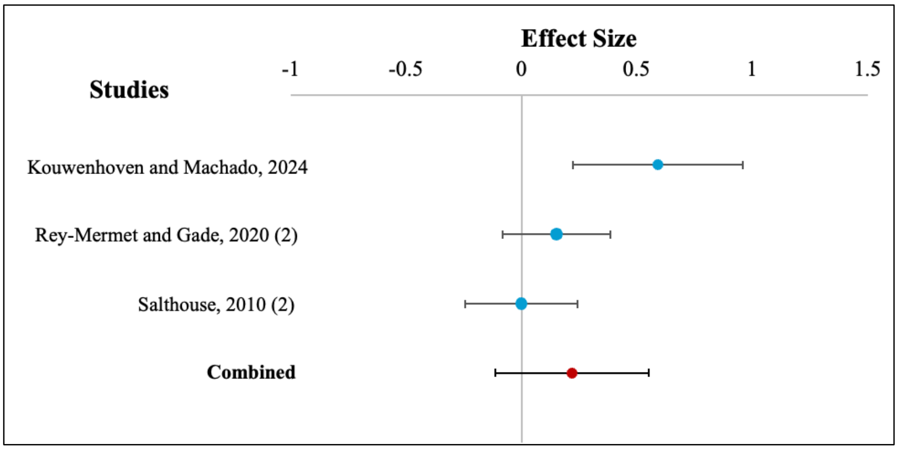

Figure 4. Forest plot of effect sizes for the interactions between age and flanker effect in the letter versions

Discussion

This systematic review and meta-analysis aimed to assess the relationship between age and the flanker effect in different versions of the task. Our findings confirm that aging is generally associated with slower RTs, and, in some cases, a larger interference effect. These results align with existing theories of cognitive aging, emphasizing the importance of task design and scoring methods in detecting age-related differences in inhibitory control. Participant characteristics, task design, methodology, and analysis methods used to examine the interaction between flanker effect and age may partially explain discrepancies in the results.

Crucially, unlike other conflict tasks such as the Stroop or Simon paradigms, the flanker task isolates interference generated by distractors that fall outside the spatial focus of attention. This distinction is essential, as it reveals age-related vulnerabilities in filtering out peripheral information, a mechanism that is less taxed in tasks where distractors and targets are more spatially or semantically integrated. The meta-analytic findings therefore do not merely replicate what is known from general inhibitory decline in aging but instead highlight a specific attentional weakness: older adults' reduced ability to ignore external, spatially distinct distractions. This makes the flanker task a sensitive tool for detecting selective impairments in interference control that are not always observable in other paradigms. Thus, the results of this meta-analysis, therefore, contribute uniquely to the literature by underscoring how aging affects attentional selectivity in spatially complex environments, which are frequent in real-world situations.

Study Selection and Characteristics

A total of 22 studies met the inclusion criteria for this systematic review, spanning from 2007 to 2024. These studies varied significantly in terms of participant demographics, task designs, and methodologies. Among them, 12 studies incorporated both behavioral and physiological measures, such as electroencephalography and functional magnetic resonance imaging, while 10 studies focused exclusively on behavioral data. Most studies explored the effects of aging on inhibitory or interference control, particularly using the flanker task, and were primarily conducted in North America and Western Europe.

The sample size of participants varied widely across studies, ranging from 26 to 302 participants, which may have impacted statistical power and generalizability, with the age spanning from 6 to 87 years. Although the primary focus was on healthy older adults, a few studies also included younger populations and even children, allowing for a broader perspective on age-related changes. However, it is important to note that while some cognitive functions show gradual decline with age, research suggests that executive control abilities, including inhibitory control, tend to remain relatively stable until around 70-75 years old, after which more pronounced declines are observed (Veríssimo et al., 2022). In contrast, the mean age of the older adult groups was generally lower, which may partly explain some differences across studies. Most studies reported mean years of education for each age group and those that included participants with lower education (e.g. Endrass et al., 2012; Gamboz et al., 2010; Zhou et al., 2011) tended to show a stronger age-related effect on inhibitory control, where older adults had more difficulty suppressing distractors. This finding is consistent with the cognitive reserve hypothesis, which suggests that individuals with higher education develop more efficient neural networks and cognitive strategies that help compensate for age-related declines in inhibitory control (Stern, 2009). Higher education levels are also associated with greater exposure to cognitively demanding tasks, which may strengthen executive functions such as attentional control and inhibition. In contrast, individuals with lower education levels may have had fewer opportunities to engage in cognitively stimulating activities, making them more vulnerable to age-related declines.

Cognitive screening was commonly implemented, with most studies excluding participants who exhibited signs of dementia or neurological or psychiatric disorders. However, some studies (Erb et al., 2020; Lemire et al., 2024; Wild-Wall et al., 2008; Zhu et al., 2010) did not use standardized assessments for dementia screening, relying instead on self-reported health declarations. Among studies that screened for cognitive functioning, a variety of assessment tools were used, including the Mini-Mental State Examination, the Mini-Mental Status Test, and the Addenbrooke's Cognitive Examination-III. Variability in cognitive assessment tools may contribute to inconsistencies in findings, as different instruments rely on distinct methodologies, scoring systems, and interpretations. Nonetheless, the inclusion of cognitively healthy individuals enhances the generalizability of findings to typical aging populations.

Flanker Task variations

The Flanker task was used in various forms across the studies, which likely contributed to discrepancies in results. The most common version of the task involved arrow stimuli (e.g., >>>>>) with congruent or incongruent distractors. However, several studies used modified versions, such as letter stimuli (e.g., HHGHH), colored arrows, or even moving stimuli (e.g., de Bruin & Sala, 2018). Different types of stimuli engage distinct cognitive mechanisms. For instance, the colored flanker task introduces an additional early cognitive processing stage, requiring participants to analyze stimulus color before responding (Korsch et al. 2016). This increased complexity alters cognitive demand and may lead to performance differences between age groups.

Stimulus timing and presentation also varied considerably across studies. While most studies (19 out of 29, accounting for multiple studies per paper) presented the target and distractors simultaneously, others (e.g., Di Chiaro & Holmes, 2024) introduced a delay between the presentation of distractors and the target. This distinction is crucial because presenting distractors before the target allows participants to preprocess visual information and anticipate their response. In contrast, simultaneous presentation requires participants to process all elements at once, potentially increasing interference effects.

To better understand these discrepancies, we conducted a meta-analysis on the three versions of the task (arrows, cued, and letters). The results confirm that age affects the flanker effect differently depending on the stimulus type. The only condition that was not significantly heterogeneous was the arrow version. This can be explained by the widespread use of this version, which has led to more standardized methodologies, including presentation time, number of trials, and task instructions.

The effect size of the flanker effect for the arrow version was small to moderate, suggesting that an age-related effect is present when using raw RTs. Arrows are shapes with a universal directional meaning, commonly encountered in daily life (Ristic & Kingstone, 2006). This version of the flanker task therefore primarily relies on automatic processing, which tends to remain intact during aging and likely explains why older adults still struggle with suppressing interference, particularly when distractors are positioned close to the target (Wild-Wall et al., 2011). Perceptual characteristics, such as the spatial distance between stimuli, play a crucial role in modulating interference. When distractors are closer to the target, they are more likely to be processed automatically, increasing the likelihood of interference (Ridderinkhof et al., 2021). Conversely, greater spatial separation reduces competition between stimuli, making it easier to focus on the relevant information (Eriksen & Eriksen, 1974; Maylor & Lavie, 1998). This sensitivity to spatial proximity may be particularly relevant in aging, as older adults often experience declines in selective attention and inhibitory control, leading to stronger interference effects when stimuli are closely spaced.

Four studies applied transformed RT scores to account for age-related slowing. In two of these, the age-related effect on the flanker effect remained significant.

Interestingly, the findings from the arrow version of the flanker task were successfully replicated in an unsupervised, at-home environment, suggesting it can be effectively administered online, expanding research possibilities beyond laboratory settings.

The cued version of the flanker task employs different cognitive mechanisms compared to the arrow version. In fact, the cued version, also known as the Attentional Network Task, is specifically designed to assess three distinct attentional networks: alerting, orientating, and executive control (Fan et al., 2002). These networks collectively influence the dynamics of the task and add another level of processing (Ridderinkhof et al., 2021). When using raw RTs, studies reported an age-related effect on the flanker effect, suggesting that even when cues were presented, older adults were more affected by the incongruent trials compared to younger adults. However, one study (Williams et al., 2016) stands out as an outlier potentially lowering the observed effect size. Interestingly, when researchers used transformed RTs to correct for age-related slowing, four studies found that the flanker effect in the executive control part of the task was no longer significant. This suggests that the age-related difference initially observed might be attributed to general slowing in processing speed rather than inhibitory control difficulties. However, when examined on a physiological level, findings suggest that older adults may be using mechanisms to compensate for their difficulties. In fact, Hsieh and Fang (2012) found age-related differences on several ERP components. For instance, N1, a component associated with sustained covert attention, was increased for the central target in older adults, suggesting that they paid more attention to the target increased the visual processing in a top-down way, while limiting the flanker processing (Wild-Wall et al., 2008). Conversely, N2, a component associated with response-related conflict, was reduced in older adults (Hsieh & Fang, 2012; Wild-Wall et al., 2008). Older adults may have reduced conflict by increasing the processing of central targets and paying more attention to the stimuli (Wild-Wall et al., 2008).

As for the letter version of the task, results were less consistent. In fact, the range of effect sizes calculated in this subgroup varied between 0.00 and 0.59, hence the small combined effect size. This variability can be explained by differences in experimental design across studies, as they did not use the same protocol. In fact, Kouwenhoven and Machado (2024) used letters displayed along a vertical axis, with only one distractor positioned either above or below the target. This approach aimed to minimize the effects of spatial compatibility with the response buttons, which were displayed along a horizontal axis.

One possible explanation for these inconsistencies is the way spatial attention and response selection are engaged. The horizontal presentation is more commonly used in reading and other everyday visual tasks, which may facilitate interference resolution, especially in older adults (Hsieh et al., 2012). Additionally, when stimuli are aligned horizontally, there is a direct spatial correspondence between stimulus position and response buttons, which may reduce cognitive load and enhance response efficiency. In contrast, when stimuli are presented vertically, this automatic mapping is absent, requiring greater cognitive control to inhibit interference. This increased demand on inhibitory control may explain why age-related differences are more pronounced in the vertical presentation condition (de Bruin & Sala, 2018; Wild-Wall et al., 2008).

Additionally, the number of experimental trials varied considerably, ranging from 80 to 1200 trials, with some studies incorporating practice trials of varying lengths. Studies also differed in the proportion of congruent versus incongruent trials, which likely influenced the magnitude of the interference effect and contributed to variability across studies. While the majority of the studies used a balanced 50/50 design, others employed unbalanced proportions, such as blocks with a high proportion of congruent trials (e.g Di Chiaro & Holmes, 2024; Hsieh & Fang, 2012; Reuter et al., 2017). This design variation is particularly relevant, as it modulates participants' expectations and engagement of cognitive control.

For example, Di Chiaro and Holmes (2024) included both response-balanced and response-imbalanced blocks, allowing for a more nuanced analysis. While the overall results did not show an opposite pattern when using combined efficiency scores, further separate analyses revealed an interaction between age and congruency proportion specifically for RTs (but not for accuracy). Older adults showed a relative advantage in the response-imbalanced condition, which may reflect a differential sensitivity to proportion manipulation. These results highlight how the choice of proportion and the outcome variable used can significantly alter the observed effects.

This finding underscores the need for caution when interpreting interference effects across studies, as congruency proportion appears to act as a source of heterogeneity in task demands, rather than reflecting consistent age-related differences per se. When a high proportion of trials are congruent, participants expect that most trials will be easy, leading to reduced cognitive control and increased reliance on automatic responses. As a result, when an incongruent trial unexpectedly appears, the interference effect is larger, particularly in older adults, who exhibit greater reliance on habitual responses and slower reactive control (Mutter et al., 2005). Unlike younger adults, who can dynamically modulate cognitive control, older adults experience persistent interference due to a reduced ability to disengage from dominant responses, making them more vulnerable to conditions with a high proportion of congruent trials (Bugg et al., 2008).

One study (de Bruin & Sala,2018) even included a moving stimulus version of the task, providing a novel approach to understanding how motion-based stimuli might differentially impact inhibitory control across age groups. This study did not find a significant age-related effect, but further investigation into this approach could yield valuable insights.

Analysis and Results

A variety of statistical techniques were employed across the studies, with analysis of variance (ANOVA) being the most commonly used method to assess the interaction of flanker congruency and age. Seven studies also employed interference scoring to specifically measure inhibitory control, often by calculating the difference in RTs between congruent and incongruent trials. Additionally, some studies utilized more sophisticated methods, such as Bayesian analysis (e.g., Rey-Mermet & Gade, 2020) to evaluate the strength of the evidence for both null and alternative hypotheses.

An important source of heterogeneity across studies stems from the variability in how performance was measured and analyzed. While most studies relied primarily on RT as the dependent variable, others used accuracy (Rey-Mermet & Gade, 2020; Wild-Wall et al., 2008; Zhu et al., 2010b), or combined both into a score such as the IES (e.g. Bowie et al., 2021; Di Chiaro & Holmes, 2024). This inconsistency in the choice of outcome variable can significantly influence the observed effects. For instance, because older adults tend to slow down to maintain accuracy, some researchers prioritized accuracy as a more reliable indicator of performance, particularly when concerned about age-related declines in processing speed. However, this variation in measurement approach leads to difficulties in comparing findings across studies, and may partly explain why some studies report robust age-related differences in inhibitory control, while others do not. In particular, studies emphasizing RTs may find stronger interference effects among older adults (Endrass et al., 2012), whereas those focusing on accuracy or IES may yield more nuanced or inconsistent results.

In terms of results, most studies, such as Bowie et al. (2021), Di Chiaro and Holmes (2024), and Endrass et al. (2012), found a significant increase in RTs among older adults compared to younger adults (especially during incongruent trials), regardless of the task design. These findings support the well-documented age-related decline in general processing speed, consistent with theories such as the processing-speed theory of cognitive aging. Significant age-related differences in inhibitory control were found in nearly half of the studies, such as Lemire et al. (2024) and Zhu et al. (2010), where older adults exhibited larger interference effects than younger participants. However, findings were more mixed when it came to accuracy. While some studies, like Endrass et al. (2012) and Reuter et al., (2017) reported significant differences in error rates between younger and older adults, others, including Bowie et al. (2021), Hsieh and Fang (2012) and Jennings et al. (2007), found no age-related changes in error rates. These mixed results further reflect the methodological variability described above and suggest that older adults may adopt a strategy of slowing down to maintain accuracy, as seen in the findings of Hsieh and Fang (2012). This speed-accuracy trade-off allows older adults to compensate for age-related declines in cognitive processing speed and inhibitory control by taking more time to process stimuli and select the correct response, thereby reducing errors. By prioritizing accuracy over speed, older adults can partially offset the effects of slower neural processing, aligning with the compensatory hypothesis of aging. While this strategy helps preserve performance in accuracy-based tasks, it may not always be advantageous in real-world situations where both speed and accuracy are crucial, such as driving or rapid decision-making under time constraints.

Methods to measure interference effect

There has been a variety of basic statistical methods used to assess the interference effect, such as T-tests and ANOVAs. On the other hand, some studies, such as Bowie et al. (2021), preferred calculating a score in order to mitigate the effect of age-related slowing in processing speed, which may explain to some extent the variability in results. Most studies employed a simple measure by computing the raw difference between incongruent and congruent RTs. The majority of studies using this method reported that the age-related effect persisted (e.g., Lemire et al., 2024; Zhu et al., 2010). However, this measure does not account for individual differences in overall processing speed and accuracy, which is why normalized measures, such as dividing by congruent RTs or squaring the result, are often preferred.

Studies like Bowie et al. (2021) and Di Chiaro and Holmes (2024) used the IES, which accounts for both RTs and accuracy. Since older adults tend to have longer RTs due to general cognitive slowing, using raw RTs alone may overestimate inhibitory difficulties. The IES mitigates this issue by normalizing RTs based on accuracy, allowing for a more accurate assessment of inhibitory abilities across age groups. This method provides a more refined measure of cognitive efficiency, suggesting that older adults are not only slower but also less efficient in handling interference. These findings imply that traditional RT measures might not fully capture the extent of these deficits in older adults.

Moreover, the use of Bayesian statistical methods can offer complementary insights by quantifying the strength of evidence for or against the presence of an effect, rather than relying solely on traditional significance testing.

Conclusion

In conclusion, this systematic review and meta-analysis highlights the significant impact of aging on inhibitory control, particularly in terms of longer RTs and increased interference effects. However, the mixed findings on accuracy and the impact of task design underscore the complexity of cognitive aging. Differences in stimulus types cognitive screening methods, and statistical approaches contribute to the variability observed across studies.

Future research should focus on standardizing task protocols and implementing robust cognitive screening measures to improve the comparability of findings. Additionally, further exploration of novel task adaptations, such as moving stimuli, could provide valuable insights into how dynamic environments influence inhibitory control across the lifespan. By addressing these gaps, future studies can contribute to a more comprehensive understanding of cognitive aging and its effects on inhibitory control.

Funding

S.G. is a recipient of scholarships from the Research, Social Sciences and Humanities Research Council of Canada (SSHRC) and Fonds de recherche du Québec (FRQ) Société et culture. S.R. is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC; RGPIN-2020-06706). B.B. is supported by the FRQ Société et culture (2023-NP-311368).

References

Bowie, D. C., Low, K. A., Fabiani, M., & Gratton, G. (2021). Event-related brain potentials reveal strategy selection in younger and older adults. *Biological Psychology*, *164*. psyh. https://doi.org/10.1016/j.biopsycho.2021.108163

Bruyer, R., & Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a Better Dependent Variable than the Mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychologica Belgica*, *51*(1), 5. https://doi.org/10.5334/pb-51-1-5

Bugg, J. M., Jacoby, L. L., & Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Memory & Cognition*, *36*(8), 1484–1494. https://doi.org/10.3758/MC.36.8.1484

de Bruin, A., & Sala, S. D. (2018). Effects of age on inhibitory control are affected by task-specific features. *The Quarterly Journal of Experimental Psychology*, *71*(5), 1219–1233. psyh. https://doi.org/10.1080/17470218.2017.1311352

Di Chiaro, N. V., & Holmes, N. P. (2024). Flanker interference at both stimulus and response levels decreases with age. *Experimental Brain Research*, *242*(3), 757–767. psyh. https://doi.org/10.1007/s00221-023-06773-9

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, *12*(10), 374–380. https://doi.org/10.1016/j.tics.2008.07.001

Endrass, T., Schreiber, M., & Kathmann, N. (2012). Speeding up older adults: Age-effects on error processing in speed and accuracy conditions. *Biological Psychology*, *89*(2), 426–432. https://doi.org/10.1016/j.biopsycho.2011.12.005

Erb, C. D., Welhaf, M. S., Smeekens, B. A., Moreau, D., Kane, M. J., & Marcovitch, S. (2021). Linking the dynamics of cognitive control to individual differences in working memory capacity: Evidence from reaching behavior. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(9), 1383–1402. https://doi.org/10.1037/xlm0001018

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the Efficiency and Independence of Attentional Networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fu, J., Yu, G., & Zhao, L. (2021). Effect of aging on visual attention: Evidence from the Attention Network Test. *Social Behavior and Personality: An International Journal*, *49*(3), 1–8. https://doi.org/10.2224/sbp.9806

Galvao-Carmona, A., González-Rosa, J. J., Hidalgo-Muñoz, A. R., Páramo, D., Benítez, M. L., Izquierdo, G., & Vázquez-Marrufo, M. (2014). Disentangling the attention network test: Behavioral, event related potentials, and neural source analyses. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00813

Gamboz, N., Zamarian, S., & Cavallero, C. (2010). Age-related differences in the Attention Network Test (ANT). *Experimental Aging Research*, *36*(3), 287–305. psyh. https://doi.org/10.1080/0361073X.2010.484729

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. https://doi.org/10.1038/nn1543

Grady, C., & Craik, F. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, *10*(2), 224–231. https://doi.org/10.1016/S0959-4388(00)00073-8

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The Psychology of Learning and Motivation*, *22*.

Hsieh, S., & Fang, W. (2012). Elderly adults through compensatory responses can be just as capable as young adults in inhibiting the flanker influence. *Biological Psychology*, *90*, 113–126.

Hsieh, S., Liang, Y., & Tsai, Y.-C. (2012). Do age-related changes contribute to the flanker effect? *Clinical Neurophysiology*, *123*, 960–972.

Jennings, J. M., Dagenbach, D., Engle, C. M., & Funke, L. J. (2007). Age-related changes and the attention network task: An examination of alerting, orienting, and executive function. *Aging, Neuropsychology, and Cognition*, *14*(4), 353–369. psyh. https://doi.org/10.1080/13825580600788837

Kaufman, D. A. S., Sozda, C. N., Dotson, V. M., & Perlstein, W. M. (2016). An Event-Related Potential Investigation of the Effects of Age on Alerting, Orienting, and Executive Function. *Frontiers in Aging Neuroscience*, *8*. https://doi.org/10.3389/fnagi.2016.00099

Kawai, N., Kubo-Kawai, N., Kubo, K., Terazawa, T., & Masataka, N. (2012). Distinct aging effects for two types of inhibition in older adults: A near-infrared spectroscopy study on the Simon task and the flanker task. *NeuroReport: For Rapid Communication of Neuroscience Research*, *23*(14), 819–824. psyh. https://doi.org/10.1097/WNR.0b013e3283578032

Kok, A. (1999). Varieties of inhibition: Manifestations in cognition, event-related potentials and aging. *Acta Psychologica*, *101*(2–3), 129–158. https://doi.org/10.1016/S0001-6918(99)00003-7

Korsch, M., Frühholz, S., & Herrmann, M. (2016). Conflict-specific aging effects mainly manifest in early information processing stages—An ERP study with different conflict types. *Frontiers in Aging Neuroscience*, *8*. psyh. https://biblioproxy.uqtr.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2016-13706-001&site=ehost-live

Kouwenhoven, M., & Machado, L. (2024). Age differences in inhibitory and working memory functioning: Limited evidence of system interactions. *Aging, Neuropsychology, and Cognition*, *31*(3), 524–555. psyh. https://doi.org/10.1080/13825585.2023.2214348

Kramer, A. F., Hahn, S., & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica*, *101*(2), 339–378. https://doi.org/10.1016/S0001-6918(99)00011-6

Kramer, A. F., Humphrey, D. G., Larish, J. F., & Logan, G. D. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging*, *9*(4), 491–512. https://doi.org/10.1037/0882-7974.9.4.491

Lemire, M., Soulières, I., & Saint-Amour, D. (2024). The effect of age on executive functions in adults is not sex specific. *Journal of the International Neuropsychological Society*. psyh. https://doi.org/10.1017/S1355617723011487

Ludwig, C., Borella, E., Tettamanti, M., & de Ribaupierre, A. (2010). Adult age differences in the Color Stroop Test: A comparison between an Item-by-item and a Blocked version. *Archives of Gerontology and Geriatrics*, *51*(2), 135–142. https://doi.org/10.1016/j.archger.2009.09.040

Maylor, E. A., & Lavie, N. (1998). The influence of perceptual load on age differences in selective attention. *Psychology and Aging*, *13*(4), 563–573. https://doi.org/10.1037//0882-7974.13.4.563

Mutter, S. A., Naylor, J. C., & Patterson, E. R. (2005). The effects of age and task context on Stroop task performance. *Memory & Cognition*, *33*(3), 514–530. https://doi.org/10.3758/BF03193068

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Proctor, R. W., Pick, D. F., Vu, K.-P. L., & Anderson, R. E. (2005). The enhanced Simon effect for older adults is reduced when the irrelevant location information is conveyed by an accessory stimulus. *Acta Psychologica*, *119*(1), 21–40. https://doi.org/10.1016/j.actpsy.2004.10.014

Reuter, E.-M., Voelcker-Rehage, C., Vieluf, S., Lesemann, F. H. P., & Godde, B. (2017). The P3 Parietal-To-Frontal Shift Relates to Age-Related Slowing in a Selective Attention Task. *Journal of Psychophysiology*, *31*, 49–66.

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review*, *25*(5), 1695–1716. psyh. https://doi.org/10.3758/s13423-017-1384-7

Rey-Mermet, A., & Gade, M. (2020). Age-related deficits in the congruency sequence effect are task-specific: An investigation of nine tasks. *Psychology and Aging*, *35*(5), 744–764. https://doi.org/10.1037/pag0000414

Ridderinkhof, K. R., Forstmann, B. U., Wylie, S. A., Burle, B., & van den Wildenberg, W. P. M. (2011). Neurocognitive mechanisms of action control: Resisting the call of the Sirens. *Wiley Interdisciplinary Reviews. Cognitive Science*, *2*(2), 174–192. https://doi.org/10.1002/wcs.99

Ridderinkhof, K. R., Wylie, S. A., van den Wildenberg, W. P. M., Bashore, T. R., & van der Molen, M. W. (2021). The arrow of time: Advancing insights into action control from the arrow version of the Eriksen flanker task. *Attention, Perception, & Psychophysics*, *83*(2), 700–721. https://doi.org/10.3758/s13414-020-02167-z

Ristic, J., & Kingstone, A. (2006). Attention to Arrows: Pointing to a New Direction. *Quarterly Journal of Experimental Psychology*, *59*(11), 1921–1930. https://doi.org/10.1080/17470210500416367

Salthouse, T. A. (1994). The nature of the influence of speed on adult age differences in cognition. *Developmental Psychology*, *30*(2), 240–259. https://doi.org/10.1037/0012-1649.30.2.240

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2010). Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. *Brain and Cognition*, *73*(1), 51–61. psyh. https://doi.org/10.1016/j.bandc.2010.02.003

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*(5), 763–776. https://doi.org/10.1037/0012-1649.27.5.763

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*(10), 2015–2028. https://doi.org/10.1016/j.neuropsychologia.2009.03.004

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of: A free and simple tool for meta-analysis. *Research Synthesis Methods*, *8*(4), 537–553. https://doi.org/10.1002/jrsm.1260

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic Modeling of Elementary Psychological Processes*. CUP Archive.

van der Lubbe, R. H. J., & Verleger, R. (2002). Aging and the Simon task. *Psychophysiology*, *39*(1), 100–110. https://doi.org/10.1111/1469-8986.3910100

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*(3), 174–180. psyh. https://doi.org/10.1177/0963721411408772

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience & Biobehavioral Reviews*, *26*(7), 849–857. https://doi.org/10.1016/S0149-7634(02)00071-4

Veríssimo, J., Verhaeghen, P., Goldman, N., Weinstein, M., & Ullman, M. T. (2022). Evidence that ageing yields improvements as well as declines across attention and executive functions. *Nature Human Behaviour*, *6*(1), 97–110. https://doi.org/10.1038/s41562-021-01169-7

Wild-Wall, N., Falkenstein, M., & Hohnsbein, J. (2008). Flanker interference in young and older participants as reflected in event-related potentials. *Brain Research*, *1211*, 72–84.

Williams, R. S., Biel, A. L., Wegier, P., Lapp, L. K., Dyson, B. J., & Spaniol, J. (2016). Age differences in the Attention Network Test: Evidence from behavior and event-related potentials. *Brain and Cognition*, *102*, 65–79. https://doi.org/10.1016/j.bandc.2015.12.007

Wylie, G. R., Yao, B., Sandry, J., & DeLuca, J. (2020). Using Signal Detection Theory to Better Understand Cognitive Fatigue. *Frontiers in Psychology*, *11*, 579188. https://doi.org/10.3389/fpsyg.2020.579188

Zhou, S., Fan, J., Lee, T. M. C., Wang, C., & Wang, K. (2011). Age-related differences in attentional networks of alerting and executive control in young, middle-aged, and older Chinese adults. *Brain and Cognition*, *75*(2), 205–210. https://doi.org/10.1016/j.bandc.2010.12.003

Zhu, D. C., Zacks, R. T., & Slade, J. M. (2010a). Brain activation during interference resolution in young and older adults: An fMRI study. *NeuroImage*, *50*(2), 810–817. psyh. https://doi.org/10.1016/j.neuroimage.2009.12.087

Zhu, D. C., Zacks, R. T., & Slade, J. M. (2010b). Brain activation during interference resolution in young and older adults: An fMRI study. *NeuroImage*, *50*(2), 810–817. https://doi.org/10.1016/j.neuroimage.2009.12.087