



HAL
open science

Exercise caution: some methodological and theoretical considerations when examining improvements in attentional selectivity following exercise

Nadège Rochat, Pierpaolo Iodice, Ludovic Ferrand, Benjamin Parris, Maria Augustinova

► **To cite this version:**

Nadège Rochat, Pierpaolo Iodice, Ludovic Ferrand, Benjamin Parris, Maria Augustinova. Exercise caution: some methodological and theoretical considerations when examining improvements in attentional selectivity following exercise. *International Review of Sport and Exercise Psychology*, 2024, pp.1-22. 10.1080/1750984X.2024.2431506 . hal-04808026

HAL Id: hal-04808026

<https://hal.science/hal-04808026v1>

Submitted on 27 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Exercise caution:

Some methodological and theoretical considerations when examining improvements in attentional selectivity following exercise

Nadège Rochat^{1,2,3*}, Pierpaolo Iodice^{2,4}, Ludovic Ferrand⁵, Benjamin A. Parris⁶, and Maria Augustinova^{1*}

¹Normandie Université, UNIROUEN, CRFDP, UR 7475, 76000, Rouen, France

²Normandie Université, UNIROUEN, CETAPS, UR 3832, 76000, Rouen, France

³Université Clermont Auvergne, ACTé, UR 4281, 63000 Clermont-Ferrand, France

⁴Le Mans Université, MIP, UR 4334, 72000 Le Mans, France

⁵Université Clermont Auvergne, CNRS, LAPSCO, UMR 6024, 63000 Clermont-Ferrand, France

⁶Department of Psychology, Bournemouth University, Poole, UK

*Corresponding authors:

Nadège Rochat

Université Clermont Auvergne

Laboratoire Activité, Connaissance, Transmission, Éducation (ACTé, UR 4281),

INSPÉ Clermont-Auvergne, 36 avenue Jean-Jaurès CS 20001, 63407 Chamalières Cedex, FRANCE

E-mail: nadege.rochat@uca.fr

Phone: + 33 473 318 769

or

Maria Augustinova

Université de Rouen Normandie

Le Centre de Recherches sur les Fonctionnements et Dysfonctionnements Psychologiques (CRFDP, EA 7475),

UFR SHS – Bâtiment Freinet, Place Emile Blondel, 76821 Mont Saint Aignan Cedex, FRANCE

E-mail: maria.augustinova@univ-rouen.fr

Phone: + 33 473 406 257

Fax: + 33 473 406 114

Word count (main text and subheadings): 6186

45 **Abstract**

46 Numerous studies investigating the effects of a given laboratory or field intervention
47 on attentional selectivity have evidenced their (positive) findings with improved performance
48 in the Stroop task. In contrast to conclusions drawn in these studies, the present paper argues
49 that the extent to which this improvement actually reflects an attentional/inhibitory boost is
50 still relatively unknown. To evidence this claim, this paper first delimited conditions under
51 which Stroop performance indicators can directly measure selective attention and its changes.
52 It then assessed the use of these indicators in a specific and exemplary body of translational
53 research investigating the link (correlational or causal) between physical activity/exercise and
54 attentional selectivity with a result that only nine past studies (out of 52) actually used
55 relevant indicators of physical activity/exercise-related changes in selective attention. The
56 narrative review of these nine studies highlighted a substantial gap between what this body of
57 research shows and the conclusions it draws. To mitigate this gap that is likely to concern all
58 types of research aimed at boosting attentional selectivity, the paper also provided practical
59 guidelines to allow future studies employing the Stroop task to draw more theoretically and
60 empirically grounded inferences about the effects of any intervention on attentional
61 selectivity.

62

63 **Keywords**

64 Inhibition; Measure of interference; Physical activity; Selective attention; Stroop task

65 **Introduction**

66 Current societal changes – including population ageing, post-pandemic adoption of
67 work-from-home policies and excessive exposure to screen time – raise concerns about
68 individuals’ cognitive abilities and their preservation and replenishment. Therefore,
69 unsurprisingly, there is a growing body of research aimed at identifying lifestyles or practices
70 that are associated with enhanced cognitive abilities (e.g., level of physical fitness, amounts of
71 daily mindfulness meditation practices). Additionally, interventions that effectively preserve
72 or even boost individuals’ cognitive functioning – either immediately (e.g., a short exposure
73 to nature) or after a longer period of time (e.g., several weeks of a cognitive training program)
74 have been investigated.

75 Regardless of both the type of intervention under scrutiny (e.g., physical activity,
76 mindfulness meditation, exposure to nature, cognitive training etc.) and the nature of its link
77 to cognitive abilities (i.e., correlational vs. causal), this type of translational research
78 frequently involves the use of selective attention tasks such as the Stroop task (Stroop, 1935).
79 The Stroop task is sometimes included in a larger battery of cognitive tests or supplemented
80 by other standardized or standard cognitive tasks (see e.g., Cremen & Carson, 2017 for a
81 discussion), but frequently, the efficiency of the key components of cognitive functioning
82 such as “interference control”, “selective attention and/or the inhibition it entails” or
83 “executive functions” (i.e., a supraordinate family of cognitive processes that encompasses
84 attentional control and inhibition it entails) are assessed by the means of the Stroop task alone
85 (see e.g., Buck et al., 2008; Ishihara et al., 2018; and Best et al., 2014 for examples of these
86 respective endeavors). This is not surprising since the standard (i.e., color-word) Stroop task
87 is generally considered the “gold standard of attention measures” (MacLeod, 1992) as well as
88 “a prototypical inhibition task” (Miyake et al., 2000, p. 57). But at the risk of stating the

89 obvious, it is important to remember that the deployment of this task itself, does not guarantee
90 that this gold standard is met.

91 To provide further and hopefully more explicit grounds for the necessity of this
92 reminder, the present paper first attempts to delimit conditions that are the most favorable for
93 measuring selective attention (section “Measuring attentional selectivity in the Stroop task”)
94 and its changes (section “Assessing the link between any given intervention and attentional
95 selectivity measured in the Stroop task”). Building on Furley and colleagues (2023 in this
96 journal), the paper then assesses the extent to which these specific measurement conditions
97 are met in research examining the link between physical activity/exercise and attentional
98 abilities (see Table 1 and Tables 1S and 2S in Supplementary Materials for all assessed
99 studies). As a result of this endeavor, this section (“Assessing the link between physical
100 activity and attentional selectivity measured in the Stroop task”) argues that there is a
101 substantial gap between what this specific body of research shows and the conclusions it
102 draws about the attentional benefits of physical activity/exercise. But because this gap is not
103 limited to this particular body of research, and rather concerns all types of translational
104 research aimed at boosting selective attention and/or the inhibition it entails, the paper then
105 provides practical guidelines designed to mitigate this gap (section “Practical guidelines for a
106 more direct measure of selective attention in the Stroop task”). Indeed, Furley and colleagues
107 (2023 in in this journal) have argued that in addition to adopting a more mechanistic approach
108 (see also e.g., Diamond, 2020; Smid et al., 2020), this mitigation largely requires progress in
109 conceptualization and measurement of indicators used to appraise intervention-related
110 changes in cognitive abilities (see also e.g., Cremen & Carson, 2017). In line with this latter
111 idea – echoing directly an ongoing debate in the Stroop literature (e.g., Algom et al., 2022;
112 Chajut & Algom, 2019; Parris et al., 2022; Parris et al., 2023) –, the most important goal of
113 this paper is to help future studies using the Stroop task to make more theoretically and

114 empirically grounded inferences about the effects of any given intervention on attentional
115 selectivity and the inhibition it entails.

116

117 **Measuring attentional selectivity in the Stroop task**

118 Given the number of studies that have employed the Stroop task since its inception in
119 1935 by John Ridley Stroop (1897-1973), it is no surprise that many different variants of the
120 original task have been employed. Still, whether it is administered in a paper/pencil (i.e., list)
121 or in a computerized (i.e., item-by-item) form, the standard (i.e., color-word) Stroop task
122 (Stroop, 1935) requires individuals to identify – as quickly and as accurately possible – the
123 color of the font in which a word is presented whilst ignoring the meaning of the word itself.
124 Despite these instructions, individuals experience difficulties identifying the ink color of a
125 word when it denotes a different color (e.g., “BLUE” presented in red ink, hereafter *BLUE*_{red}).
126 This is because “The conflict between the relevant (color of the word) and irrelevant (name of
127 the word) dimensions of the stimulus on *incongruent trials* presents a particularly difficult
128 task for the selective attentional system (Spieler et al., 1996, p. 461; italics added).”
129 Consequently, color-identification of incongruent trials is substantially longer and more prone
130 to errors compared to a relevant baseline. It is this difference – called *the Stroop effect* or the
131 Stroop *interference* effect – that is used to assess the efficiency of selective attention with a
132 larger magnitude of Stroop interference being equated with lower attentional selectivity and
133 lower interference control it entails (e.g., Chajut & Algom, 2019; MacLeod, 1998; Parris et
134 al., 2022).

135 But before elaborating further on conceptual difficulties surrounding this measure of
136 selective attention, it seems important to recall some basic methodological principles, starting
137 with the fact that both speed (typically measured by RTs) and accuracy (typically measured
138 by errors-rates), need to be considered simultaneously to understand performance in the

139 Stroop task (Heitz, 2014; Vandierendonck, 2021), unless the time to complete a given list of
140 items is kept constant¹. Indeed, when only one of these two indicators is considered (i.e.,
141 measured, analyzed and/or reported; see Table 2S in Supplementary Materials for examples
142 of this selective consideration in sport and exercise research using the Stroop task), it is
143 impossible to exclude the possibility that the reported speed is not traded for accuracy, nor
144 inversely, that the reported accuracy is not traded for speed. Simply put, while the elimination
145 of this tradeoff is important for the interpretation of processes driving the Stroop effect itself,
146 and/or its changes (see here below), simultaneous consideration of both indicators allows the
147 verification that the task instructions – emphasizing both accuracy and speed (see above) –
148 were followed.

149 Another basic, yet important methodological aspect to bear in mind is that the
150 observed levels of speed and accuracy must be assessed with respect to a relevant baseline
151 (again see Table 2S in SM for examples of sport and exercise research where this type of
152 comparison is clearly missing). Now, selecting an appropriate baseline in the Stroop task is
153 itself non-trivial (see Parris et al., 2022 for discussion; see Figure 1 for a diagrammatic
154 representation of the complexity of baseline selection). For example, color-congruent trials
155 (e.g., *BLUE*_{blue}), first introduced by Dalrymple-Alford and Budayr (1966, Exp. 2), have
156 become a popular baseline condition against which performance on color-incongruent (i.e.,
157 conflict or interference) trials is compared (see Table 1S in SM for examples of this type of
158 baseline in sport and exercise research). But it is necessary to understand that the ink color of
159 color-congruent items is identified reliably faster than that of color-neutral trials (e.g., non-
160 color word “DEAL” presented in red ink, *DEAL*_{red}, Dalrymple-Alford & Budayr, 1966, for
161 the first demonstration) – leading to an effect later termed as *facilitation* (Dalrymple-Alford,
162 1972).

¹ This the case in some paper/pencil (i.e., list) versions of the Stroop task where the time to complete a given list is restricted to 45 sec for instance (see e.g., Tables 1) for examples of such versions of the Stroop).

163 Consequently, without the inclusion of color-neutral trials, the measure of the Stroop
164 effect as a difference between color-incongruent and color-congruent trials
165 (e.g., $BLUE_{red} - BLUE_{blue}$) – sometimes referred to as the Stroop congruency effect or the
166 overall Stroop effect – confounds the interfering effect experienced on color-incongruent
167 trials with the opposing effect of facilitation on color-congruent trials. It represents therefore
168 “(...) the sum of facilitation and interference, each in unknown amounts” (MacLeod, 1991,
169 p.168). Moreover, when including a color-neutral baseline, it is clear that their magnitudes are
170 often asymmetrical (with facilitation being generally smaller than interference, MacLeod &
171 MacDonald, 2000, see also e.g., Smith & Ulrich, 2023 for a more general discussion of this
172 “violation of the midpoint assumption” in several other conflict or interference tasks).

173 Furthermore, facilitation on color-congruent trials is known to increase when
174 experimental setups encourage contingent associative learning between words and responses,
175 or words and colors, as is the case when using equal proportions of color-congruent and color-
176 incongruent items (see Hasshim & Parris, 2021; Schmidt, 2013 and see Table 1 and Tables 1S
177 and 2S in Supplementary Materials, SM for examples of sport and exercise research using this
178 equal proportion). Since the numbers of incongruent and congruent stimuli are equal, the
179 word e.g., *red* more commonly predicts the response red than any other color. This is because
180 there are only four possible congruent stimuli (assuming four response colors). If congruent
181 trials are as numerous as incongruent trials (e.g., 72 of each trial type), each congruent
182 stimulus must be repeated more times than each incongruent stimulus. This contingency
183 between the word and response leads to faster responses for the contingent congruent trials.

184 To sum up, the magnitude (or amount) of *interference* produced by color-incongruent
185 trials – considered as prime indicator of attentional abilities – cannot be reliably assessed
186 when using color-congruent trials as a unique comparison baseline (see Figure 1, and see

187 Algom et al., 2022; Algom & Chajut, 2019; MacLeod, 1991; Parris et al., 2022; Quétard et
188 al., 2023, for discussions of this issue).

189

190 *Insert Figure 1 near here*

191 Figure 1. Contribution of facilitation, lexicality cost and interference effects to the overall
192 Stroop/congruency effect.

193 *Note.* Original figure based on data from Augustinova et al. (2019) reported with Verbal
194 (Panel A) and Manual (Panel B) response modality.

195

196 In order to measure interference on color-incongruent trials independently of
197 facilitation on color-congruent trials, some studies used either color-patches or colored dots,
198 as well as rows of signs or letters (e.g., colored Xs) as a comparison baseline (see Table 1 for
199 examples). As an anecdote, to derive interference – central in his work – J.R. Stroop himself,
200 used rows of colored Sanskrit swastikas as a comparison baseline (Stroop, 1935, Exp.3, see
201 also Exp.2 for the use of full squares). However, even with these different choices, the
202 magnitude of interference is still potentially overestimated (see Brown, 2011 for an ample
203 demonstration), since a row of letters (e.g., “XXXX”), is not a word which, unlike “DEAL”
204 for instance, has a meaning. As a result, color identification of color-neutral non-word items
205 (e.g., *XXXX*_{red}) is faster than that of color-neutral word items (e.g., *DEAL*_{red}). This effect,
206 known as *lexicality cost* or *neutral word interference* (e.g., Brown, 2011), is of substantially
207 greater magnitude (i.e., conflates magnitude of Stroop interference even more) when the
208 Stroop task administered with vocal as opposed to manual responses (see Figure 1).

209 Therefore, when the magnitude of interference produced by color-incongruent trials
210 (i.e., the prime indicator of attentional abilities) is under scrutiny, a color-neutral baseline
211 consisting of non-color words (e.g., *DEAL*_{blue}) seems to be the most appropriate baseline. This

212 is because the lexicality cost of non-color related words is identical to that observed on color-
213 incongruent words, and as mentioned above, they also allow the derivation of interference
214 independently of facilitation. In conclusion, non-color *words* (as opposed to other types of
215 color-neutral baseline items) permits the most accurate estimation of interference produced by
216 color-incongruent items.

217 Despite the more ideal nature of the non-color word baseline, it is still seldom used
218 and not just in research examining the link between physical activity/exercise and attentional
219 abilities considered below. In light of Cremen and Carson’s (2017) argument that in
220 intervention studies standardized tests are *de rigueur*, perhaps the most striking example is the
221 fact that the standardized Stroop tests used in clinical practice (e.g., Bayard et al., 2011;
222 Troyer et al., 2006) do not currently use non-color word baseline to derive interference².
223 Consequently, in what follows, the *Stroop interference effect* is defined more loosely as the
224 difference in speed and accuracy between color-incongruent and any kind of color-neutral
225 trial (e.g., $BLUE_{red} - DEAL/XXXX/■■■■_{red}$). In contrast, a difference between color-
226 incongruent and color-congruent trials (e.g., $BLUE_{red} - BLUE_{blue}$) – confounding interference
227 and facilitation – will be referred to as the *overall Stroop effect*. Finally, whenever both are
228 being evoked, the umbrella term of *Stroop effect* will be used.

229

230 **Assessing the link between any given intervention and attentional selectivity measured** 231 **in the Stroop task**

² In these tests, non-color words are in fact included, but color-naming on these trials is compared to that on non-word trials (e.g., colored dots for instance). This latter measure of “low interference” is different from “high interference” condition (i.e., performance difference between color-incongruent words and dots). As explained above, significant “low interference” constitutes indeed an empirical demonstration of *lexicity cost* or *neutral word interference* (e.g., Brown, 2011) discussed above. Most importantly, the “genuine” interference (i.e., performance difference between color-incongruent words and non-color words that is free of lexicity cost) is not computed.

232 As already mentioned, to test the merits of any given intervention – including physical
233 activity/exercise – one can assess the extent to which this intervention is either associated
234 with attentional abilities or causes their improvement. In the former case, the test is
235 correlational and requires that both the levels of Stroop performance and of a given
236 intervention or practice (e.g., level of physical fitness) are measured. In order to claim that
237 higher physical fitness for instance is associated with better attentional selectivity, it is the
238 magnitude of Stroop interference (preferably measured with color-neutral words trials as
239 baseline) that needs to be negatively correlated with physical fitness (see section above). Said
240 differently, in light of the conceptual and methodological issues raised in the previous section,
241 negative correlation between levels of physical fitness and the overall Stroop effect (or simply
242 RTs and/or errors observed for incongruent items etc., see Table 2S in SM for examples of
243 these other types of measures) is simply not enough. Indeed, because the overall Stroop effect
244 confounds both interference and facilitation, it is impossible to conclude whether a given
245 intervention or practice (e.g., level of physical fitness or measured levels of daily physical
246 activity) is associated with lower interference, greater facilitation, or both.

247 Experimental studies (i.e., studies in which a given intervention is considered as a
248 potential cause of changes in attentional abilities and is therefore manipulated) face this same
249 issue. Indeed, if effective, any given intervention that aims to improve attention/inhibition
250 should reduce Stroop interference specifically. Again, the distinction between interference
251 and facilitation is important because some interventions (but also a mere repetition of
252 measures in pre- vs. post-intervention designs, see below) might also, or perhaps even
253 exclusively, boost facilitation and/or contingent responding discussed in the previous section
254 (see e.g., Algom et al., 2022; Algom & Chajut, 2019; Schmidt, 2019, 2023 for discussion of
255 these issues in research on processes driving cognitive control).

256 Additionally, to evidence an attentional boost unambiguously, the aforementioned
257 post-interventional reduction of Stroop *interference* needs to specifically result from
258 shortened RTs to color-incongruent trials (i.e., increase in speed that is not traded for
259 accuracy, see previous section and see Palfi et al., 2022 for an example of this). Still, because
260 many interventions – including physical activity for instance – simultaneously influence
261 response speed (and not just attentional selectivity), a post-intervention reduction of RTs to
262 color-neutral baseline items will often occur in tandem with a reduction in RTs to incongruent
263 items.

264 This is a known and long-lasting issue in Stroop research. Because the magnitude of
265 Stroop interference is known to increase proportionally to the slow-down in response speed
266 (see e.g., Martinon et al., 2024; Roelofs, 2010, see also Faust et al., 1999; Spieler et al., 1996),
267 it has been argued that the age-related differences in this magnitude can be entirely explained
268 by the older adults' general slowing and not alterations in their attentional selectivity per se
269 (e.g., Verhaeghen & De Meersman, 1998; see also Verhaeghen, 2000, 2015). This
270 consequently prompted more recent research addressing the effect of healthy and/or
271 pathological ageing on selective attention to better control for concomitant changes in
272 response speed of older adults (e.g., Bugg et al., 2007; Burca et al., 2022). As suggested
273 below, a similar approach is needed to isolate the effect of any given intervention on selective
274 attention/inhibition from its (simultaneous and positive) effect on processing speed.

275 It is however also important to underline that any experimental intervention that
276 results in performance improvements has to be further isolated from the effects of practice
277 that comes in to play whenever pre- and -post intervention measurements of Stroop interference
278 are not sufficiently separated. Therefore, these practice-related changes in response speed also
279 need to be closely monitored when assessing the benefits of a given intervention.

280 To conclude these opening sections, it seems important recall that the Stroop task is
281 often evoked to illustrate the so-called *task-impurity problem* (MacLeod, 2003; Myiake &
282 Friedman, 2012). This notion refers specifically to the fact that any Stroop performance
283 indicator necessarily includes variance that is attributable to processes other than selective
284 attention and the interference control it entails, like the aforementioned speed of processing
285 (see also e.g., Diamond, 2020; Furley et al., 2023; Smidt et al., 2020 for more general
286 discussions of this issue). Still, the vast majority of methodological and conceptual difficulties
287 discussed above are not related to this task-impurity problem. Rather, these difficulties stem
288 from the design of a given study and/or the type of analyses performed as they consistently
289 fail to isolate interference induced by color-incongruent items (i.e., prime indicator of
290 attentional selectivity and interference control it entails) from other processes observed in the
291 Stroop or other conflict tasks (facilitation, contingent responding etc.). To illustrate this issue,
292 the following section scrutinized these latter aspects in studies assessing the link between
293 physical activity/exercise and attentional abilities.

294

295 **Assessing the link between physical activity and attentional selectivity measured in the** 296 **Stroop task**

297 In line with long-lasting research on processes involved in the Stroop task (e.g.,
298 Chajut & Algom, 2019; MacLeod, 1998; Parris et al., 2022), it was argued above that the
299 most direct measure of *interference* induced by color-incongruent trials is derived by
300 comparing RTs and errors observed with these latter items, against those observed with color-
301 neutral items consisting preferably of non-color words (e.g., *DEAL*_{red}). Therefore, the first
302 question that the present section attempts to answer concerns the extent to which this specific
303 comparison is used in research investigating the link (correlational or causal) between *fitness*,

304 *physical activity* (both *endurance* and *resistance*) or *exercise* (put together under the umbrella
305 term physical activity) and attentional selectivity (and control and/inhibition it entails).

306 To this end, we conducted a bibliographic search on Google Scholar, PsychInfo,
307 Pubmed, Scopus and Web of Science using the following keywords: Stroop task OR Stroop
308 effect OR Stroop interference AND Physical fitness OR physical activity OR physical
309 exercise. This search focused on healthy participants only. The identification, screening, and
310 inclusion/exclusion steps are presented in Figure 1S (see SM). Initially, 55 records of studies
311 – published between 2006-2021 – were identified as carrying out this type of investigation
312 employing the Stroop task (either exclusively or in addition to other cognitive tasks). One
313 record was discarded because, the reported study only used the Stroop task as a stressor. The
314 remaining 54 records – all journal articles – were retrieved and screened. As a results of the
315 first screening for eligibility, three articles were considered as ineligible and therefore
316 excluded (see Figure 1S in SM for further exclusion details).

317 Indicators of Stroop performance used in 51 remaining journal articles were then
318 systematically scrutinized. As a result of this scrutiny, only 9 studies (17,3%) were found to
319 have compared RTs and errors on color-incongruent to some kind of color-neutral baseline
320 (i.e., measured magnitudes of Stroop interference). However, none of them employed non-
321 color words (e.g., $DEAL_{red}$) as baseline. 9 other studies (17,3%) measured the magnitude of
322 the overall Stroop effect (i.e., $BLUE_{red} - BLUE_{blue}$, see Table 1S in SM) and the 34³ (65,4%)
323 remaining studies used other indicators of Stroop performance (e.g., RT and/or errors to
324 color-incongruent trials only – see Table 2S in Supplementary Materials). Although results
325 and conclusions of these latter studies were closely scrutinized (see Tables 1S and 2S in
326 Supplementary Materials for summary), in line with methodological and conceptual

³ One journal article reported two separate experiments; these are presented as two separate entries in Table 2S

327 considerations outlined in the two previous sections, they were no longer considered in what
328 follows (see Figure 1S in SM).

329 Our next goal was to assess the extent to which physical activity is indeed associated
330 with an increased efficiency and/or a boost in attentional selectivity. Given the paucity of
331 studies measuring Stroop interference (and its reduction) directly (see 9 studies depicted in
332 Table 1), their methodological differences and their remaining shortcomings (see last column
333 of Tables), this issue was assessed via a critical narrative review rather than a systematic
334 review or a metaanalysis.

335 Three out of nine included studies were correlational. Buck et al. (2008) found a
336 positive association between aerobic fitness and percentages of correct responses in each list
337 of the card- / list-version of the Stroop task, however, no association was found between
338 aerobic fitness and the magnitude of Stroop interference (e.g., $BLUE_{red} - XXXX_{red}$) in RTs.
339 Despite the absence of this latter (i.e., crucial) association, the study concluded that cognition
340 and notably executive control (interference control) in preadolescent children benefits from
341 fitness. Ishihara et al. (2018) found a negative association between physical fitness and
342 magnitudes of Stroop interference (e.g., $BLUE_{red} - XXXX_{red}$) in 9-year-old tennis players
343 showing that the more physically fit children experienced smaller magnitudes of Stroop
344 interference. Although the study concluded that there was a benefit of fitness on inhibitory
345 control, no such association was found when Stroop interference was correlated with daily
346 physical activity or tennis experience. Finally, van der Niet et al. (2015) found a small
347 negative association between time spent in sedentary behavior and magnitudes of Stroop
348 interference. In line with its correlational nature, the study concluded that the more time
349 children spent in sedentary behavior, the worse is their performance.

350 Six remaining studies were experimental. Two of them suggest a positive effect of
351 acute aerobic and/or resistance exercise on attentional selectivity of middle-aged participants

352 (Alves et al., 2012, 2014). Indeed, in Alves et al. (2014) for instance, the reduced magnitudes
353 of Stroop interference (due to reduced RTs to color-incongruent items only) were observed
354 after a cycling but not stretching (i.e., control) exercise that were administered seven days
355 apart (see Alves et al., 2012 for benefits of 30-min of either walking or strength exercise
356 compared to stretching in middle-aged women).

357 The four remaining experimental studies involved a long-term training intervention
358 (i.e., at least 22-weeks up to 12 months). Indeed, after 22-weeks of a 30-min long physical
359 training program during lunch time (twice a week), primary school children displayed a
360 reduced magnitude of Stroop interference, compared to children that kept the same lunch
361 routine (van der Niet et al., 2016). The authors consequently argued that physical activity
362 benefits executive functioning (notably inhibition) in primary school children. In older
363 women, Best et al. (2014) found a similar benefit after 12 months of 60-min classes (whether
364 these were completed once or twice a week) but also after a balance, toning and relaxation
365 program (i.e., active control program) of the same length. Also, and importantly, a post-
366 intervention reduction in magnitudes of Stroop interference positively predicted adherence to
367 physical activity during the follow-up period (13 to 25 months). While using very similar
368 program in both content and length and administered to a similar population of older women
369 (although possibly significantly older than those recruited in Best et al., 2014), Liu-Ambrose
370 et al. (2010) found no post-intervention reduction of Stroop interference in the active control
371 toning and relaxation program. Sipilä et al. (2021) showed that when coupled with a cognitive
372 training program (i.e., 15-20 minutes computer training of executive functions 3-4 times a
373 week), the selective attention ability of older women benefited from a 12-month physical
374 training program even more than a group of women following a physical training-only
375 program— as evidenced by a greater post-interventional reduction of Stroop interference in the
376 former as compared to a latter group.

377 To sum up, 8 out of 9 studies reviewed above suggest either an association between
378 the level of exercise/fitness and attentional selectivity (although only in young children) or a
379 positive change in individuals' attentional selectivity following an acute bout of activity, or
380 long-term physical activity (although only in young children or middle-aged or older adults,
381 mostly women, see also Table 1, column entitled "Main findings on Stroop performance
382 only" for further details). Nevertheless, these results need to be interpreted with caution (see
383 also Table 1, column entitled "Main conclusions drawn in the study" for proposed
384 interpretations, and column "Observations and Issues" for elements potentially precluding the
385 proposed interpretation).

386 As discussed above, none of these studies measured the magnitude of Stroop
387 interference using a non-color word baseline (e.g., *DEAL_{red}*). Since almost all of them
388 employed a vocal response modality (see column "Item presentation; Response modality;
389 Number of colors"), the derived magnitude of interference is likely to confound interference
390 on color-incongruent words with faster identification of baseline non-word items (due the
391 absence of *lexicality cost*, see Figure 1 for a magnitude of this potential confound). This
392 confound might be shown to be problematic when the reduction of interference is scrutinized
393 (i.e., in the case of experimental studies depicted above).

394 Additionally, not all of these studies effectively controlled for changes in processing
395 speed. This is even more important in studies outlined above, since the effect of physical
396 activity and/or of fitness has been extensively studied in populations presenting high intra-
397 and inter-individual variability in RTs (i.e., children, older adults) that might also evolve
398 rapidly (e.g., between the beginning and the end of 12-months intervention) as function of
399 individual developmental trajectory (Ferrand et al., 2020; Faust et al., 1999).

400 In light of current research standards, several studies reviewed above might also be
401 underpowered. This is not exclusively due to the high number of observations that

402 correlational studies require or to the high-power requirements for between-groups
403 comparisons that this type of experimental research often entails. Indeed, Brysbaert and
404 Stevens (2018) for instance recommend a minimum of 1,600 observations *per condition* in
405 designs with repeated measures (i.e., all designs including measures of the Stroop effect and
406 of its reduction). Although the number of observations can be divided over participants and
407 items (e.g., 20 participants and 80 observations, 40 participants and 40 observations, or 80
408 participants and 20 observations), some of the studies cited above are far from meeting these
409 standards (see Table 1, column entitled “*Stroop task and Item-type (with number of trials if*
410 *available*” and see the same column in Tables 1S & 2S in Supplementary Materials). In line
411 with this reasoning, Ciria et al. (2023) recently concluded that randomized control trials on
412 the effects of physical exercise on cognition indeed suffer from low statistical power.

413 Despite the fact the Stroop task is often considered as “a prototypical inhibition task”
414 (Miyake et al., 2000, p. 57, see above), perhaps the most important challenge that the
415 abovementioned studies face is the interpretation of an increased efficiency and/or a boost in
416 selective attention in terms of underlying processes, namely inhibition and/or the control that
417 attentional selectivity entails (see also Table 1, column entitled “Main conclusions drawn in
418 the study”). Indeed, research from a wide range of conflict tasks suggests that while distractor
419 inhibition certainly contributes to reducing interference from the irrelevant word-dimension
420 (i.e., so-called *inhibition hypothesis*, see Egner & Hirsch, 2005; Wyatt & Machado, 2013),
421 different interventions are also likely to enhance processing of the target (i.e., the ink-color to
422 be identified). In line with this *target amplification hypothesis*, it remains plausible that
423 physical activity actually enhances focus on the relevant color-dimension of the Stroop task
424 (see also e.g., De Jong et al., 1999 for an *increased maintenance of the relevant task-set*).
425 Accordingly, Sibley et al. (2006) reasoned that exercise is likely to improve the maintenance
426 of goal-oriented processing (i.e., it improves a more consistent focus on the color-dimension

427 instead of on the inhibition of the irrelevant word-dimension, see also Table 2S for details of
428 this study).

429 *Insert Table 1 near here*

430

431 **Practical guidelines for a more direct measure of selective attention in the Stroop task**

432 In line with conclusions that have been reached by inspecting results from other
433 inhibitory (e.g., Cremen & Carson, 2017) or executive function tasks (e.g., Furley et al.,
434 2023), the previous section (along with summaries provided in Tables 1, 1S & 2S) pointed to
435 the fact that there is a substantial gap between what a given body of research shows and the
436 conclusions it is able to draw (see namely Tables 1S & 2S in Supplementary Materials for
437 examples of this gap).

438 However, this gap is not limited to research on the effects of physical activity nor to
439 translational (or perhaps more applied) research in general. To illustrate, basic research on
440 cognitive control (see e.g., Egner, 2023 for a review) for instance, is also prone to some of the
441 shortcomings outlined above. Indeed, in the Stroop task, a boost in cognitive control is
442 inferred from the reduction of the overall Stroop effect (and not from the reduction of Stroop
443 interference, exactly like studies presented in Table 1S) and therefore disregards the role
444 facilitation plays in this reduction (see e.g., Algom & Chajut, 2019, Algom et al., 2022 for a
445 discussion of this issue). Additionally, only recently controls for contingency issues –
446 that the use of color-congruent items induces – have been employed (e.g., Braem et al., 2019,
447 but see also Schmidt, 2019, 2023). Therefore, this section attempts to provide some up-to-date
448 practical guidelines designed to eliminate issues that detract from the putative conclusion that
449 an increased efficiency and/or a boost in selective attention in the Stroop task has indeed been
450 demonstrated (via any type of laboratory and/or field intervention).

451 In line with methodological and theoretical arguments outlined above, future
452 correlational and experimental studies need to measure interference on color-incongruent
453 items directly (i.e., against a color-neutral baseline that preferably consists of non-color
454 words, **Issue 1**). If color-congruent items are used as additional (i.e., not a unique) baseline,
455 experimental designs need to control for contingency (**Issue 2**) such that, designs using equal
456 proportions of color-congruent and color-incongruent items should be avoided. Additionally,
457 for experimental studies to evidence the post-interventional boost in attentional selectivity
458 unambiguously, post-interventional reduction of Stroop interference needs to result from
459 shortened RTs to color-incongruent trials (**Issue 3**) – an increase in speed that is clearly not
460 traded for accuracy (as evidenced by simultaneous consideration of both RTs and errors,
461 **Issue 4**). Whenever RTs to a relevant baseline decrease along with those to color-incongruent
462 RTs, it becomes increasingly difficult to assess whether these changes also reflect a boost in
463 selective attention (or in inhibition and/or control it entails) and to what extent (see Palfi et al.,
464 2022 for an example of this). Thus, in order to isolate the effect of physical activity on
465 selective attention per se, future studies need to control for the effect of processing speed
466 (**Issue 5**). This is rather effectively done via proportional transformation (e.g., $[(RT_{\text{incongruent}} -$
467 $RT_{\text{baseline}}) / RT_{\text{baseline}}]$). However, it is important to understand that proportional transformation
468 might actually create an advantage for individuals presenting slower RTs (e.g., older adults,
469 children, Hedge et al., 2018). This spurious advantage (compared to populations with faster
470 RTs, e.g., young adult) is no longer present when the variability is instead taken into account
471 using z-scores (Faust et al., 1999; Hedge et al., 2018; see also Burca et al. 2022 for a direct
472 demonstration). To eliminate practice-related changes in response speed (**Issue 6**), pre- and -
473 post-interventional measurements of Stroop interference also need to be sufficiently delayed.

474 As emphasized in the previous section, a reduced magnitude of Stroop interference
475 does not necessarily result from a reduced interference of the irrelevant word-dimension of

476 incongruent trials (**Issue 7**). It could alternatively and/or in concert result from an enhanced
477 focus on their relevant color-dimension. To address this issue, several studies on the effects of
478 physical activity used the amount of negative priming (i.e., RTs to trials in which the ink
479 color to be named on N-trial corresponds the (word) stimulus that was ignored on N-1 trial) as
480 an additional indicator of an inhibitory boost (see Table 2S for examples). However, since
481 negative priming is a notoriously controversial measure of selective inhibition (see e.g.,
482 Besner, 2001), a combination of several “inhibitory” tasks⁴, or the use of moderators known
483 to promote target amplification might constitute venues for addressing this latter issue directly
484 (e.g., De Jong et al., 1999, argued that shortening an interval between a response and the
485 presentation of a new stimulus increases a focus on the relevant color-dimension of Stroop
486 stimulus). Nevertheless, post-interventional reductions of Stroop interference can still be
487 interpreted simply in terms of a “boost in attentional selectivity” or a “boost in attentional
488 selectivity and the control it entails” as it is more agnostic with respect to processes actually
489 underlying the reduction (inhibition vs. target amplification / increased goal maintenance
490 etc.).

491 Another important, as yet unaddressed, challenge is to establish whether a
492 demonstrated boost in attentional selectivity is general or limited to a specific subsystem.
493 Indeed, in the Stroop task, distractors can cause conflict at multiple stages within the
494 information-processing system – conflict that may be inhibited by different subsystems
495 (Parris et al., 2022; see also e.g., Berggren & Derakshan, 2014 who showed anxiety affected
496 only one subsystem of attentional control; and Augustinova et al., 2018; Burca et al., 2022
497 and Li & Bosman, 1996 for the same effect of healthy ageing).

⁴ Still, many (response) inhibition tasks are prone to the same limitations as those discussed above including the fact that they do not include an appropriate neutral baseline (see e.g., Smith & Ulrich, 2023 for a discussion, see also e.g., Cremen & Carson, 2017 for a discussion of other issues).

498 To address this point, future studies might also consider using Stroop tasks that permit
499 the measurement of the independent contributions of *response*, *semantic* and *task* conflicts to
500 the Stroop interference effect (see e.g., Hasshim & Parris, 2014 and see Parris et al., 2022 for
501 a review; although see Parris et al., 2023 for a critical discussion of task conflict). Indeed, the
502 ultimate goal of this tutorial section is to help future studies using the Stroop task to make
503 more theoretically and empirically grounded inferences not only about the effects of any
504 given intervention on attentional selectivity (and/or different processes it entails) but also
505 about the levels of processing where these effects take place.

506

507 **Concluding remarks**

508 In times where sedentary lifestyles are increasingly observed in both children and
509 adults (World Health Organization, 2022; Magnon et al., 2018), questioning the positive
510 effect of physical activity and/or of fitness in the Stroop task might seem a bit
511 counterproductive (although see e.g., Ciria et al., 2023). It needs to be remembered at this
512 point that several meta-analyses still agree on a positive (although modest) exercise-related
513 benefit on cognition (e.g., Colcombe & Kramer, 2003; Hillman et al., 2008; Barenberg et al.,
514 2011; Guiney, & Machado, 2013; Gomez-Pinilla & Hillman, 2013; Hötting and Röder, 2013;
515 Gajewski & Falkenstein, 2016; Kramer & Colcombe, 2018; Ludyga et al., 2020). Importantly,
516 these meta-analyses also included studies of other cognitive functions and/or studies of attention
517 using other tasks (not just the standard color-word Stroop task but also or alternatively a
518 flanker, stop-signal, cuing tasks etc.) – results of which are not considered above (or in Tables
519 1S and 2S). Although Furley et al. (2023, p.14) convincingly argued “that methodological
520 flaws and measurement problems have been a great issue in the primary studies that have fed
521 into (...)” these meta-analyses and the present paper hopefully illustrated in greater details the
522 type of flaws and measurement problems that occur in the Stroop task specifically, it is still

523 essential to insist on the importance of physical activity and fitness. In addition to health
524 benefits (see e.g. Warburton & Bredin, 2017; Melzer et al., 2004), these activities are also
525 known to enhance psychological and social health outcomes in younger populations (i.e.,
526 children and adolescents, Eime et al., 2013), and also in older adults (Knight et al., 2022).
527 Therefore, by no means does the present section of this paper seek to discourage regular
528 physical activity and/or of high fitness level, or discourage research into their benefits. Rather,
529 it seeks to draw attention to the uncertainty of the knowledge we currently have on the link
530 between physical activity and/or exercise on attentional selectivity (see also Furley et al.,
531 2023 for a similar idea). Indeed, many more carefully designed and powered studies –
532 conducted in populations covering the entire life-span – are still needed to address the issue of
533 whether there is a boost in attentional selectivity following physical activity (or any other
534 intervention) in robust systematic reviews and metaanalyses. We only hope that the practical
535 guidelines outlined above will prove useful in this endeavor.

536

537 **Declarations**

538 *Funding details*

539 The work reported was supported by RIN Tremplin Grant 19E00851 from Normandie
540 Region, France and by ANR-19-CE28-0013 Grant from Agence Nationale de la recherche
541 awarded to all senior authors.

542

543 *Disclosure statement*

544 The authors report there are no competing interests to declare.

545

546 **Data availability statement**

547 Data sharing is not applicable to this article as no new data were created or analyzed in this
548 study

549

550 ***Supplemental online material***

551 This paper includes supplementary materials submitted with the main article.

552

553 ***Authors' contributions***

554 NR and MA jointly drafted the manuscript. PI, LF, and BP provided substantial intellectual
555 content. All authors read and approved the final version of the manuscript.

556

557 **References**

558 Algom, D., & Chajut, E. (2019). Reclaiming the Stroop Effect Back From Control to Input-
559 Driven Attention and Perception. *Frontiers in Psychology, 10*, 1683.

560 <https://doi.org/10.3389/fpsyg.2019.01683>

561 Algom, D., Fitousi, D., & Chajut, E. (2022). Can the Stroop effect serve as the gold standard
562 of conflict monitoring and control? A conceptual critique. *Memory & Cognition,*
563 *50(5)*, 883–897. <https://doi.org/10.3758/s13421-021-01251-5>

564 Alves, C. R. R., Gualano, B., Takao, P. P., Avakian, P., Fernandes, R. M., Morine, D., &
565 Takito, M. Y. (2012). Effects of Acute Physical Exercise on Executive Functions: A
566 Comparison Between Aerobic and Strength Exercise. *Journal of Sport and Exercise*
567 *Psychology, 34(4)*, 539–549. <https://doi.org/10.1123/jsep.34.4.539>

568 Alves, C. R. R., Tessaro, V. H., Teixeira, L. A. C., Murakava, K., Roschel, H., Gualano, B., &
569 Takito, M. Y. (2014). Influence of Acute High-Intensity Aerobic Interval Exercise
570 Bout on Selective Attention and Short-Term Memory Tasks. *Perceptual and Motor*
571 *Skills, 118(1)*, 63–72. <https://doi.org/10.2466/22.06.PMS.118k10w4>

- 572 Augustinova, M., Banovic, I., Burca, M., Ferrand, L., Iodice, P., Junker, C., Kernivinen, V.,
573 & Parris, B. A. (2022). Does exposure to pictures of nature boost attentional control in
574 the Stroop task? *Journal of Environmental Psychology*, *84*, 101901.
575 <https://doi.org/10.1016/j.jenvp.2022.101901>
- 576 Augustinova, M., Clarys, D., Spatola, N., & Ferrand, L. (2018). Some further clarifications on
577 age-related differences in Stroop interference. *Psychonomic Bulletin & Review*, *25*,
578 767-774. <https://doi.org/10.3758/s13423-017-1427-0>
- 579 Augustinova, M., & Ferrand, L. (2012). Suggestion does not de-automatize word reading:
580 Evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*,
581 *19*(3), 521–527. <https://doi.org/10.3758/s13423-012-0217-y>
- 582 Augustinova, M., Parris, B. A., & Ferrand, L. (2019). The Loci of Stroop Interference and
583 Facilitation Effects With Manual and Vocal Responses. *Frontiers in Psychology*, *10*,
584 1786. <https://doi.org/10.3389/fpsyg.2019.01786>
- 585 Augustinova, M., Silvert, L., Spatola, N., & Ferrand, L. (2018). Further investigation of
586 distinct components of Stroop interference and of their reduction by short response-
587 stimulus intervals. *Acta Psychologica*, *189*, 54–62.
588 <https://doi.org/10.1016/j.actpsy.2017.03.009>
- 589 Barenberg, J., Berse, T., & Dutke, S. (2011). Executive functions in learning processes: Do
590 they benefit from physical activity? *Educational Research Review*, *6*(3), 208–222.
591 <https://doi.org/10.1016/j.edurev.2011.04.002>
- 592 Bayard, S., Erkes, J., & Moroni, C. (2011). Victoria Stroop Test: normative data in a sample
593 group of older people and the study of their clinical applications in the assessment of
594 inhibition in Alzheimer's disease. *Archives of Clinical Neuropsychology*, *26*(7), 653-
595 661. <https://doi.org/10.1093/arclin/acr053>

- 596 Berggren, N., & Derakshan, N. (2014). Inhibitory deficits in trait anxiety: Increased stimulus-
597 based or response-based interference? *Psychonomic Bulletin & Review*, *21*(5), 1339–
598 1345. <https://doi.org/10.3758/s13423-014-0611-8>
- 599 Besner, D. (2001). The myth of ballistic processing: Evidence from Stroop's paradigm.
600 *Psychonomic Bulletin & Review*, *8*(2), 324–330. <https://doi.org/10.3758/BF03196168>
- 601 Best, J. R., Nagamatsu, L. S., & Liu-Ambrose, T. (2014). Improvements to executive function
602 during exercise training predict maintenance of physical activity over the following
603 year. *Frontiers in Human Neuroscience*, *8*, 353.
604 <https://doi.org/10.3389/fnhum.2014.00353>
- 605 Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J., Weissman, D. H., Notebaert, W., &
606 Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in cognitive*
607 *sciences*, *23*(9), 769-783. <https://doi.org/10.1016/j.tics.2019.07.002>
- 608 Brown, T. L. (2011). The relationship between stroop interference and facilitation effects:
609 Statistical artifacts, baselines, and a reassessment. *Journal of Experimental*
610 *Psychology: Human Perception and Performance*, *37*(1), 85–99.
611 <https://doi.org/10.1037/a0019252>
- 612 Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects
613 Models: A Critical. *Journal of Cognition*, *1*(1), 9. <https://doi.org/10.5334/joc.10>
- 614 Buck, S. M., Hillman, C. H., & Castelli, D. M. (2008). The Relation of Aerobic Fitness to
615 Stroop Task Performance in Preadolescent Children. *Medicine & Science in Sports &*
616 *Exercise*, *40*(1), 166–172. <https://doi.org/10.1249/mss.0b013e318159b035>
- 617 Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age Differences in Stroop
618 Interference: Contributions of General Slowing and Task-Specific Deficits. *Aging,*
619 *Neuropsychology, and Cognition*, *14*(2), 155–167.
620 <https://doi.org/10.1080/138255891007065>

- 621 Burca, M., Chausse, P., Ferrand, L., Parris, B. A., & Augustinova, M. (2022). Some further
622 clarifications on age-related differences in the Stroop task: New evidence from the
623 two-to-one Stroop paradigm. *Psychonomic Bulletin & Review*, 29(2), 492–500.
624 <https://doi.org/10.3758/s13423-021-02011-x>
- 625 Ciria, L. F., Román-Caballero, R., Vadillo, M. A., Holgado, D., Luque-Casado, A., Perakakis,
626 P., & Sanabria, D. (2023). An umbrella review of randomized control trials on the
627 effects of physical exercise on cognition. *Nature Human Behaviour*, 1–14.
628 <https://doi.org/10.1038/s41562-023-01554-4>
- 629 Colcombe, S., & Kramer, A. F. (2003). Fitness Effects on the Cognitive Function of Older
630 Adults: A Meta-Analytic Study. *Psychological Science*, 14(2), 125–130.
631 <https://doi.org/10.1111/1467-9280.t01-1-01430>
- 632 Cremen, I. A., & Carson, R. G. (2017). Have standard tests of cognitive function been
633 misappropriated in the study of cognitive enhancement?. *Frontiers in human*
634 *neuroscience*, 11, 276. <https://doi.org/10.3389/fnhum.2017.00276>
- 635 Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-
636 word task. *Perception & Psychophysics*, 11(4), 274–276.
637 <https://doi.org/10.3758/BF03210377>
- 638 Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of Some Aspects of the Stroop
639 Color-Word Test. *Perceptual and Motor Skills*, 23(3_suppl), 1211–1214.
640 <https://doi.org/10.2466/pms.1966.23.3f.1211>
- 641 De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations:
642 Dissociable causes of interference effects in conflict situations. *Acta psychologica*,
643 101(2-3), 379-394. [https://doi.org/10.1016/S0001-6918\(99\)00012-8](https://doi.org/10.1016/S0001-6918(99)00012-8)
- 644 Diamond, A. (2020). Executive functions. In J. L. Michaud, C. Bulteau, D. Cohen, & A.
645 Gallagher (Eds.), *Handbook of clinical neurology* (Vol. 173, pp. 225–240). Elsevier.

- 646 Egner, T. (2023). Principles of cognitive control over task focus and task switching. *Nature*
647 *Reviews Psychology*, 1-13. <https://doi.org/10.1038/s44159-023-00234-4>
- 648 Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical
649 amplification of task-relevant information. *Nature Neuroscience*, 8(12), 1784–1790.
650 <https://doi.org/10.1038/nn1594>
- 651 Eime, R. M., Young, J. A., Harvey, J. T., Charity, M. J., & Payne, W. R. (2013). A systematic
652 review of the psychological and social benefits of participation in sport for children
653 and adolescents: Informing development of a conceptual model of health through
654 sport. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 98.
655 <https://doi.org/10.1186/1479-5868-10-98>
- 656 Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in
657 information-processing rate and amount: Implications for group differences in
658 response latency. *Psychological Bulletin*, 125(6), 777–799.
659 <https://doi.org/10.1037/0033-2909.125.6.777>
- 660 Ferrand, L., Ducrot, S., Chausse, P., Maïonchi-Pino, N., O'Connor, R. J., Parris, B. A., ... &
661 Augustinova, M. (2020). Stroop interference is a composite phenomenon: Evidence
662 from distinct developmental trajectories of its components. *Developmental Science*,
663 23(2), e12899. <https://doi.org/10.1111/desc.12899>
- 664 Furley, P., & Goldschmied, N. (2021). Systematic vs. narrative reviews in sport and exercise
665 psychology: Is either approach superior to the other?. *Frontiers in psychology*, 12,
666 685082. <https://doi.org/10.3389/fpsyg.2021.685082>
- 667 Furley, P., Schütz, L. M., & Wood, G. (2023). A critical review of research on executive
668 functions in sport and exercise. *International Review of Sport and Exercise*
669 *Psychology*, 1-29. <https://doi.org/10.1080/1750984X.2023.2217437>

- 670 Gajewski, P. D., & Falkenstein, M. (2016). Physical activity and neurocognitive functioning
671 in aging—A condensed updated review. *European Review of Aging and Physical*
672 *Activity*, 13(1), 1–7. <https://doi.org/10.1186/s11556-016-0161-3>
- 673 Gomez-Pinilla, F., & Hillman, C. (2013). The Influence of Exercise on Cognitive Abilities.
674 *Comprehensive Physiology*, 403–428. <https://doi.org/10.1002/cphy.c110063>
- 675 Guiney, H., & Machado, L. (2013). Benefits of regular aerobic exercise for executive
676 functioning in healthy populations. *Psychonomic Bulletin & Review*, 20(1), 73–86.
677 <https://doi.org/10.3758/s13423-012-0345-4>
- 678 Hasshim, N., & Parris, B. A. (2014). Two-to-one color-response mapping and the presence of
679 semantic conflict in the Stroop task. *Frontiers in Psychology*, 5:1157.
680 <https://doi.org/10.3389/fpsyg.2014.01157>
- 681 Hasshim, N., & Parris, B. A. (2021). The role of contingency and correlation in the Stroop
682 task. *Quarterly Journal of Experimental Psychology*, 74(10), 1657–1668.
683 <https://doi.org/10.1177/17470218211032548>
- 684 Hedge, C., Powell, G., & Sumner, P. (2018). The mapping between transformed reaction time
685 costs and models of processing in aging and cognition. *Psychology and Aging*, 33(7),
686 1093–1104. <https://doi.org/10.1037/pag0000298>
- 687 Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and
688 behavior. *Frontiers in Neuroscience*, 8, 150. <https://doi.org/10.3389/fnins.2014.00150>
- 689 Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart:
690 Exercise effects on brain and cognition. *Nature Reviews Neuroscience*, 9(1), 58–65.
691 <https://doi.org/10.1038/nrn2298>
- 692 Hötting, K., & Röder, B. (2013). Beneficial effects of physical exercise on neuroplasticity and
693 cognition. *Neuroscience & Biobehavioral Reviews*, 37(9, Part B), 2243–2257.
694 <https://doi.org/10.1016/j.neubiorev.2013.04.005>

- 695 Ishihara, T., Sugasawa, S., Matsuda, Y., & Mizuno, M. (2018). Relationship between sports
696 experience and executive function in 6-12-year-old children: Independence from
697 physical fitness and moderation by gender. *Developmental Science*, *21*(3), e12555.
698 <https://doi.org/10.1111/desc.12555>
- 699 Jackson, J. D., & Balota, D. A. (2013). Age-related changes in attentional selection: Quality
700 of task set or degradation of task set across time? *Psychology and Aging*, *28*(3), 744–
701 753. <https://doi.org/10.1037/a0033159>
- 702 Knight, R. L., Chalabaev, A., McNarry, M. A., Mackintosh, K. A., & Hudson, J. (2022). Do
703 age stereotype-based interventions affect health-related outcomes in older adults? A
704 systematic review and future directions. *British Journal of Health Psychology*, *27*(2),
705 338–373. <https://doi.org/10.1111/bjhp.12548>
- 706 Kramer, A. F., & Colcombe, S. (2018). Fitness Effects on the Cognitive Function of Older
707 Adults: A Meta-Analytic Study—Revisited. *Perspectives on Psychological Science*,
708 *13*(2), 213–217. <https://doi.org/10.1177/1745691617707316>
- 709 Li, K. Z., & Bosman, E. A. (1996). Age differences in Stroop-like interference as a function
710 of semantic relatedness. *Aging, Neuropsychology, and Cognition*, *3*(4), 272-284.
711 <https://doi.org/10.1080/13825589608256630>
- 712 Liu-Ambrose, T. (2010). Resistance Training and Executive Functions: A 12-Month
713 Randomized Controlled Trial. *Archives of Internal Medicine*, *170*(2), 170–178.
714 <https://doi.org/10.1001/archinternmed.2009.494>
- 715 Ludyga, S., Gerber, M., Pühse, U., Looser, V. N., & Kamijo, K. (2020). Systematic review
716 and meta-analysis investigating moderators of long-term effects of exercise on
717 cognition in healthy individuals. *Nature Human Behaviour*, *4*(6), 603–612.
718 <https://doi.org/10.1038/s41562-020-0851-8>

- 719 MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative
720 review. *Psychological Bulletin*, *109*(2), 163–203. [https://doi.org/10.1037/0033-](https://doi.org/10.1037/0033-2909.109.2.163)
721 [2909.109.2.163](https://doi.org/10.1037/0033-2909.109.2.163)
- 722 MacLeod, C. M. (1992). The Stroop task: The “gold standard” of attentional measures.
723 *Journal of Experimental Psychology: General*, *121*(1), 12–14.
724 <https://doi.org/10.1037/0096-3445.121.1.12>
- 725 MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition
726 to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 43, pp. 163–214). Elsevier Science.
727
- 728 MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop
729 effect: Uncovering the cognitive and neural anatomy of attention. *Trends in cognitive*
730 *sciences*, *4*(10), 383-391. [https://doi.org/10.1016/S1364-6613\(00\)01530-8](https://doi.org/10.1016/S1364-6613(00)01530-8)
- 731 Magnon, V., Vallet, G. T., & Auxiette, C. (2018). Sedentary Behavior at Work and Cognitive
732 Functioning: A Systematic Review. *Frontiers in Public Health*, *6*, 239.
733 <https://doi.org/10.3389/fpubh.2018.00239>
- 734 Martinon L.M., Ferrand, L. Burca M, Hasshim N, Lakzoum D, Parris B.A., Silvert L.,
735 Augustinova M. (2024). Distributional Analyses reveal the Polymorphic Nature of the
736 Stroop Interference Effect: It’s about (Response) Time. *Memory & Cognition*
737 (accepted)
- 738 Melzer, K., Kayser, B., & Pichard, C. (2004). Physical activity: the health benefits outweigh
739 the risks. *Current Opinion in Clinical Nutrition & Metabolic Care*, *7*(6), 641-647.
- 740 Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences
741 in executive functions: Four general conclusions. *Current Directions in Psychological*
742 *Science*, *21*(1), 8–14. <https://doi.org/10.1177/0963721411429458>

- 743 Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D.
744 (2000). The Unity and Diversity of Executive Functions and Their Contributions to
745 Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*,
746 *41*(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- 747 Palfi, B., Parris, B. A., Collins, A. F., & Dienes, Z. (2022). Strategies that reduce Stroop
748 interference. *Royal Society Open Science*, *9*(3), 202136.
749 <https://doi.org/10.1098/rsos.202136>
- 750 Parris, B. A., Hasshim, N., & Dienes, Z. (2021). Look into my eyes: Pupillometry reveals that
751 a post-hypnotic suggestion for word blindness reduces Stroop interference by
752 marshalling greater effortful control. *European Journal of Neuroscience*, *53*(8), 2819–
753 2834. <https://doi.org/10.1111/ejn.15105>
- 754 Parris, B.A., Hasshim, N., Ferrand, L. & Augustinova, M. (2023). Do Task Sets Compete in
755 the Stroop Task and Other Selective Attention Paradigms? *Journal of Cognition*, *6*(1),
756 p.23. <https://doi.org/10.5334/joc.272>
- 757 Parris, B. A., Hasshim, N., Wadsley, M., Augustinova, M., & Ferrand, L. (2022). The loci of
758 Stroop effects: A critical review of methods and evidence for levels of processing
759 contributing to color-word Stroop effects and the implications for the loci of
760 attentional selection. *Psychological Research*, *86*(4), 1029–1053.
761 <https://doi.org/10.1007/s00426-021-01554-x>
- 762 Parris, B. A., Hasshim, N., Augustinova, M., & Ferrand, L. (2022). Do task sets compete in
763 the Stroop task and other selective attention paradigms. *Journal of Cognition*, *6*(1), 23.
764 <https://doi.org/10.5334/joc.272>
- 765 Parris, B. A., Sharma, D., Weekes, B. S. H., Momenian, M., Augustinova, M., & Ferrand, L.
766 (2019). Response Modality and the Stroop Task: Are There Phonological Stroop

- 767 Effects With Manual Responses? *Experimental Psychology*, 66(5), 361–367.
768 <https://doi.org/10.1027/1618-3169/a000459>
- 769 Quétyard, B., Spatola, N., Parris, B. A., Ferrand, L., & Augustinova, M. (2023). A mouse-
770 tracking study of the composite nature of the Stroop effect at the level of response
771 execution. *PLOS ONE*, 18(1), e0279036.
772 <https://doi.org/10.1371/journal.pone.0279036>
- 773 Roelofs, A. (2010). Attention and facilitation: converging information versus inadvertent
774 reading in Stroop task performance. *Journal of Experimental Psychology: Learning,*
775 *Memory, and Cognition*, 36(2), 411-422. <https://doi.org/10.1037/a0018523>
- 776 Schmidt, J. R. (2013). Questioning conflict adaptation: Proportion congruent and Gratton
777 effects reconsidered. *Psychonomic Bulletin & Review*, 20(4), 615–630.
778 <https://doi.org/10.3758/s13423-012-0373-0>
- 779 Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated
780 review. *Psychonomic Bulletin & Review*, 26, 753-771.
781 <https://link.springer.com/article/10.3758/s13423-018-1520-z>
- 782 Schmidt, J. R. (2023). Is conflict adaptation adaptive? An introduction to conflict monitoring
783 theory and the ecological problems it faces. *Quarterly Journal of Experimental*
784 *Psychology*. <https://doi.org/10.1177/17470218231161555>
- 785 Sipilä, S., Tirkkonen, A., Savikangas, T., Hänninen, T., Laukkanen, P., Alen, M., Fielding, R.
786 A., Kivipelto, M., Kulmala, J., Rantanen, T., Sihvonen, S. E., Sillanpää, E., Stigsdotter
787 Neely, A., & Törmäkangas, T. (2021). Effects of physical and cognitive training on
788 gait speed and cognition in older adults: A randomized controlled trial. *Scandinavian*
789 *Journal of Medicine & Science in Sports*, 31(7), 1518–1533.
790 <https://doi.org/10.1111/sms.13960>

- 791 Smid, C. R., Karbach, J., & Steinbeis, N. (2020). Toward a science of effective cognitive
792 training. *Current Directions in Psychological Science*, 29(6), 531–537.
793 <https://doi.org/10.1177/0963721420951599>
- 794 Smith, P., & Ulrich, R. (2023). The neutral condition in conflict tasks: On the violation of the
795 midpoint assumption in reaction time trends. *Quarterly Journal of Experimental*
796 *Psychology*. <https://doi.org/10.1177/17470218231201476>
- 797 Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop Performance in Healthy Younger
798 and Older Adults and in Individuals with Dementia of the Alzheimer’s Type. *Journal*
799 *of Experimental Psychology: Human Perception and Performance*, 22(2), 461–479.
800 Scopus. <https://doi.org/10.1037/0096-1523.22.2.461>
- 801 Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental*
802 *Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- 803 Troyer, A. K., Leach, L., & Strauss, E. (2006). Aging and response inhibition: Normative data
804 for the Victoria Stroop Test. *Aging, Neuropsychology, and Cognition*, 13(1), 20-35.
805 <http://dx.doi.org/10.1080/138255890968187>
- 806 van der Niet, A. G., Smith, J., Oosterlaan, J., Scherder, E. J. A., Hartman, E., & Visscher, C.
807 (2016). Effects of a Cognitively Demanding Aerobic Intervention During Recess on
808 Children’s Physical Fitness and Executive Functioning. *Pediatric Exercise Science*,
809 28(1), 64–70. <https://doi.org/10.1123/pes.2015-0084>
- 810 van der Niet, A. G., Smith, J., Scherder, E. J. A., Oosterlaan, J., Hartman, E., & Visscher, C.
811 (2015). Associations between daily physical activity and executive functioning in
812 primary school-aged children. *Journal of Science and Medicine in Sport*, 18(6), 673–
813 677. <https://doi.org/10.1016/j.jsams.2014.09.006>

- 814 Vandierendonck, A. (2021). On the Utility of Integrated Speed-Accuracy Measures when
815 Speed-Accuracy Trade-off is Present. *Journal of Cognition*, 4(1), 22.
816 <https://doi.org/10.5334/joc.154>
- 817 Verhaeghen, P. (2000). The parallels in beauty's brow: Time-accuracy functions and their
818 implications for cognitive aging theories. In T. J. Perfect & E. A. Maylor (Eds.),
819 *Models of cognitive aging* (pp. 50-86). Oxford: Oxford University Press.
- 820 Verhaeghen, P. (2015). Aging and executive control: Reports of a demise greatly exaggerated.
821 *Current Directions in Psychological Science*, 20, 174-180.
822 <https://doi.org/10.1177/0963721411408772>
- 823 Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: a meta-analysis.
824 *Psychology and aging*, 13(1), 120-126. <https://doi.org/10.1037/0882-7974.13.1.120>
- 825 Warburton, D. E., & Bredin, S. S. (2017). Health benefits of physical activity: a systematic
826 review of current systematic reviews. *Current opinion in cardiology*, 32(5), 541-556.
827 <https://doi.org/10.1097/HCO.0000000000000437>
- 828 World Health Organization (2022). Global status report on physical activity 2022. Geneva
- 829 Wyatt, N., & Machado, L. (2013). Distractor inhibition: Principles of operation during
830 selective attention. *Journal of Experimental Psychology: Human Perception and*
831 *Performance*, 39(1), 245–256. <https://doi.org/10.1037/a0027922>
- 832

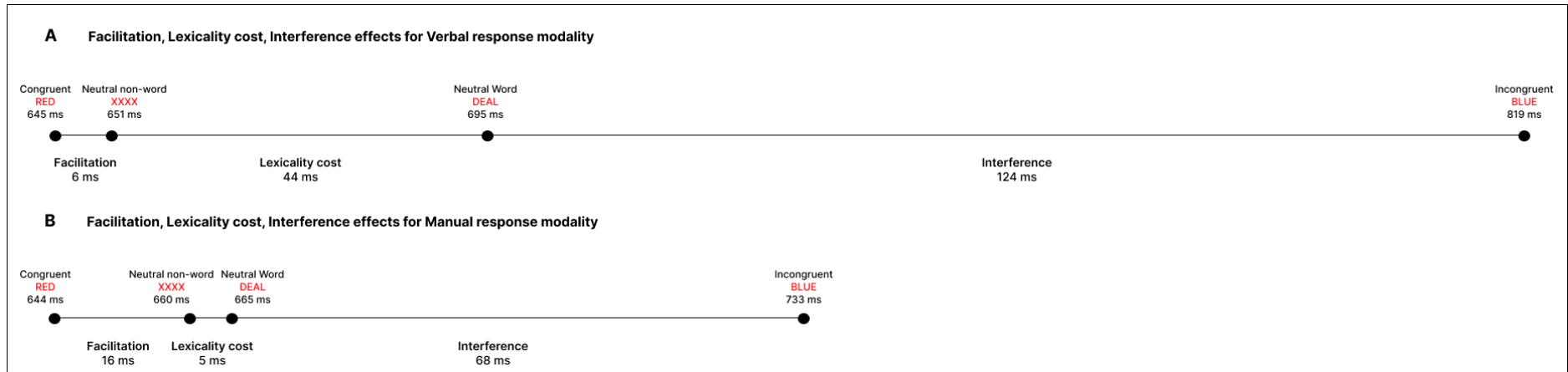


Figure 1. Contribution of facilitation, lexicality cost and interference effects to the overall Stroop/congruency effect.

Note. Original figure based on data from Augustinova et al. (2019) reported with Verbal (Panel A) and Manual (Panel B) response modality.

Table 1. Studies measuring the Stroop interference (studies that used color-neutral non-word items as baseline)

Authors (year) - Journal	Sample size (N participants, groups)	Physical activity/exercise and its control if used	Item presentation; Response modality; Number of colors	Stroop task and Item-type (with number of trials if available)	Indicators of Stroop performance	Study type with a design	Main findings on Stroop performance only	Main conclusions drawn in the study	Observations and Issues
Buck et al. (2008) – Medicine & Science in Sports & Exercise	74 participants (41 boys, 33 girls; M _{age} =9.3)	Measured level of physical fitness (via aerobic capacity, muscle fitness, body mass index)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	% of correct responses (%CR) completed in 45 sec for each list and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral (non-word) condition)	Correlational using aerobic fitness as predictor of (single measure of) Stroop performance	Positive association between aerobic fitness and %CR in each list, but no such association with the Stroop interference	Benefit of fitness for cognition and notably executive control (interference control) in preadolescent children	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) No direct indicator of inhibition (Issue 7)
Ishihara et al. (2018) – Developmental Science	68 participants, all tennis players (M _{age} =9.6)	Measured level of physical fitness (via accelerometry)	Computerized item-by-item presentation; Manual (keys on a computer keyboard); 4 (red, yellow, blue, and green)	2 blocks each consisting of 64 trials: 1 block with color-neutral items (colored patches) and 1 block with color-words items printed in incongruent colors (1-min rest between the two blocks, 128 trials in total)	%ER, mean RTs and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral (patch) condition)	Correlational using physical fitness as predictor of (single measure of) Stroop performance	Negative association between physical fitness and magnitudes of interference that is specifically due to reduction of RTs to incongruent items only	Benefit of fitness (but not of daily physical activity or tennis experience) on inhibitory control	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) No direct indicator of inhibitory control (Issue 7)
van der Niet et al. (2015) – Journal of Science and Medicine in Sport	80 participants (8-12 years old)	Measured level of physical fitness and time spent in sedentary behavior (via accelerometry)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored rectangles)	Number of CR completed in 45 sec for each list (not reported) and the magnitudes of the Stroop interference ratio (= CR in the color-incongruent condition / CR in the color-neutral condition => higher	Correlational using physical fitness and time spent in sedentary behavior as predictors of (single measure of) Stroop performance	Small negative association between time spent in sedentary behavior and interference ratios	The more time children spent in sedentary behavior, the worse is their performance on this inhibition task	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) Additional control for response speed via interference ratio

Running head: Examining attentional boost in the Stroop task

					ratio=better performance)				Conclusion about inhibition not supported by the results (Issue 7)
Alves et al. (2014) – Perceptual & Motor Skills: Learning & Memory	22 participants (M _{age} =53.7)	10x1-min cycling bouts interspersed by 1 min. active (acute aerobic exercise) vs 10 min instruction about the benefits of regular exercise followed by 15 min of stretching (active control activity)	Paper and pencil/list presentation (Stroop Victoria); Vocal; 3	3 color-naming lists each consisting of 24 colored items: 1 with color-neutral items (colored dots), 1 with color-neutral words items, and 1 with color-words items printed in incongruent colors	% of correct responses (%CR) and completion times (in sec) in each list and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral dot (non-word) condition)	Experimental comparing Stroop performance collected after acute aerobic exercise AND after control activity (administered in counter-balanced order 7 days apart)	Reduced magnitudes of the Stroop interference after exercise (but not after control) activity due to reduced RTs to color-incongruent items only	Results suggest that acute aerobic exercise improves selective attention and susceptibility to interference from conflicting stimuli	Low N participants Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) Possible practice effect (Issue 6)
Alves et al. (2012) – Journal of Sport and Exercise Psychology	42 female participants (M _{age} =52)	30 min walking (acute aerobic exercise) vs 2 sets of 15 repetitions for 6 exercises of the lower and upper limbs (acute resistance exercise) vs 10-min instruction about the benefits of regular exercise followed by a 15 min of stretching (active control activity)	Paper and pencil/list presentation (Stroop Victoria); Vocal; 4	3 color-naming lists each consisting of 24 colored items: 1 with color-neutral items (colored dots), 1 with color-neutral words items, and 1 with color-words items printed in incongruent colors	% of correct responses (%CR) and completion times (in sec) in each list and magnitudes of two ratios: color-neutral words/dot and incongruent word/color-neutral word	Experimental comparing Stroop performance collected after aerobic exercise AND strength exercise AND control activity (administered in counter-balanced order 7 days apart)	Reduced completion times to color-neutral and incongruent items and reduced interference ratios after both aerobic and strength exercise (but not after control) activity	Both types of exercises improve selective attention and susceptibility to interference from conflicting stimuli. Both types of exercises are beneficial for executive functioning	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) Computation of ratios to account for change in response speed Possible practice effect (Issue 6)
van der Niet et al. (2016) – Pediatric Exercise Science	105 participants (8-12 years old): assigned to physical training (PT, N=53) vs. control group (N=52)	22-weeks physical training program during lunch time (30-min sessions, twice a week) in the PT group vs. the same lunch routine (passive control group)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists (no information about the contingencies): 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items	Number of CR completed in 45 sec for each list (not reported) and interference ratio (= CR in the color-incongruent condition / CR in the color-neutral condition => higher ratio=better performance)	Experimental* analyzing pre- vs. post-intervention Stroop performance (collected 22 weeks apart) as a function of PT vs. control group (*children group assignment based on parental consent)	Post-intervention reduction of interference ratio in the PT (vs. control) group	Benefit of physical activity program on executive functioning (notably inhibition) in primary school children	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) Additional control for response speed via interference ratios

Running head: Examining attentional boost in the Stroop task

				(colored rectangles)					
Best et al. (2014) – Frontiers in Human Neurosciences	125 female participants (M _{age} =69.5) assigned to Once-(N=44) vs. Twice-weekly training group (N=45) vs. Control group (N=36)	12-month high-intensity resistance training program (60-min classes) once vs twice a week vs balance, toning and relaxation program (active control group)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists each consisting of 80 items: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Completion times (in sec) in each list and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral (non-word) condition). And a Δ between Stroop interference at month 12 and month 0	Experimental analyzing pre- and post-intervention Stroop performance (collected 12 months apart) as a function of once vs. twice a week exercise group vs active control	Reduced magnitudes the Stroop interference after intervention as compared to before in all groups (including control) positively related to adherence to physical activity during the follow-up period (months 13 to 25)	Improving executive functions (in particular, inhibition of prepotent responses) plays an important role in whether older women maintain higher levels of physical activity following exercise training	No direct indicator of inhibition (Issue 7) Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) No results about changes in the incongruent condition (possibly Issue 3) No measure/analysis of errors (possibly Issue 4) Lack of control for changes in response speed (possibly Issue 5) No direct indicator of inhibition (Issue 7)
Liu-Ambrose et al. (2010) – Archives of Internal Medicine	135 female participants (M _{age} =82.2) assigned to Once-(N=47) vs. Twice-weekly training group (N=46) vs. Control group (N=42)	12-month high-intensity resistance training (60-min classes) once vs twice a week vs balance, toning and relaxation program (active control group)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists each consisting of 80 items: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Completion times (in sec) in each list (not reported in the MS) and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral (non-word) condition)	Experimental analyzing pre- vs. midterm vs. post-intervention Stroop performance (collected 6 months apart) as a function of once vs. twice a week exercise group vs active control	Reduced magnitudes the Stroop interference after intervention as compared to before following both the once- and twice-weekly training program (but not after control activity)	Benefit of resistance training on executive functioning of selective attention and conflict resolution in senior women	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) No results about changes in the incongruent condition (possibly Issue 3) No measure/analysis of errors (possibly Issue 4)

Running head: Examining attentional boost in the Stroop task

									No direct indicator of inhibition (Issue 7)
Sipilä et al. (2021) – Scandinavian Journal of Medicine and Science in Sports	314 participants (M _{age} =74.5) assigned to Physical and Cognitive training group (PTCT, N=155) vs Physical training (PT, N=159) group	12-month PT (aerobic, balance, resistance exercise, 2-3 sessions/week) vs 12-month PTCT (PT with additional computer training of executive functions (3-4 sessions/week); no active or passive control group)	Paper and pencil/list presentation; Vocal; Not mentioned	3 lists each consisting of 72 items: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs) (referred to as congruent)	Completion times (in sec) in each list (not reported in the MS) and magnitudes of the Stroop interference (= time difference between color-incongruent and color-neutral (non-word) condition)	Experimental analyzing pre- vs. midterm vs. post-intervention Stroop performance (collected 6 months apart) as a function of training type (physical + cognitive vs. physical training only)	Post-intervention reduction of interference in both training groups but greater in the PTCT than the PT group	Greater benefit of combined physical and cognitive training on response inhibition than that of physical training alone	Magnitude of Stroop interference inflated by the use of a non-word baseline, creating a lexicality cost confound (Issue 1) No results about changes in the incongruent condition (possibly Issue 3) No measure/analysis of errors (possibly Issue 4) Lack of control for changes in response speed (possibly Issue 5) No direct indicator of inhibition (Issue 7)

Note. Light gray background indicates an “overinterpretation” of reported results and dark gray indicates a clear mismatch between empirical results and their interpretation.

Supplementary Materials

Table of contents

Figure 1S..... 2
Table 1S 3
References (Table 1S)..... 10
Table 2S 12
References (Table 2S)..... 29

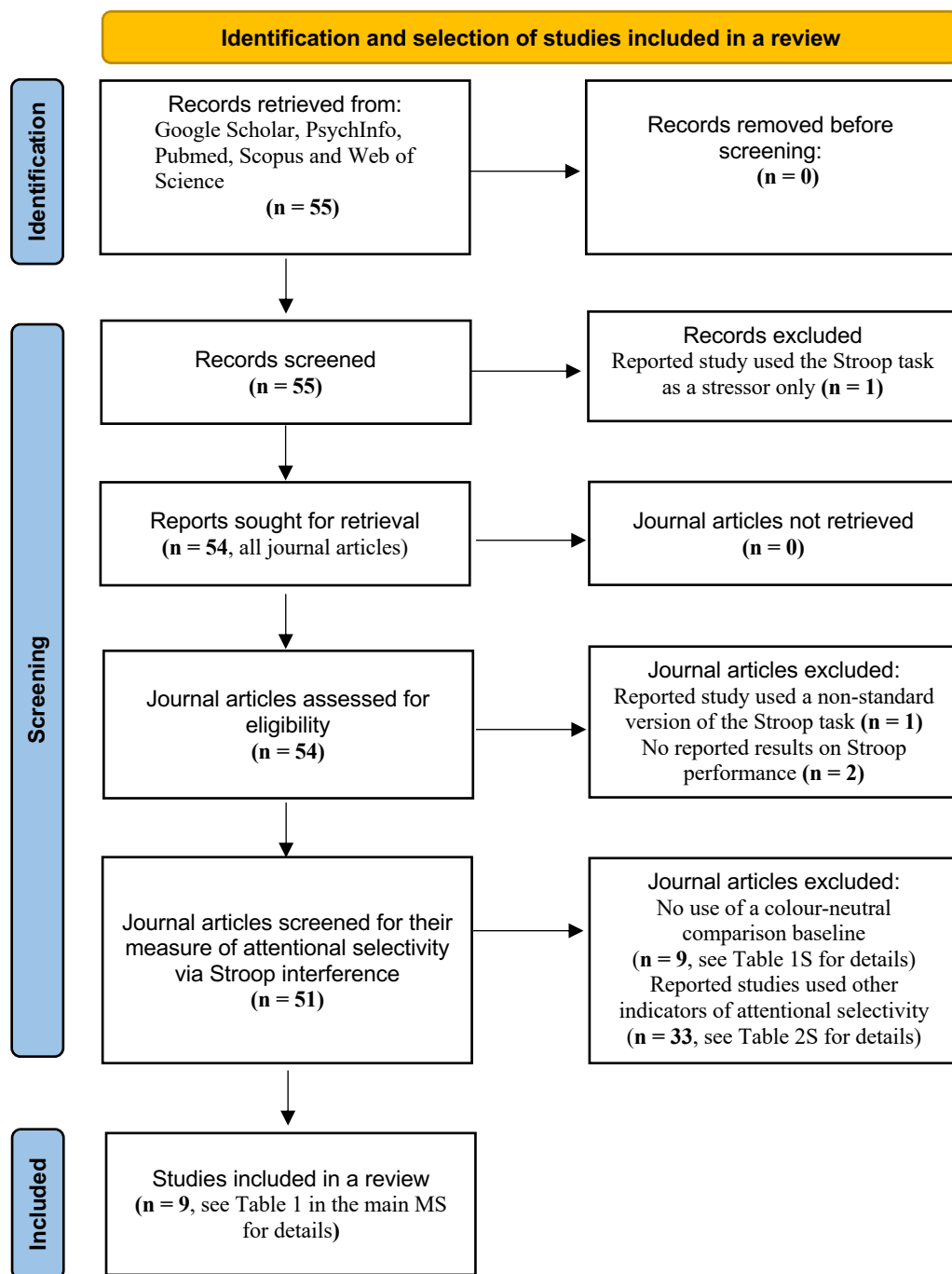


Figure 1S. Flowchart (adapted from Page et al., 2021) depicting the selection process of the included studies. The results of the included studies are discussed in the main text of the manuscript, while the results of excluded studies are displayed in the supplementary tables 1S and 2S in the Supplementary Materials.

References (Figure 1S)

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372.

Table 1S. Studies measuring the overall Stroop effect (studies that used congruent baseline items as baseline)

Authors (year) - Journal	Sample size (N participants, groups)	Physical activity/exercise and its control if used	Item presentation; Response modality; Number of colors	Stroop task and Item-type (with number of trials if available)	Indicators of Stroop performance	Study type with a design	Main findings on Stroop performance only	Main conclusions drawn in the study	Observations and issues
Bento-Torres et al. (2019) – Journal of Sport and Health Science	48 participants (M _{age} =24.6)	Measured level of cardiorespiratory fitness (CRF) (via a field test) and physical activity (via accelerometry)	Computerized item-by-item presentation; Manual (Pressing a key on a standard keyboard with the index, middle, and ring fingers of the right hand); Not mentioned	A single block consisting of 54 color-words items printed in congruent color, 48 words items printed in incongruent color, and 63 color-neutral words items printed in color, presented randomly (165 trials in total)	%ER, RTs (in ms), SD (indicator of inter-individual variability, IIV) and the Stroop effect ratio (= (RTs to color-incongruent trials - color-congruent trials) / RTs to color-congruent trials))	Correlational using physical fitness and CRF as predictors of (single measure of) Stroop performance	No association between physical activity and IIV for either Stroop condition and Stroop effect ratios No other associations with Stroop performance are reported	No conclusions with respect to Stroop performance	Magnitude of Stroop interference inflated by the use of a color-congruent baseline (despite the use of a color-neutral words condition), creating a facilitation confound (Issue 1) No control of the color-congruent items contingency (Issue 2) Computation of ratios to account for change in response speed
Li et al. (2021) – Brain Sciences	180 participants (M _{age} =20.2)	Measured level of physical activity (PA) (via the International Physical Activity Questionnaire, IPAQ)	Computerized item-by-item presentation; Manual (color-coded keys on a computer keyboard); 4 (blue, green, red, and green)	2 blocks each consisting of 8 color-words items printed in congruent color and 24 color-words items printed in incongruent color, presented randomly (64 trials in total)	%ER, RTs (in ms), and magnitudes of the Stroop effect (= RTs to color-incongruent trials – color-congruent trials)	Correlational using physical activity as predictor of (single measure of) Stroop performance	Negative association between physical activity and the magnitude of the Stroop effect. No significant association between physical activity level and RTs to congruent and	A higher level of physical activity is linked to better inhibitory control in young adults	Magnitude of Stroop interference inflated by the use of a color-congruent baseline, creating a facilitation confound (Issue 1) No specific correlation with RTs on

Running head: Examining attentional boost in the Stroop task

							incongruent conditions. No significant changes in the %ER		incongruent items (Issue 3) No direct indicator of inhibitory control (Issue 7)
Peven et al. (2018) – Journal of Cognitive Enhancement	96 participants (M _{age} =65.8)	Measured level of physical activity (PA, via accelerometry) resulting in the identification of two categories of activity bouts duration (< 10 or ≥ 10 min)	Computerized item-by-item presentation; not mentioned; not mentioned	54 color-words items printed in congruent color, 63 color-neutral words items printed in color and 48 color-words items printed in incongruent color, presented randomly (165 trials in total, no information about the number of block and item randomization)	%ER (reported, but integrated into the statistical analyses), RTs (in ms) and magnitudes the Stroop effect (= RTs to color-incongruent trials - color-congruent trials)	Correlational using physical activity and number of < 10 or ≥ 10 min bouts of activity as predictors of (single measure of) Stroop performance	Hierarchical linear regression models revealed that higher levels of PA and more bouts of ≥ 10 min were associated with smaller magnitude of the Stroop effect and lower RTs to incongruent items	Physical activity level is associated with better executive function (inhibitory function)	Magnitude of Stroop interference inflated by the use of a color-congruent baseline (despite the use of a color-neutral words condition), creating a facilitation confound (Issue 1) No control of the color-congruent items contingency (Issue 2) No analysis of errors (possibly Issue 4) No direct indicator of inhibitory control (Issue 7)
Stillman et al. (2016) – PloS ONE	41 participants (M _{age} =23.9)	Measured level of cardiorespiratory fitness (CRF) (via exercise test) and of physical activity (PA) (via accelerometry)	Computerized item-by-item; Manual (response glove); Not mentioned	A single block consisting of 40 color-words items printed in congruent color, 40 color-neutral words items printed in color and 40 color-words items printed in incongruent color, presented	%ER, RTs (in ms, not reported in the results), no errors (due to “high accuracy”) and magnitudes of and Stroop effect ratios (= (RTs to color-incongruent trials - color-congruent trials)	Correlational using PA as predictor of (single measure of) Stroop performance (no analyses of CRF reported)	Negative association between physical activity and the magnitude of the Stroop effect ratio. No analyses of RTs and %ER	Higher physical activity levels are associated with better executive functioning	Magnitude of Stroop interference inflated by the use of a color-congruent baseline (despite the use of a color-neutral words condition), creating a

Running head: Examining attentional boost in the Stroop task

				randomly (120 trials in total)	/ RTs to color-congruent trials))				<p>facilitation confound (Issue 1)</p> <p>No control of the color-congruent items contingency (Issue 2)</p> <p>No results about changes in the incongruent condition (possibly Issue 3)</p> <p>Computation of ratios to account for change in response speed</p> <p>No direct indicator of inhibition (Issue 7)</p> <p>Low N participants for a correlational study</p>
Peven et al. (2019) – Brain Plasticity	50 participants (M _{age} =25.2)	Measured level of cardiorespiratory fitness (CRF) via exercise test resulting in the comparison of higher vs lower CRF group	Computerized item-by-item; Manual (response glove); Not mentioned	54 color-words items printed in congruent color, 63 color-neutral words items printed in color and 48 color-words items printed in incongruent color, presented randomly (165 trials in total, no information about the number of block and item randomization)	RTs (in ms), no errors (due to “high accuracy”) and magnitudes of and Stroop effect ratios (= (RTs to color-incongruent trials - color-congruent trials) / RTs to color-congruent trials)	Experimental* comparing (single measure of) Stroop performance as a function of higher or lower CRF group *determination of groups based on the measured levels of cardiorespiratory fitness	Negative association between CRF and RTs to color-incongruent and -neutral conditions as compared to lower CRF. No such effect found on color-congruent items and Stroop effect ratio	Cardiorespiratory fitness is associated with enhanced executive functions and notably selective attention	<p>Magnitude of Stroop interference inflated by the use of a color-congruent baseline (despite the use of a color-neutral words condition), creating a facilitation confound (Issue 1)</p> <p>No control of the color-congruent</p>

Running head: Examining attentional boost in the Stroop task

								items contingency (Issue 2)
								Computation of ratios to account for change in response speed
								No direct indicator of inhibition (Issue 7)
								Low N participants
Brush et al. (2016) – Journal of Sport and Exercise Psychology	28 participants (M _{age} =20.5)	Acute resistance exercise of low (40% 10-RM), moderate (70% 10-RM), high intensity (100% 10-RM) vs watching a 45-min instructional video on how to perform various resistance exercises)	Computerized item-by-item presentation; Not mentioned; Not mentioned	2 color-naming blocks each consisting of 36 color-words items printed in congruent color, and 36 color-words items printed in incongruent color, presented randomly (144 trials in total)	%ER, RTs (in ms), and magnitude of the Stroop effect (= RTs/%ER to color-incongruent trials - color-congruent trials)	Experimental comparing Stroop performance collected 15 and 180mn after acute resistance exercise AND after control activity (administered in counter-balanced order at least 48h apart)	At 15 min postexercise: Post-intervention reduction of the Stroop effect increased linearly as a function training intensity (but the effect is not driven by the change in RTs to incongruent trials) No effect at 180 min postexercise	Limited and task-specific influence of acute resistance exercise on executive function, especially inhibition in healthy young adults
								Magnitude of Stroop interference inflated by the use of a color-congruent baseline, creating a facilitation confound (Issue 1)
								No control of the color-congruent items contingency (Issue 2)
								No specific correlation with RTs on incongruent items (Issue 3)
								Possible practice issue (Issue 6)
								No direct indicator of inhibition (Issue 7)

Running head: Examining attentional boost in the Stroop task

									No pre-exercise measure of Stroop performance Low N participant for a within between interaction
Dimitrova et al. (2017) – Aging Clinical and Experimental Research	58 participants: younger ($M_{age}=23.2$, $N=29$) and older ($M_{age}=70.7$, $N=29$) adults. Participants of each age group were randomly assigned to exergaming vs standard physical exercise	20 min of cycling with a cognitive strategy demand (Acute aerobic “exergaming” exercise) vs 20-min exercise on a virtual reality track without cognitive stimulation (standard physical exercise serving as active control)	Computerized item-by-item presentation; Manual (color-coded keys on a computer keyboard); 3 (red, blue, green)	2 blocks each consisting of 138 color-words items printed in congruent color and 60 color-words items printed in incongruent color (198 trials in total)	%ER, RTs (in ms), and magnitude of the Stroop effect (= RTs to color-incongruent trials - color-congruent trials)	Experimental* analyzing pre- vs. post-intervention Stroop performance and of training type (exergaming vs standard physical training only) *Comparison of 2 age groups	Post-intervention boost in response speed in both age-groups with a greater magnitude in older adults. No significant changes in magnitudes of the Stroop effect	The two forms of acute exercise interventions had generalized benefits on cognitive performance, especially on speed of processing in older adults	Magnitude of Stroop interference inflated by the use of a color-congruent baseline, creating a facilitation confound (Issue 1) No control for the color-congruent items contingency, leading to a facilitation issue due to contingent responding (Issue 2) Lack of control for changes in response speed (possibly Issue 5) Possible practice effect (Issue 6) but greater benefit in older adults argues against simple practice effect Low N participants in the 4 groups
Hsieh et al. (2018) –	44 participants: young ($M_{age}=24$,	20-min of moderate intensity aerobic	Computerized item-by-item	6 blocks each consisting of 42	%ER, RTs (in ms), and	Experimental* analyzing pre- vs	Post-intervention	Acute aerobic exercise is	Magnitude of Stroop

Running head: Examining attentional boost in the Stroop task

Journal of Clinical Medicine	N=24) and older (M _{age} =70, N=20) adults, all assigned to both aerobic exercise and control activity	exercise at 60-70% of HRR (Acute aerobic exercise) vs 20-min of sitting quietly and watch a video related to sport science (passive control activity)	presentation; Manual (color-coded keys on a computer keyboard); 3 (red, blue, green)	color-words items printed in congruent color and 18 color-words items printed in incongruent color (360 trials in total)	magnitudes of the Stroop effect (= RTs/%ER to color-incongruent trials - color-congruent trials)	post-intervention Stroop performance collected for both acute exercise and control conditions (counterbalanced and administered at least 96h apart) *Comparison of 2 age groups	reduction of the Stroop effect in the exercise compared to control condition (however due to shorter RTs to incongruent trials)	beneficial to inhibition in young and older adults	interference inflated by the use of a color-congruent baseline, creating a facilitation confound (Issue 1) No control for the color-congruent items contingency, leading to a facilitation issue due to contingent responding (Issue 2) Lack of control for changes in response speed (possibly Issue 5) Possible practice effect (Issue 6) although differences in interventions suggest more than that No direct indicator of inhibition (Issue 7) Low N participants in both age groups
Konijnenberg and Fredriksen (2018) – Scandinavian Journal of Public Health	1173 participants (AGE=7-12) assigned to physical training (PT, N=793) vs. control group. Only the 9-12 years olds	12-month physical training program (45 min of moderate to high (25–30% of the time) intensity physical activity a day) vs. keeping the same routine	Computerized item-by-item presentation; Manual (color-coded keys on a computer keyboard); 4	A single block consisting of 84 trials including color-words items printed in congruent color, color-words items printed in	%ER, RTs (in ms), and magnitudes of the Stroop effect (= RTs/%ER to color-incongruent	Experimental analyzing pre- vs. post-intervention Stroop performance as function of physical training vs. control	Post-intervention reduction of the Stroop effect in both the exercise and control groups (due to the overall	The physical activity intervention program had no positive effect on children’s executive control	Magnitude of Stroop interference inflated by the use of a color-congruent baseline, creating a facilitation

Running head: Examining attentional boost in the Stroop task

performed the Stroop task, the 7-8 years olds performed the Flanker task (no information about the contingencies in these age groups)	(pseudo-passive control)	(blue, green, red, and black)	incongruent color and color-neutral items (colored squares) (contingencies not mentioned), presented randomly	trials – color-congruent trials)	reduction of RTs to all items)	<p>confound (Issue 1)</p> <p>Possible contingency issue (Issue 2)</p> <p>Improvement consistent with a normal developmental trajectory; Lack of control for changes in response speed (possibly Issue 5)</p> <p>No direct indicator of inhibition (Issue 7)</p>
---	---------------------------------	-------------------------------	---	----------------------------------	--------------------------------	---

Note. Light gray background indicates an “overinterpretation” of reported results and dark grey indicates a clear mismatch between empirical results and their interpretation.

References (Table 1S)

- Bento-Torres, J., Bento-Torres, N. V. O., Stillman, C. M., Grove, G. A., Huang, H., Uyar, F., Watt, J. C., Wollam, M. E., & Erickson, K. I. (2019). Associations between cardiorespiratory fitness, physical activity, intraindividual variability in behavior, and cingulate cortex in younger adults. *Journal of Sport and Health Science*, 8(4), 315–324. <https://doi.org/10.1016/j.jshs.2019.03.004>
- Brush, C. J., Olson, R. L., Ehmann, P. J., Osovsky, S., & Alderman, B. L. (2016). Dose–Response and Time Course Effects of Acute Resistance Exercise on Executive Function. *Journal of Sport and Exercise Psychology*, 38(4), 396–408. <https://doi.org/10.1123/jsep.2016-0027>
- Dimitrova, J., Hogan, M., Khader, P., O’Hora, D., Kilmartin, L., Walsh, J. C., Roche, R., & Anderson-Hanley, C. (2017). Comparing the effects of an acute bout of physical exercise with an acute bout of interactive mental and physical exercise on electrophysiology and executive functioning in younger and older adults. *Aging Clinical and Experimental Research*, 29(5), 959–967. <https://doi.org/10.1007/s40520-016-0683-6>
- Hsieh, S.-S., Huang, C.-J., Wu, C.-T., Chang, Y.-K., & Hung, T.-M. (2018). Acute Exercise Facilitates the N450 Inhibition Marker and P3 Attention Marker during Stroop Test in Young and Older Adults. *Journal of Clinical Medicine*, 7(11), 391. <https://doi.org/10.3390/jcm7110391>
- Konijnenberg, C., & Fredriksen, P. M. (2018). The effects of a school-based physical activity intervention programme on children’s executive control: The Health Oriented Pedagogical Project (HOPP). *Scandinavian Journal of Public Health*, 46(21_suppl), 82–91. <https://doi.org/10.1177/1403494818767823>

- Li, L., Yu, Q., Zhao, W., Herold, F., Cheval, B., Kong, Z., Li, J., Mueller, N., Kramer, A. F., Cui, J., Pan, H., Zhan, Z., Hui, M., & Zou, L. (2021). Physical Activity and Inhibitory Control: The Mediating Role of Sleep Quality and Sleep Efficiency. *Brain Sciences*, *11*(5), Article 5. <https://doi.org/10.3390/brainsci11050664>
- Peven, J. C., Grove, G. A., Jakicic, J. M., Alessi, M. G., & Erickson, K. I. (2018). Associations Between Short and Long Bouts of Physical Activity with Executive Function in Older Adults. *Journal of Cognitive Enhancement*, *2*(2), 137–145. <https://doi.org/10.1007/s41465-018-0080-5>
- Peven, J. C., Litz, G. A., Brown, B., Xie, X., Grove, G. A., Watt, J. C., & Erickson, K. I. (2019). Higher Cardiorespiratory Fitness is Associated with Reduced Functional Brain Connectivity During Performance of the Stroop Task. *Brain Plasticity*, *5*(1), 57–67. <https://doi.org/10.3233/BPL-190085>
- Stillman, C. M., Watt, J. C., Grove, G. A., Wollam, M. E., Uyar, F., Mataro, M., Cohen, N. J., Howard, D. V., Howard, J. H., & Erickson, K. I. (2016). Physical Activity Is Associated with Reduced Implicit Learning but Enhanced Relational Memory and Executive Functioning in Young Adults. *PLOS ONE*, *11*(9), e0162100. <https://doi.org/10.1371/journal.pone.016210>

Table 2S. Studies using other (perhaps less conclusive) performance indicators of attentional selectivity

Authors, year - Journal	Sample size (N participants, groups)	Physical activity/exercise and its control if used	Item presentation; Response modality; Number of colors	Stroop task and Item-type (with number of trials if available)	Indicators of Stroop performance	Study type with a design	Main findings on Stroop performance only	Main conclusions drawn in the study	Observations and Issues
Aadland et al. (2014) – Mental Health and Physical Activity	697 participants (357 girls and 340 boys, M _{age} =10.2)	Measured level of physical activity , sedentary time, aerobic fitness and motor skills	Paper and pencil/list presentation; Vocal; 3 (red, blue, green)	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Number of correct responses in 45 seconds in the incongruent color-naming list only (no use of baseline)	Correlational using physical activity sedentary time, aerobic fitness and motor skills levels as predictors of (single measure of) Stroop performance	No association between physical activity and Stroop performance , Positive associations between motor skills sedentary , time, aerobic fitness and Stroop performance	Physical activity (as opposed to motor skills) is not a way to improve executive functions , especially inhibition	No comparison to a non-color words neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)
Bixby et al. (2007) – Medicine & Science in Sports & Exercise	120 participants (M _{age} =78.9)	Measured level of physical activity (via the Yale Physical Activity Survey, YPAS)	Not mentioned; Vocal; 3 (red, green, blue)	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Number of correct responses in 45 seconds for each condition; Calculation of a “ predicted color-word score ”: word reading score*neutral score / word reading score + neutral score); Calculation of an “ interference score ”: Incongruent score – predicted color score (A higher score reflecting more resistance to interference)	Correlational using physical activity as predictor of (single measure of) Stroop performance	Positive association between physical activity and performance in the incongruent items and the “interference score” No association between physical activity and performance in the word-reading and neutral items	Higher physical activity levels are beneficial to inhibitory executive function in older adults	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)
Prakash et al. (2011) – Frontiers in Psychology	70 participants (M _{age} =65)	Measured level of cardiorespiratory	Computerized; Manual (keyboard	A single block consisting of 36 non-color words	RTs (in ms) and error rates	Correlational using CRF as predictor of	Negative associations between CFR and	Higher levels of fitness are associated with	No comparison to a non-color

Running head: Examining attentional boost in the Stroop task

		fitness (CRF via a field test)	with three colored buttons); 3 (green, purple, red)	items printed in color, 36 color-words items printed in congruent color, 36 eligible color-words items printed in incongruent color, 36 ineligible color-words items printed in incongruent color presented in a counterbalanced order (144 trials in total)	for each condition	(single measure of) Stroop performance	both RTs and error rates to the incongruent-eligible condition	better attentional control	words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2)
Ramnath et al. (2018) – PLoS ONE	70 participants (M _{age} =71)	Measured level of physical activity (via the Yale Physical Activity Survey, YPAS) and measured level of fitness and functional performance (via a field test battery assessing grip strength, functional reach, and balance)	Computerized; Manual (color-coded response button); 4 (red, blue, green, yellow)	2 blocks consisting of a word-reading condition (color words printed in gray ink), and a color-naming conditions with color-words items printed in incongruent colors (no information about the contingencies)	Mean RTs and accuracy for the incongruent condition (correct responses, errors, missed responses)	Correlational using physical activity, physical fitness and functional performance as predictors of (single measure of) Stroop performance	Positive association between functional performance and the number of correct responses to the incongruent items. No association between physical activity levels and all Stroop Task variables	Individuals with higher functional performance have better executive functions	No comparison to a non-color words neutral baseline (Issue 1) Possible contingency issue (Issue 2) No direct indicator of inhibition (Issue 7)
Salas-Gomez et al. (2020) – Frontiers in Human Neuroscience	206 participants (university students, M _{age} =19.6)	Measured level of physical activity (PA, via the International Physical Activity Questionnaire Short Version, IPAQ-SF)	Paper and pencil; Vocal; 3 (red, blue, green)	3 lists (no information about the contingencies): 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Number of CR completed in 45 sec for each list	Correlational using PA as predictor of (single measure of) Stroop performance	Positive association between PA level and the number of CR in each Stroop condition	Beneficial relationship between PA and cognition in young adults. Physical activity might improve aspects of executive function such as the ability to inhibit stimuli that trigger automatic responses, cognitive control ability, and mental flexibility in university student	No comparison to a non-color words neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)
Aly and Kojima (2020) – Psychophysiology	30 participants (M _{age} =21)	Measured level of physical activity (via International Physical Activity Questionnaire, IPAQ)	Computerized item-by-item presentation; Manual (Response box)	A single block consisting of 60 color-words items printed in congruent color and	%ER, mean RTs (in msec) and intraindividual variability (as	Experimental* comparing (single measure of) Stroop performance as a	General decrease in response speed (i.e., shortened RTs in each condition and item-type) and	Physical activity is associated with lower distractibility and better inhibition in the Stroop task	No comparison to a non-color words neutral

Running head: Examining attentional boost in the Stroop task

		resulting in comparison of more (N=15) vs. less (N=15) active group	with 4 response keys labeled (in Kanji-characters) as red, green, blue and yellow)	60 color-words items printed in incongruent color, requiring either color or word meaning identification (cued by a square vs. diamond flanker), presented randomly (240 items in total)	reflected by mean SDs) observed separately for each condition and item-type	function of more or less active group *determination of groups based on measured levels of PA	Reduced variability in more active vs. less group No significant differences in errors		baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) No direct indicator of inhibition (Issue 7)
Boucard et al. (2012) – Journal of Sport and Exercise Psychology	123 participants: Young (Age=18–28, N=63), young-old (60–70 years old, N=30), and old (71–81 years old, N=30) adults. Each age group subdivided into active and non-active subgroups	Median split applied to measured level of physical fitness (via accelerometry) resulting in comparison of active and non-active subgroups	Paper and pencil/list presentation; Vocal; 3 (red, green, blue)	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	%CR completed in 45 sec for each list and magnitudes of “interference” computed as [(word reading score × neutral score) / (word reading score + neutral score)] ratio. Conversion of the raw scores into Z-scores	Experimental* comparing (single measure of) Stroop performance as function of age and active vs non-active group Additional analyses of mediation *determination of groups based on measured levels of physical fitness	Reduced magnitudes of the “interference” in active as compared non-active group - effect mediated by the level of cardiovascular fitness	Effect of physical activity on specific executive functions , especially inhibition . Positive link between physical activity, cardiovascular fitness , and inhibition in aging	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) Calculation of the interference not based on the incongruent condition No direct indicator of inhibition (Issue 7)
Chang et al. (2017) – Journal of Sport and Health Science	60 participants (M _{age} =21.1): Endurance (N=20), martial arts (N=20), and recreational (N=20) athletes	Measured levels of physical fitness (cardiovascular fitness, muscular endurance and flexibility, agility and power of motor fitness) (via field tests)	Paper and pencil/list presentation; Vocal; Not mentioned	5 lists each consisting of 50 trials: 1 word-reading list with color-words items printed in black; 4 color-naming lists 1 with color-words items printed in congruent colors 1 with color-words items printed in incongruent colors, 1 with color-neutral	Completion times for each condition	Experimental* comparing (single measure of) Stroop performance as function of Endurance vs martial arts vs recreational sport group *determination of groups based on sport practice	No effect of physical fitness on any Stroop condition	No positive relationship between sport expertise and general cognitive abilities (notably perceptual speed and executive function as measured by the Stroop task)	No computation of the Stroop interference despite the use of a color-neutral words condition No control of the color-congruent items

Running head: Examining attentional boost in the Stroop task

				word items printed in color and 1 with color-neutral items (colored squares)					contingency (Issue 2)
									No analysis of errors (possibly Issue 4)
Gajewski and Falkenstein (2015) – Brain and Cognition	40 participants (M _{age} =73.2):	Measured level of physical activity (assessed via the Lüdenscheid Activity Questionnaire), and measured level of physical fitness (assessed via a test on a bicycle ergometer), resulting in the identification of a “ highly active ” group (N=20) and a “ low active ” group (N=20) The groups were matched regarding gender, age, and education	Computerized version, item-by-item presentation + Paper and pencil/list presentation; Manual for the computerized version with buttons corresponding to each of the four colors; Not mentioned for the paper and pencil version; 4 (red, green, yellow, blue)	2 blocks each consisting of 52 trials: 26 color-words items printed in congruent colors and 26 color-words items printed in incongruent colors, presented randomly. The first block was a word-reading condition and the second block was a color-naming condition (color-words in colored ink) (104 trials in total) These two blocks were performed in both the paper and pencil and computerized versions of the task	Error rate, mean RTs, and individual SDs. An interference index was calculated only in the paper and pencil version by subtracting the time to perform the word-reading condition (block 1) to the time to perform the color-naming condition (block 2)	Experimental* comparing (single measure of) Stroop performance as function of highly active vs low active group *determination of groups based on measured levels of physical activity	Higher error rates in the color-naming block and incongruent items for the low-active group No group difference in the RTs Higher interference index for the low active group in comparison to their active counterparts	These results question the assumption that physical activity enhances inhibitory processing ; rather, physical activity is associated with lower distractibility in Stroop interference	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) Possible practice effect (issue 6) due to the completion of both computerized and paper version No direct indicator of inhibition (Issue 7)
Gajewski et al. (2020) – NeuroImage	246 participants: young (M _{age} =25.1, N=36), middle-aged (M _{age} =46.5, N=58), older (M _{age} =70.6, N=152) adults	Measured level of physical activity in the past two years (assessed with the Lüdenscheid Activity Questionnaire), and measured level of physical fitness (assessed with a test on a bicycle ergometer), resulting in the identification of	Computerized version + Paper and pencil/list presentation; Manual for the computerized version with buttons corresponding to each of the four colors;	3 blocks: 1 word-reading block consisting of 52 trials: 26 color-words items printed in congruent colors and 26 color-words items printed in incongruent colors, presented randomly;	Error rate, mean RTs, and individual SDs. An interference index was calculated only in the paper and pencil version by subtracting the time to perform the word-	Experimental* comparing (single measure of) Stroop performance as function of young vs middle-aged vs older-low performing vs older-mid performing vs	High performing old adults did not differ from middle aged-adults in RTs to incongruent items and interference index	Old high performing adults display executive function levels that are comparable to middle-aged adults	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent

Running head: Examining attentional boost in the Stroop task

		low-, mid- and high-performing subgroups in old adults	Not mentioned for the paper and pencil version; 4 (red, green, yellow, blue)	1 color-naming block consisting of 146 trials: 26 color-words items printed in congruent colors incongruent and 26 color-words items printed in incongruent colors; 1 switching block consisting of 146 trials including word-reading and color-naming items in function of a cue on the screen (146 items). The two first blocks were also performed in the paper and pencil version	reading condition to the time to perform the color-naming condition	older-high performing adults group		items contingency (Issue 2)
						*determination of groups based on age and measured levels of physical activity		Possible practice effect (issue 6) due to the completion of both computerized and paper version
								No control of the color-congruent items contingency (Issue 2)
								No direct indicator of inhibition (Issue 7)
Goenarjo et al. (2020) – International Journal of Environmental Research and Public Health	56 male participants (M _{age} =22.1)	Measured level of physical activity (via the Global Physical Activity Questionnaire, GPAQ), and Measured level of cardiorespiratory fitness (CRF) (via a graded maximal exercise test) resulting in the identification of an “ active ” (N=26) and “ inactive ” (N=30) groups	Computerized; Manual (color-coded keys on a computer keyboard); 4 (red, blue, green, yellow)	3 blocks each consisting of 60 trials: 1 with color-words items printed in congruent colors, 1 with color-words items printed in incongruent colors and 1 switching block with color-words items printed in incongruent color, requiring either color or word meaning in function of a cue on the screen	% of correct responses (%CR) and mean RTs for each condition	Experimental* comparing (single measure of) Stroop performance as function of active vs inactive group	Reduced RTs to the switching items for the active group in comparison to the inactive group No other group difference in RTs in the other conditions and in accuracy	Physical activity levels are beneficial to executive functioning , especially switching but not inhibition (as reflected by the absence of group effect in the incongruent condition)
						*determination of groups based on measured levels of physical activity		No comparison to a non-color words neutral baseline (Issue 1)
								No control of the color-congruent items contingency (Issue 2)
								No direct indicator of inhibition (Issue 7)
								Change in switching performance difficult to interpret in the

Running head: Examining attentional boost in the Stroop task

									absence of change in the incongruent items
Goenarjo et al. (2021) – Physiology & Behavior	24 male participants (M _{age} =60.4), 2 age groups: 55-60 years old (N=14) and 61-69 years old (N=10)	Measured level of cardiorespiratory fitness (CRF) (via a graded maximal exercise test) resulting in the identification of a “ higher-fit ” and “ lower-fit ” subgroups for each age groups	Computerized; Manual (color-coded keys on a computer keyboard); 4 (red, blue, green, yellow)	3 blocks each consisting of 60 trials: 1 with color-words items printed in congruent colors, 1 with color-words items printed in incongruent colors and 1 switching block with color-words items printed in incongruent color, requiring either color or word meaning in function of a cue on the screen	% of correct responses (%CR) and mean RTs for each condition	Experimental* comparing (single measure of) Stroop performance as function of higher-fit vs lower-fit group *determination of groups based on measured levels of cardiorespiratory fitness and age	Reduced RTs to the switching items for the higher-fit 61-69 years old in comparison to lower-fit group No other group difference in RTs in the other conditions and in accuracy	Positive association between cardiorespiratory fitness and executive functions in older adults	No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) Change in switching performance difficult to interpret in the absence of change in the incongruent items Low N participants
O’Brien et al. (2021) – GeroScience	35 participants	Measured level of physical activity (PA) (assessed via a step-count device), Measured level of cardiorespiratory fitness (CRF) (via a graded maximal exercise test) resulting in the identification of an Inactive group (N=16, M _{age} =68) and Active group (N=19, M _{age} =66)	Computerized, item-by-items presentation; Manual (color-coded keys on a computer keyboard); 4 (red, blue, green, yellow)	3 blocks each consisting of 60 trials: 1 with color-words items printed in congruent colors, 1 with color-words items printed in incongruent colors and 1 switching block with color-words items printed in incongruent color, requiring either color or word meaning	Mean RTs (in ms) in each condition, accuracy (not reported in the results because of high accuracy)	Experimental* comparing (single measure of) Stroop performance as function of Inactive vs Active group *determination of groups based on measured levels of PA and CRF	Reduced RTs to switching items for the active group in comparison to the inactive group Negative associations between levels of physical activity and CRF and RTs to the switching condition	Older active adults had superior cognitive flexibility than their inactive peers; habitual physical activity is an important contributor to executive functions in cognitively healthy older adults	No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) Change in switching performance difficult to interpret in the

Running head: Examining attentional boost in the Stroop task

									absence of change in the incongruent items Low N participants
Aguirre-Loaiza et al. (2019) – Frontiers in Psychology, Experiment 1*	54 physically active participants ($M_{age} = 20.7$) assigned to the exercise (N=25) vs. control (N=29) group	45-min session of indoor cycling at 75-85% HRmax (acute aerobic exercise) vs rest (passive control group)	Paper and pencil/list presentation; Vocal; Not mentioned	2 lists each consisting of 84 trials: 1 word-reading list with switch to color-naming for highlighted items (Stroop A) and 1 word-reading list with switch to color-naming on the experiment's signal (Stroop B). No information about the word color congruency of the items	% of correct responses (%CR) and completion times (in sec) in each list	Experimental analyzing pre- and post-test Stroop performance as a function of acute aerobic exercise vs. control group	Post-test improvement in %CR on Stroop A list in both groups ; Post-test improvement in completion time on Stroop B list in control group	Results reflect a cognitive reserve in young active participants; the exercise intensity turned out detrimental to inhibitory control	No comparison to a non-color words neutral baseline (Issue 1) Possible contingency issue (Issue 2) Performance of the control group suggests a practice effect (Issue 6) No direct indicator of inhibition (Issue 7)
Aguirre-Loaiza et al. (2019) – Frontiers in Psychology, Experiment 2*	36 physically inactive participants ($M_{age} = 21.6$) assigned to the exercise (N=10) vs control (N=26) group	30-min session of indoor cycling at 75-85% HRmax (acute aerobic exercise) vs rest (passive control group)	Paper and pencil/list presentation; Vocal; Not mentioned	2 lists each consisting of 84 trials: 1 word-reading list with switch to color-naming for highlighted items (Stroop A) and 1 word-reading list with switch to color-naming on the experiment's signal (Stroop B). No information about the word color congruency of the items	% of correct responses (%CR) and completion times (in sec) in each list	Experimental analyzing pre- and post-test Stroop performance as a function of acute aerobic exercise vs control group	Post-test improvement in %CR on Stroop A list in exercise group ; Post-test improvement in %CR on Stroop B list in control group	Acute physical exercise tends improve inhibitory control in physically inactive participants,	No comparison to a non-color words neutral baseline (Issue 1) Possible contingency issue (Issue 2) Performance of the control group suggests a practice effect (Issue 6) No direct indicator of

Running head: Examining attentional boost in the Stroop task

<p>Barella et al. (2010) – Journal of Aging and Physical Activity</p>	<p>40 participants (M_{age}=69.5) assigned to an Aerobic exercise (N=20) vs Control (N=20) group</p>	<p>20 min of walking on a treadmill (Acute aerobic exercise group) vs sitting quietly and having a general conversation with the investigator during 25 min (passive control group)</p>	<p>Computerized; Vocal (RTs were recorded with a sound detection program); 4 (red, green, blue, yellow)</p>	<p>1 block consisting of 20 color-neutral items (colored Os) and 20 color-words items printed in incongruent color, presented randomly 1 “negative priming” block consisting of 20 color-words items printed in incongruent color (the ink color of each word was the same as the color word stimulus on the previously presented item) (60 items in total)</p>	<p>Mean RTs per condition and per block</p>	<p>Experimental analyzing pre- vs immediately after vs 12 times post-intervention* Stroop performance as a function of acute aerobic exercise vs. control activity *the time intervals were: immediately after and 5, 10, 15, 20, 30, 45, 60, 75, 90, 105, and 120 min post-exercise</p>	<p>Immediate effects: Reduced RTs to the neutral, incongruent, and negative priming items immediately after exercise in comparison to control activity Delayed effects: Reduced RTs to the neutral, incongruent, and negative priming items after both exercise and control activity</p>	<p>Acute exercise has short-term benefits for speed of processing but does not affect other types of cognitive functioning such as executive function and inhibition in healthy older adults; the absence of difference between the exercise and control conditions in the delayed effects suggests a practice effect</p>	<p>inhibition (Issue 7) No comparison to a non-color words neutral baseline (Issue 1) No analysis of errors (possibly Issue 4) Lack of control for changes in response speed (possibly Issue 5) Performance of the control group in the delayed effects suggests a practice effect (Issue 6) No direct indicator of inhibition (Issue 7)</p>
<p>Chang et al. (2017) – Psychophysiology</p>	<p>30 participants (M_{age}=22.7): all performed aerobic exercise and control sessions</p>	<p>20-min steady-state exercise performed at 60–70% of HRR (acute aerobic exercise) vs reading physical activity related books for approximately 30 min (passive control session)</p>	<p>Computerized; Manual (color-coded buttons on a response pane); 3 (red, blue, green)</p>	<p>6 color-naming blocks each consisting of 38 color-words items printed in congruent color and 22 color-words items printed in incongruent color (360 trials in total)</p>	<p>Mean RTs and accuracy for each condition</p>	<p>Experimental comparing Stroop performance collected after aerobic exercise AND control activity (administered in counter-balanced order 72 hours apart)</p>	<p>Reduced completion times for congruent and incongruent items after the aerobic exercise activity compared to control No change in accuracy</p>	<p>Moderate intensity acute exercise results in a generalized rather than selective (i.e., executive) improvement in cognition</p>	<p>No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) No direct indicator of</p>

Running head: Examining attentional boost in the Stroop task

									executive function due to the lack of relevant comparisons (Issue 7)
Chang and Etnier (2009a) – Psychology of Sport and Exercise	41 participants (M _{age} =49.1 assigned to a Resistance exercise (N=22) vs Control (N=19) group)	2 sets of 10 repetitions for 6 exercises (acute resistance exercise , 45 min) vs reading a pamphlet related to resistance exercise (passive control) session	Not mentioned; Vocal; Not mentioned	3 lists (no information about the contingencies): 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored rectangles)	Completion times for each condition	Experimental analyzing pre- and post-test Stroop performance as a function of acute resistance exercise vs. control group	Reduced completion times for the word-reading and neutral items lists after the resistance exercise activity compared to control	Acute resistance exercise benefits speed of processing but not inhibition	No comparison to a non-color words neutral baseline (Issue 1) No analysis of errors (possibly Issue 4) Lack of control for changes in response speed (possibly Issue 5)
Chang and Etnier (2009b) – Journal of Sport and Exercise Psychology	65 participants (M _{age} =25.9) assigned to a 40% intensity (N=16), 70% intensity (N=16), 100% intensity (N=17) vs Control (N=16) group	2 sets of 10 repetitions for each of six muscle groups (Acute resistance exercise , 3 separate experimental groups, 40%, 70%, and 100%) vs watching a video on resistance exercise training (passive control group)	Paper and pencil/list presentation; Vocal – In case of mistake, the examiner said “incorrect”, pointed to the stimulus that resulted in the mistake and the participant was instructed correct his/her mistake and then to continue; Not mentioned	3 lists (50 trials each): 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored rectangles)	Completion times for each condition	Experimental analyzing pre- and post-test Stroop performance as a function of acute resistance exercise vs. control group	Reduced completion times for the neutral items lists after all exercise conditions compared to control and reduced completion times for the incongruent items list after the moderate (70%) intensity exercise	There is a dose-response relationship between the intensity of resistance exercise and cognitive performance: high-intensity exercise benefits speed of processing, but moderate intensity exercise is most beneficial for executive function (as assessed by the incongruent condition)	No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) No analysis of errors (possibly Issue 4) No direct indicator of executive function due to the lack of relevant

Running head: Examining attentional boost in the Stroop task

									comparisons (Issue 7)
Chang et al. (2014) – Journal of Science and Medicine in Sport	30 participants (M _{age} =58.1): All performed Resistance exercise and Control sessions	2 sets of 7 resistance exercises (Acute resistance exercise) vs reading materials related to physical activity and mental health (passive control) session	Paper and pencil/list presentation; Vocal; Not mentioned	5 lists each consisting of 50 trials: 1 word-reading list with color-words items printed in black; 4 color-naming lists 1 with color-words items printed in congruent colors, 1 with color-words items printed in incongruent colors, 1 with color-neutral word items printed in color and 1 with color-neutral items (colored squares)	Completion times for each condition	Experimental comparing Stroop performance collected after resistance exercise AND control activity (administered in counter-balanced order 48 hours apart)	Reduced completion times for all Stroop lists after the exercise session compared to the control session; largest difference between pre-test RT and post-test RT as a function of exercise for the incongruent condition	Acute resistance exercise facilitates general cognition but has a more beneficial effect on processes involving executive control	No computation of the Stroop interference despite the use of a color-neutral words condition No control of the color-congruent items contingency (Issue 2) No analysis of errors (possibly Issue 4) No direct indicator of executive function due to the lack of relevant comparisons (Issue 7)
Johnson et al. (2016) – Journal of Aging and Physical Activity	31 participants (M _{age} =71.7) assigned to the Aerobic group (N=17) vs Resistance group (N=14)	Intense cycling session on an ergometer (acute aerobic exercise) vs 5 strength exercises on machines (acute resistance exercise): No control group Each group performed 10- min and 30-min duration sessions one week apart	Computerized; Vocal (RTs were recorded with a sound detection program); 4 (red, blue, green, yellow)	3 blocks each consisting of 20 trials: 1 with color-words items printed in incongruent colors, 1 with color-neutral items (colored Xs), and 1 with two successive similar color-word items printed in incongruent colors (negative priming list)	Mean RTs per condition and per block	Experimental comparing Stroop performance collected immediately, 30 minutes, and 60 minutes after 10min duration AND 30min duration acute aerobic exercise vs 10min duration AND 30min duration resistance exercise	Reduced RTs to the negative priming condition immediately post-exercise, independently of the group (aerobic or resistance) or exercise bout duration At 60 min post-exercise, reduced RTs to the neutral items after the 10 min exercise bout	The improved performance in the negative priming condition suggests that an acute bout of exercise can improve cognitive performance, and in particular, the more complex executive functioning (i.e., inhibition) in older adults	No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) No analysis of errors (possibly Issue 4)

Running head: Examining attentional boost in the Stroop task

									<p>Possible practice effect (Issue 6)</p> <p>Change in negative priming difficult to interpret in the absence of change in the incongruent items</p> <p>Low N participants</p>
<p>Sibley et al. (2006) – Journal of Sport and Exercise Psychology</p>	<p>79 participants (M_{age}=22.5)</p>	<p>20 min of self-paced moderate-intensity exercise on a treadmill (Acute aerobic exercise) vs sitting quietly and reading material (passive control activity)</p>	<p>Computerized (Stimuli were presented in list format); Vocal; 4 (red, blue, green, yellow)</p>	<p>3 color-naming lists each consisting of 30 items: 1 with color-neutral items (colored Xs), 1 with color-words items printed in incongruent colors, and 1 with two successive similar color-word items printed in incongruent colors (negative priming list)</p>	<p>Completion times (in sec) in each list, errors (not reported because of high accuracy)</p>	<p>Experimental comparing Stroop performance collected after acute aerobic exercise AND after control activity (administered in counter-balanced order 7 days apart)</p>	<p>Reduced completion times in the incongruent list after exercise. Significant order*session interaction for all Stroop conditions showing that participants had faster responses in the second session; No significant change in the two other lists</p>	<p>The session effect suggests a learning or practice effect. Acute exercise leads to an improvement in executive functioning related to maintenance of goal-oriented processing, but not in processing speed and inhibition</p>	<p>No comparison to a non-color words neutral baseline (Issue 1)</p> <p>Practice effect (Issue 6)</p> <p>No analyses of pre-session RTs</p>
<p>Sugimoto et al. (2020) – Physiological Reports</p>	<p>20 participants (M_{age}=21); all performed Repeated exercise bout and Single exercise bout sessions</p>	<p>2 20-min bouts, separated by a 20-min sitting rest (Repeated Exercise, R-EX)) vs 1 40-min bout (Single Exercise bout, S-EX) (Acute moderate intensity aerobic exercise); no control group</p>	<p>Computerized, item-by-item presentation; Manual (color-labeled keypad); 4 (red, blue, green, yellow)</p>	<p>3 lists each consisting of 24 trials: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in congruent colors and 1 with color-words items printed in incongruent colors</p>	<p>Reaction times and response accuracy; Calculation of the “reverse-Stroop interference” score: (RTincongruent – RTword reading) / RTword reading x 100</p>	<p>Experimental comparing Stroop performance collected before, immediately after, 10 min after vs 30 min after the Repeated AND Single exercise (administered in counter-balanced order 72 hours apart)</p>	<p>Reduced RTs to the word-reading and congruent list immediately post exercise; this reduction remained after until 10min post-exercise</p> <p>Reduction of the reverse-Stroop interference score, immediately post exercise</p>	<p>Both repeated exercise and single exercise bouts have beneficial immediate effects on inhibitory control</p>	<p>Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1)</p> <p>No control of the color-congruent items contingency (Issue 2)</p>

Running head: Examining attentional boost in the Stroop task

									Possible practice effect (Issue 6) No direct indicator of inhibition (Issue 7)
Tomoo et al. (2020) – Physiological Reports	22 male participants (M _{age} =21.2); all performed the High-Intensity resistance exercise condition and the High-Volume low-intensity resistance exercise	High-Intensity resistance exercise condition (HRE) (70% of 1-RM) vs High-Volume low-intensity resistance exercise condition (HV-LRE) (35 % of 1-RM) (Acute resistance exercise); no control group	Computerized, item-by-item presentation; Not mentioned; 4 (red, blue, green, yellow)	3 lists each consisting of 24 trials: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in congruent colors and 1 with color-words items printed in incongruent colors	Reaction time and response accuracy; Calculation of the “reverse-Stroop interference” score: (RTincongruent – RTword reading) / RTword readingl x 100	Experimental comparing Stroop performance collected before, immediately after, 10 min, 20 min after, after vs 30 min after the High Intensity AND Low intensity exercise (administered in counter-balanced order 72 hours apart)	Reduced RTs to all lists immediately post exercise; this reduction remained after until 10min post-exercise for the word-reading and incongruent list and 20min for the congruent list Reduction of the reverse-Stroop interference score; this reduction remained after until 30 min post-exercise	The two resistance training interventions have beneficial immediate effects on inhibitory control	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) Possible practice effect (Issue 6) No direct indicator of inhibition (Issue 7)
Zimmer et al. (2016) – European Journal of Applied Physiology	121 participants (M _{age} =23.8) assigned to Low intensity (N=30), Moderate intensity (N=30), High intensity (N=31) vs control group (N=30)	Low vs moderate vs high intensity exercise bouts (acute aerobic exercise) vs supervised low-intensity myofascial relaxation training (active control group)	Computerized item-by-item presentation; Manual (keyboard with four colored buttons); 4 (red, green, blue, yellow)	4 blocks each consisting of 128 trials per condition: 1 word-reading block consisting of color-words presented in colored ink; 2 color-naming blocks: 1 consisting of colored rectangles and 1 with color-words items printed in incongruent colors; 1 reverse condition	Mean RTs (in ms); medians of RTs at word-reading baseline and naming baseline (colored rectangles). A software computes the Stroop effect by subtracting RTnaming from median RT of each Stroop	Experimental analyzing pre- vs post-intervention Stroop performance as a function of acute aerobic exercise (3 intensities) vs control Additional analyses of serum 5-HT concentrations - not discussed here	Reduced RTs to the word-reading condition after high-intensity exercise in comparison to low-intensity exercise A non-significant tendency for a RT reduction to the “Reverse Stroop effect” after Low Intensity in comparison to the oderate intensity exercise	Exercise intensity (and subsequent enhanced serum 5-HT levels) leads to improved response inhibition as reflected by the changes in the reading and the reverse Stroop effect	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2)

Running head: Examining attentional boost in the Stroop task

				(word-reading incongruent items).	condition; Computation of the “ Reverse Stroop effect ” as the balance between median RT at Stroop condition and RT at reading condition				No analysis of errors (possibly Issue 4) Possible practice effect (Issue 6) No direct indicator of inhibition (Issue 7)
Kim (2015) – Journal of Exercise Nutrition & Biochemistry	14 participants (M _{age} =20.6) assigned to the Exercise group (N=7) vs the Control group (N=7)	8-week Taekwondo training program (85-minute sessions, 5 sessions per week); No information about the control group intervention	Paper and pencil/list presentation; Vocal; 3 (red, green, blue)	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Number of correct responses in 45 seconds for each condition	Experimental analyzing pre- and post-test Stroop performance as a function of training program vs. control group	Post-test improvement in number of correctly identified items in the word-reading and neutral lists No significant change in the incongruent condition	The exercise training intervention improved cognitive functions	No comparison to a non-color words neutral baseline (Issue 1) Lack of control for changes in response speed (possibly Issue 5) Low N participants
Ludyga et al. (2019) – Journal of Science and Medicine in Sport	35 participants (Age=12-15) assigned to Physical training group (N=19) vs Control group (N=16)	8-week physical training program (20-min exercise sessions, moderate intensity aerobic and coordinative exercise), 5 days a week vs social interaction with classmates during the intervention groups’ exercise sessions (passive control group)	Computerized; Manual (button on a response box); Not mentioned	4 blocks each consisting of 36 trials including 18 color-words items printed in congruent color and 18 color-words items printed in incongruent color presented randomly (144 trials in total)	Accuracy rates and RTs (in ms) in the incongruent condition	Experimental analyzing pre- vs. post-intervention Stroop performance as function of physical training vs passive control	Reduced post-test RTs to incongruent items in the exercise group compared to the control group. No change in accuracy	The combined aerobic and coordinative exercise program led to improvements in inhibitory control	No comparison to a non-color words neutral baseline (Issue 1) No control of the color-congruent items contingency (Issue 2) Improvement consistent with a normal developmental trajectory;

Running head: Examining attentional boost in the Stroop task

									Lack of control for changes in response speed (possibly Issue 5)
									No direct indicator of inhibition (Issue 7)
									Low N participants
Iuliano et al. (2015) – Experimental Gerontology	80 participants (M _{age} =67) assigned to the Resistance training group (N=20), Cardiovascular training group (N=20), Postural training group (N=20) vs Control group (N=20)	12-week training program (cardiovascular, resistance, postural), 3 sessions/week vs no intervention (passive control group)	Not mentioned; Vocal; 3 (red, green, blue)	3 lists: 1 word-reading list with color-words items printed in black (T1); 2 color-naming lists: 1 with color-neutral items (colored circles) (T2) and 1 with color-words items printed in incongruent colors (T3)	Number of correct responses in 45 seconds for each condition; the Stroop interference was calculated as: (T3 - (T1 + T2))/2; only the interference scores are reported	Experimental analyzing pre- and post-test Stroop performance as a function of cardiovascular vs resistance vs postural training programs vs control	No change in Stroop performance for either group	The interventions did not lead to significant improvements in executive functions (and notably selective attention, visual attention, and inhibitory control as measured by the Stoop task)	Word-reading condition as baseline; No comparison to a non-color words neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)
Maillot et al. (2012) – Psychology and Aging	30 participants (M _{age} =73.5) assigned to Intervention group (N=15) vs Control group (N=16)	12-week active video game (exergames) training program (one-hour sessions, twice a week) vs no intervention (passive control group)	Not mentioned; Vocal; Not mentioned	4 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 consisting of colored Xs; 1 switching list including word-reading and color-naming incongruent items in function of a cue	Number of correct responses in 45 seconds to the incongruent and flexibility conditions; A “change score” was calculated by subtracting the pretest score from the posttest score for the incongruent and flexibility conditions	Experimental analyzing pre- vs. post-intervention (6 months) Stroop performance as function of physical training vs passive control	Greater change score in both the incongruent and flexibility lists after the training program than control for both	Exergames training is beneficial to executive functions	No comparison to a non-color words neutral baseline (Issue 1) No direct indicator of executive function due to the lack of relevant comparisons (Issue 7) Low N participants
Predovan et al. (2012) – Journal	67 participants (M _{age} =67.9) assigned to the	12-week aerobic training program (60-min sessions,	Paper and pencil/list presentation;	4 lists each consisting 100 trials: 1 word-	Mean RTs and accuracy for each condition;	Experimental* analyzing pre- vs. post-	Reduced post-test RTs and greater change score in the	The benefits of an aerobic training program occur	No comparison to a non-color

Running head: Examining attentional boost in the Stroop task

of Aging Research	Intervention group (N=32) vs Control group (N=35)	three sessions per week) vs no training intervention (passive control group)	Vocal; 5 (black, red, blue, green, yellow)	reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored rectangles); 1 switching list with color-words items printed in incongruent color, requiring either color or word meaning identification (cued by a square)	For each Stroop condition, a score measuring the amount of change between the pre- and post-test was computed as follows: pre-test minus post-test score	intervention Stroop performance as function of physical training vs passive control * determination of groups based on participants' order of recruitment and their willingness to engage in training program	switching list after exercise in comparison to control	primarily on executive functions (inhibition and shifting). The absence of change in the incongruent condition suggests that the improvement occurs in the switching domain	words neutral baseline (Issue 1) Lack of control for changes in response speed (possibly Issue 5) No direct indicator of inhibition (Issue 7)
Castelli et al. (2011) – Preventive Medicine	59 participants (M _{age} =8.8)	9-month 120-minute daily school-based physical training program (Participants wore a heart rate monitor during the physical exercise session); no control group	Not mentioned; Vocal; Not mentioned	3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored Xs)	Number of correct responses in 45 seconds for each condition; Computation of an interference score with a software (Not reported in the results; no details about its calculation)	Experimental analyzing pre- and post-test Stroop performance as a function of physical training	Post-test improvement in correct responses in all Stroop conditions. Positive associations between the percentage of time spent above the target heart zone during the training program and performance in the incongruent condition	The time spent at high heart rate during physical activity sessions (suggesting vigorous physical activity) is beneficial to executive functions	No comparison to a non-color words neutral baseline (Issue 1) Improvement consistent with a normal developmental trajectory; Lack of control for changes in response speed (possibly Issue 5) No direct indicator of executive function due to the lack of relevant comparisons (Issue 7)

Running head: Examining attentional boost in the Stroop task

<p>Kvalø et al. (2017) – Scandinavian Journal of Medicine & Science in Sports</p>	<p>429 participants (Age=10-11) assigned to Physical training group (N=217) vs Control group (N=212)</p>	<p>10-month school-based physical training program (325 minutes of physical activities per week) vs regular mandatory national curriculum of 135 minutes of physical activity (active control group)</p>	<p>Not mentioned; Vocal; Not mentioned</p>	<p>3 lists: 1 word-reading list with color-words items printed in black; 2 color-naming lists: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (colored patches)</p>	<p>Number of CR completed in 45 sec for each list</p>	<p>Experimental analyzing pre- vs post-intervention Stroop performance as function of physical training vs active control</p>	<p>Increased number of CR to colored patches items list in both the physical training and control groups. No significant change in the other lists. Non-significant tendency for a time×group interaction in the total number of CR in the 3 cumulated lists</p>	<p>Increased physical activity in school resulted in a non-significant tendency to benefit children’s executive function</p>	<p>No comparison to a non-color words neutral baseline (Issue 1) Improvement consistent with a normal developmental trajectory; Lack of control for changes in response speed (possibly Issue 5) No direct indicator of executive function due to the lack of relevant comparisons (Issue 7)</p>
<p>Smiley-Oyen et al. (2008) – Annals of Behavioral Medicine</p>	<p>57 participants (Age=65-79) assigned to a Cardiovascular exercise training group (N=28) vs a Flexibility and Resistance training group (N=29)</p>	<p>10-month training program including cardiovascular exercise vs flexibility and resistance exercise (no passive control group)</p>	<p>Computerized; Manual (pressing 1, 2, 3 keys on a computer keyboard); 3 (red, blue, green)</p>	<p>3 blocks each consisting of 100 trials: 1 word-reading block consisting of color-words presented in colored ink; 2 color-naming blocks: 1 with color-words items printed in incongruent colors and 1 with color-neutral items (undefined)</p>	<p>Error rates and mean RTs for each condition</p>	<p>Experimental analyzing pre- vs midpoint vs post-intervention Stroop performance as function of cardiovascular training vs flexibility and resistance training</p>	<p>Reduced post-test RTs to incongruent items in the cardiovascular training group compared to the flexibility and resistance group. No significant change in the two other Stroop blocks Increased error rates in the incongruent items at post intervention in comparison to midpoint in the flexibility and resistance group.</p>	<p>Cardiovascular (aerobic) exercise in older adults has beneficial effect on executive control, especially inhibition</p>	<p>No comparison to a color-neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)</p>

Running head: Examining attentional boost in the Stroop task

							Reduced error rates in the incongruent items at mid-point and post intervention in comparison to pre-intervention in the cardiovascular training group		
Liu-Ambrose et al. (2008) – Journal of the American Geriatrics Society	52 participants (M _{age} =69.6) assigned to the Intervention group (N=28) vs the Control group (N=24)	12-month home-based balance and strength exercise training program vs administration of clinical care recommendations (passive control group)	Paper and pencil/list presentation; Vocal; Not mentioned	A single block of color-words items printed in incongruent colors (112 trials)	Number of correct responses and the total time to complete the task	Experimental analyzing pre- vs. midpoint intervention (6 months) Stroop performance as function of physical training vs passive control (no measure at post-test)	Increased number of correctly identified items at 6 months in comparison to pre-intervention after the training program vs passive control	The home-based resistance and balance training program improved executive functioning , specifically response inhibition , after 6 months	No comparison to a color-neutral baseline (Issue 1) No direct indicator of inhibition (Issue 7)

Note. Light gray background indicates an “overinterpretation” of reported results and dark grey indicates a clear mismatch between empirical results and their interpretation.

*Same study but two different experiments; the general design of the study is quasi-experimental: participants completed a physical activity questionnaire (IPAQ) to determine whether participants were physically active (study 1) or physically inactive (study 2).

References (Table 2S)

- Aadland, K. N., Moe, V. F., Aadland, E., Anderssen, S. A., Resaland, G. K., & Ommundsen, Y. (2017). Relationships between physical activity, sedentary time, aerobic fitness, motor skills and executive function and academic performance in children. *Mental Health and Physical Activity, 12*, 10–18. <https://doi.org/10.1016/j.mhpa.2017.01.001>
- Aguirre-Loaiza, H., Arenas, J., Arias, I., Franco-Jimenez, A., Barbosa-Granados, S., Ramos-Bermúdez, S., Ayala-Zuluaga, F., Núñez, C., & García-Mas, A. (2019). Effect of Acute Physical Exercise on Executive Functions and Emotional Recognition: Analysis of Moderate to High Intensity in Young Adults. *Frontiers in Psychology, 10*, 2774. <https://doi.org/10.3389/fpsyg.2019.02774>
- Aly, M., & Kojima, H. (2020). Relationship of regular physical activity with neuroelectric indices of interference processing in young adults. *Psychophysiology, 57*(12), e13674. <https://doi.org/10.1111/psyp.13674>
- Barella, L. A., Etnier, J. L., & Chang, Y.-K. (2010). The Immediate and Delayed Effects of an Acute Bout of Exercise on Cognitive Performance of Healthy Older Adults. *Journal of Aging and Physical Activity, 18*(1), 87–98. <https://doi.org/10.1123/japa.18.1.87>
- Bixby, W. R., Spalding, T. W., Haufler, A. J., Deeny, S. P., Mahlow, P. T., Zimmerman, J. B., & Hatfield, B. D. (2007). The Unique Relation of Physical Activity to Executive Function in Older Men and Women. *Medicine & Science in Sports & Exercise, 39*(8), 1408–1416. <https://doi.org/10.1249/mss.0b013e31806ad708>
- Boucard, G. K., Albinet, C. T., Bugaiska, A., Bouquet, C. A., Clarys, D., & Audiffren, M. (2012). Impact of Physical Activity on Executive Functions in Aging: A Selective Effect on Inhibition Among Old Adults. *Journal of Sport and Exercise Psychology, 34*(6), 808–827. <https://doi.org/10.1123/jsep.34.6.808>

- Castelli, D. M., Hillman, C. H., Hirsch, J., Hirsch, A., & Drollette, E. (2011). FIT Kids: Time in target heart zone and cognitive performance. *Preventive Medicine, 52*, S55–S59.
<https://doi.org/10.1016/j.ypmed.2011.01.019>
- Chang, E. C.-H., Chu, C.-H., Karageorghis, C. I., Wang, C.-C., Tsai, J. H.-C., Wang, Y.-S., & Chang, Y.-K. (2017). Relationship between mode of sport training and general cognitive performance. *Journal of Sport and Health Science, 6*(1), 89–95.
<https://doi.org/10.1016/j.jshs.2015.07.007>
- Chang, Y.-K., Alderman, B. L., Chu, C.-H., Wang, C.-C., Song, T.-F., & Chen, F.-T. (2017). Acute exercise has a general facilitative effect on cognitive function: A combined ERP temporal dynamics and BDNF study. *Psychophysiology, 54*(2), 289–300.
<https://doi.org/10.1111/psyp.12784>
- Chang, Y.-K., & Etnier, J. L. (2009a). Effects of an acute bout of localized resistance exercise on cognitive performance in middle-aged adults: A randomized controlled trial study. *Psychology of Sport and Exercise, 10*(1), 19–24.
<https://doi.org/10.1016/j.psychsport.2008.05.004>
- Chang, Y.-K., & Etnier, J. L. (2009b). Exploring the Dose-Response Relationship between Resistance Exercise Intensity and Cognitive Function. *Journal of Sport and Exercise Psychology, 31*(5), 640–656. <https://doi.org/10.1123/jsep.31.5.640>
- Chang, Y.-K., Tsai, C.-L., Huang, C.-C., Wang, C.-C., & Chu, I.-H. (2014). Effects of acute resistance exercise on cognition in late middle-aged adults: General or specific cognitive improvement? *Journal of Science and Medicine in Sport, 17*(1), 51–55.
<https://doi.org/10.1016/j.jsams.2013.02.007>
- Gajewski, P. D., & Falkenstein, M. (2015). Long-term habitual physical activity is associated with lower distractibility in a Stroop interference task in aging: Behavioral and ERP

evidence. *Brain and Cognition*, 98, 87–101.

<https://doi.org/10.1016/j.bandc.2015.06.004>

Gajewski, P. D., Falkenstein, M., Thönes, S., & Wascher, E. (2020). Stroop task performance across the lifespan: High cognitive reserve in older age is associated with enhanced proactive and reactive interference control. *NeuroImage*, 207, 116430.

<https://doi.org/10.1016/j.neuroimage.2019.116430>

Goenarjo, R., Bosquet, L., Berryman, N., Metier, V., Perrochon, A., Fraser, S. A., & Dupuy, O. (2020). Cerebral Oxygenation Reserve: The Relationship Between Physical Activity Level and the Cognitive Load During a Stroop Task in Healthy Young Males. *International Journal of Environmental Research and Public Health*, 17(4), Article 4.

<https://doi.org/10.3390/ijerph17041406>

Goenarjo, R., Dupuy, O., Fraser, S., Berryman, N., Perrochon, A., & Bosquet, L. (2021). Cardiorespiratory fitness and prefrontal cortex oxygenation during Stroop task in older males. *Physiology & Behavior*, 242, 113621.

<https://doi.org/10.1016/j.physbeh.2021.113621>

Iuliano, E., di Cagno, A., Aquino, G., Fiorilli, G., Mignogna, P., Calcagno, G., & Di Costanzo, A. (2015). Effects of different types of physical activity on the cognitive functions and attention in older people: A randomized controlled study. *Experimental Gerontology*, 70, 105–110. <https://doi.org/10.1016/j.exger.2015.07.008>

Johnson, L., Addamo, P. K., Selva Raj, I., Borkoles, E., Wyckelsma, V., Cyarto, E., & Polman, R. C. (2016). An Acute Bout of Exercise Improves the Cognitive Performance of Older Adults. *Journal of Aging and Physical Activity*, 24(4), 591–598. <https://doi.org/10.1123/japa.2015-0097>

- Kim, Y. (2015). The effect of regular Taekwondo exercise on Brain-derived neurotrophic factor and Stroop test in undergraduate student. *Journal of Exercise Nutrition & Biochemistry, 19*(2), 73–79. <https://doi.org/10.5717/jenb.2015.15060904>
- Kvalø, S. E., Bru, E., Brønnick, K., & Dyrstad, S. M. (2017). Does increased physical activity in school affect children’s executive function and aerobic fitness? *Scandinavian Journal of Medicine & Science in Sports, 27*(12), 1833–1841. <https://doi.org/10.1111/sms.12856>
- Liu-Ambrose, T., Donaldson, M. G., Ahamed, Y., Graf, P., Cook, W. L., Close, J., Lord, S. R., & Khan, K. M. (2008). Otago Home-Based Strength and Balance Retraining Improves Executive Functioning in Older Fallers: A Randomized Controlled Trial. *Journal of the American Geriatrics Society, 56*(10), 1821–1830. <https://doi.org/10.1111/j.1532-5415.2008.01931.x>
- Ludyga, S., Köchli, S., Pühse, U., Gerber, M., & Hanssen, H. (2019). Effects of a school-based physical activity program on retinal microcirculation and cognitive function in adolescents. *Journal of Science and Medicine in Sport, 22*(6), 672–676. <https://doi.org/10.1016/j.jsams.2018.11.029>
- Maillot, P., Perrot, A., & Hartley, A. (2012). Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. *Psychology and Aging, 27*(3), 589–600. <https://doi.org/10.1037/a0026268>
- O’Brien, M. W., Kimmerly, D. S., & Mekari, S. (2021). Greater habitual moderate-to-vigorous physical activity is associated with better executive function and higher prefrontal oxygenation in older adults. *GeroScience, 43*(6), 2707–2718. <https://doi.org/10.1007/s11357-021-00391-5>
- Prakash, R. S., Voss, M. W., Erickson, K. I., Lewis, J. M., Chaddock, L., Malkowski, E., Alves, H., Kim, J., Szabo, A., White, S. M., Wójcicki, T. R., Klamm, E. L., McAuley,

- E., & Kramer, A. F. (2011). Cardiorespiratory fitness and attentional control in the aging brain. *Frontiers in Human Neuroscience*, *4*, 229.
<https://doi.org/10.3389/fnhum.2010.00229>
- Predovan, D., Fraser, S. A., Renaud, M., & Bherer, L. (2012). The Effect of Three Months of Aerobic Training on Stroop Performance in Older Adults. *Journal of Aging Research*, *2012*, 1–7. <https://doi.org/10.1155/2012/269815>
- Ramnath, U., Rauch, L., Lambert, E. V., & Kolbe-Alexander, T. L. (2018). The relationship between functional status, physical fitness and cognitive performance in physically active older adults: A pilot study. *PLOS ONE*, *13*(4), e0194918.
<https://doi.org/10.1371/journal.pone.0194918>
- Salas-Gomez, D., Fernandez-Gorgojo, M., Pozueta, A., Diaz-Ceballos, I., Lamarain, M., Perez, C., Kazimierczak, M., & Sanchez-Juan, P. (2020). Physical Activity Is Associated With Better Executive Function in University Students. *Frontiers in Human Neuroscience*, *14*, 11. <https://doi.org/10.3389/fnhum.2020.00011>
- Sibley, B. A., Etnier, J. L., & Le Masurier, G. C. (2006). Effects of an Acute Bout of Exercise on Cognitive Aspects of Stroop Performance. *Journal of Sport and Exercise Psychology*, *28*(3), 285–299. <https://doi.org/10.1123/jsep.28.3.285>
- Smiley-Oyen, A. L., Lowry, K. A., Francois, S. J., Kohut, M. L., & Ekkekakis, P. (2008). Exercise, Fitness, and Neurocognitive Function in Older Adults: The “Selective Improvement” and “Cardiovascular Fitness” Hypotheses. *Annals of Behavioral Medicine*, *36*(3), 280–291. <https://doi.org/10.1007/s12160-008-9064-5>
- Sugimoto, T., Suga, T., Tsukamoto, H., Tomoo, K., Dora, K., Hashimoto, T., & Isaka, T. (2020). Effect of repeated bouts versus a single bout of moderate-intensity exercise on postexercise inhibitory control. *Physiological Reports*, *8*(15), e14528.
<https://doi.org/10.14814/phy2.14528>

Tomoo, K., Suga, T., Sugimoto, T., Tanaka, D., Shimoho, K., Dora, K., Mok, E., Matsumoto, S., Tsukamoto, H., Takada, S., Hashimoto, T., & Isaka, T. (2020). Work volume is an important variable in determining the degree of inhibitory control improvements following resistance exercise. *Physiological Reports*, 8(15), e14527.
<https://doi.org/10.14814/phy2.14527>

Zimmer, P., Stritt, C., Bloch, W., Schmidt, F.-P., Hübner, S. T., Binneböbel, S., Schenk, A., & Oberste, M. (2016). The effects of different aerobic exercise intensities on serum serotonin concentrations and their association with Stroop task performance: A randomized controlled trial. *European Journal of Applied Physiology*, 116(10), 2025–2034. <https://doi.org/10.1007/s00421-016-3456-1>