

Research Article

The Effect of Mental Fatigue and Gender on Working Memory Performance during Repeated Practice by Young and Older Adults

Valentina Pergher ^{1,2}, Nele Vanbilsen ², and Marc Van Hulle ²

¹Department of Cognitive Neuropsychology, Harvard University, Cambridge, MA, USA

²Department of Neurosciences, Laboratory for Neuro- & Psychophysiology, KU Leuven-University of Leuven, Belgium

Correspondence should be addressed to Valentina Pergher; valentinapergher@fas.harvard.edu

Received 8 December 2020; Revised 9 August 2021; Accepted 3 September 2021; Published 4 October 2021

Academic Editor: Nicoletta Berardi

Copyright © 2021 Valentina Pergher et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Working memory (WM) is one of the most investigated cognitive functions albeit the extent to which individual characteristics impact on performance is still unclear, especially when older adults are involved. The present study considers repeated practice of a visual *N*-Back task with three difficulty levels (1-, 2-, and 3-Back) in healthy young and older individuals. Our results reveal that, for both age groups, the expected mental fatigue was countered by a learning effect, in terms of accuracies and reaction times, which turned out to benefit females more than males, for all three *N*-Back levels. We conclude that future WM studies, in particular when relying on repeated *N*-Back sessions, should account for learning effects in relation to mental fatigue and gender, in both young and older adults.

1. Introduction

Working memory (WM), one of the most investigated cognitive functions, refers to the ability to temporarily store and manipulate information necessary to complete complex tasks for a short period of time [1, 2]. Literature suggests that WM is the outcome of a series of processes that manage two domain-specific slave systems [3]. One subdomain is the phonological store that elaborates verbal information; the other is the visual cache that elaborates visuospatial information [3].

Previous studies have indicated that individual differences, such as age, gender, and education, may have a substantial impact on several cognitive domains [4]. In the results after a single session of cognitive performance, the impact of age was observed by lower accuracies and increased reaction times for older individuals compared to younger ones [5–7]. Some studies found lower accuracies based on task demand [8], while in the study of Hale et al. [9], age-related performance decline was independent from task demand. In general, it has been shown that WM perfor-

mance becomes worse with aging, indicating an increased reaction time due to a general slowdown in information processing [10, 11] or due to a higher difficulty in suppressing interference from irrelevant information [12]. Moreover, education level has been shown to provide a protective effect on age-related changes during a verbal learning task [13–15]. Meijer et al. [13] showed that only participants with a high educational level were able to inhibit irrelevant verbal information. Additionally, a higher education level indicated to play a significant positive role when individuals aged [16].

Albeit there is a broad consensus on the impact of age and education level on cognitive functioning, this is less in the case with gender. For instance, there are stark discrepancies between studies as several WM tasks failed to reveal any gender differences [17], whereas others observed significant age-related changes in task performance [18] and neuro-functional responses [19]. Several studies that focused on adolescents [20], young adults [21, 22], and older adults [23] did not observe any significant gender differences. In contrast, other studies demonstrated a gender difference for WM tasks by showing that women perform better than

men, especially in spatial-based tasks [24, 25], although the picture could change when considering higher WM loads [26]. Furthermore, a recent meta-analysis that investigated gender differences in WM showed a female advantage for object location memory and a male advantage for *N*-Back tasks [27]. Several studies reported a superior female performance for object-location memory tasks [28–30] and for verbal tasks [31, 32]. However, not all high-load WM studies reported a gender-related difference. For instance, Goldstein et al. [33] observed no effects for gender when using a task similar to the letter *N*-Back task in a behavioral fMRI study, although female subjects had a higher neural activation in the prefrontal cortex. Despite the evidence, gender has been much less researched in the context of the visual (object) *N*-Back task, especially in older individuals, and the few studies that have been conducted showed mixed outcomes. The study of Speck et al. [34] showed a female advantage for a verbal *N*-Back task that uses sequences of displayed words and numbers. Using a digit *N*-Back task, Pliatsikas et al. [18] showed that, with increasing age, males exhibited a steeper WM decline than females and that, with increasing education levels, females reached greater WM gains than males. However, the study of Lejbak et al. [28] revealed a male advantage when using an object-location memory task (*N*-Back), and Schmidt et al. [17], using a verbal *N*-Back task in young adults, demonstrated that WM is not affected by gender at behavioral and neural levels.

Considering that males and females have different pathways to encode memories, solve problems, and make decisions, these gender-related functional differences may be associated with gender-specific structures of the brain [35]. Xin et al. [36] showed, using a 3D “Pure” Convolutional Neural Network (PCNN), brain structure differences between males and females. Specifically, they found 25 brain features that were significantly different between men and women. These structures included the left precuneus, left postcentral gyrus, left cingulate gyrus, right orbital gyrus of the frontal lobe, left occipital thalamus in the gray matter volume, middle cerebellum peduncle, genu of corpus callosum, right anterior corona radiata, right superior corona radiata, and left anterior limb of the internal capsule. Therefore, Xin et al. [36] confirm that gender-related differences in the brain structure exist and that they might explain gender-related differences in cognition. Furthermore, another potential source of gender differences in WM task performance is the gender-specific nature of the adopted strategy as shown by Alarcón et al. [20]. In this study, using a spatial WM task in adolescents, no gender differences in performance were found, although significant differences in patterns of neural activity measured by fMRI were detected. Females showed a reduced activation of the default mode network, while males an activation in regions associated with spatial WM function, suggesting that there might be multiple networks associated to WM and that changes in information-type such as spatial WM tasks might reveal the adoption of gender-specific behavioral strategies [37]. Furthermore, females showed an advantage relative to males on tests that measure recollection, processing speed, object location memory, verbal memory, verbal learning, and pre-

cision tasks [38, 39]. In contrast, other studies showed a male advantage in tests that require mental rotation or manipulation of an object [40] and in reaction time and finger tapping [38]. Moreover, Upadhayay and Guragain [41] investigated the effect of progesterone on cognitive functions when comparing several cognitive task performances between males (one time) and females (two times: during preovulatory and postovulatory phases of the menstrual cycle). While studies have shown that estrogen and testosterone accentuate cognitive functions, the role of progesterone on cognitive functions is still controversial. The results indicated that male cognitive functions were comparable to female preovulatory phase cognitive functions. However, females, during the postovulatory phase of their cycle, may have advantages in executive tasks (Stroop test), but disadvantages in attentional tasks.

One of the most common WM paradigms used to study complex cognitive functions is *N*-Back, a task that requires deciding whether the current stimulus matches the one presenting *N* stimuli before. It calls upon different mental processes [42], in particular WM updating that involves the encoding of incoming stimuli, monitoring, maintenance and updating of the sequence, and matching the current stimulus to the one presenting *N* positions back. Furthermore, the *N*-Back task also reflects other high-order cognitive functions, such as inhibitory control, cognitive flexibility, problem solving, and decision making [2]. Recent functional imaging studies focused on the *N*-Back task and investigated structural correlates for different difficulty levels. Indeed, the *N*-Back task requires a continuous variable load and, consequently, different cognitive efforts [43]. For instance, using a visual *N*-Back task, Pesonen et al. [44] observed differences in alpha band power when participants were performing a higher *N*-back level (2- and 3-Back) compared to a lower one (1-Back), confirming that information maintenance and manipulation load change with increasing *N* [45]. Moreover, it has been shown that with age, cognitive performance becomes slower and less accurate, and performance accuracy in more complex WM tasks decreases [46]. Another factor of interest, albeit only rarely investigated, is mental fatigue resulting from prolonged periods of cognitive activity. It is associated with tiredness or exhaustion resulting in a decrease in task performance and commitment [47]. Käthner et al. [48] and Kohl-morgen et al. [49] highlighted that, along with a lower performance, alpha band power increased with mental fatigue in older individuals, respectively, during dichotic listening tasks and in real operating environments for the last run compared to the first run. Furthermore, also, the study of Pergher et al. [50] attributed an important role to mental fatigue, comparing the first and last run of an object *N*-Back task, and revealed significant differences in alpha band power of older adults.

In the present study, we investigated accuracy and reaction time during repeated sessions (lasting less than one hour) of an *N*-Back task with three difficulty levels by healthy young and older adults. Both age groups performed the 1-, 2-, and 3-Back levels, for a total of four rounds each. We hypothesize that the last part of the experiment (last

round) compared to the first part (first round) has a higher accuracy level due to a learning effect. Indeed, it has been shown that repetition of visual stimuli in certain cognitive tasks can promote improvements in implicit memory and attention [51]. However, based on the studies of Käthner et al. [48] and Kohlmorgen et al. [49], demonstrating the effect of mental fatigue, especially in older individuals, we investigated the effect of the latter performing our analyses on the first (1st) and last (4th) rounds of the N -Back task. We expected that the effect of mental fatigue was stronger compared to the learning effect, especially in older adults, showing a lower accuracy at the end of the 1-hour N -Back task performance. Besides the possible impact of fatigue, we are aware that learning could also interfere with task performance. Moreover, taking into consideration the controversial results of gender on the WM domain [23, 24], especially in the elderly population, we also explored the impact of gender on an object N -Back task. Based on previous literature on N -Back studies [18], we hypothesize an advantage for females compared to males.

2. Materials and Methods

2.1. Participants. We recruited 23 young healthy subjects (14 females, mean age 24.78, age range 19–34 years) via flyers and advertisements, all undergraduate or graduate students from KU Leuven University, and 16 older healthy subjects (8 females, mean age 62.12 years, range 52–69 years). Demographics are shown in Table 1.

Participants were included in the study only if they met the following inclusion criteria: no history of psychiatric or neurological disease, normal or corrected vision, and not on any medication that could interfere with cognitive functioning. Participants were divided according to age and gender as follows: female ($N = 14$, education = $14.86 \text{ y} \pm 2.88$) and male ($N = 9$, education = $12.88 \text{ y} \pm 2.52$) for young adults; female ($N = 8$, education = $7.25 \text{ y} \pm 1.77$) and male ($N = 8$, education = $7.62 \text{ y} \pm 2.13$) for older adults. All participants, except four, were part of previous studies on the effect of WM training on behavioral and EEG responses ([52] for older adults; [53] for young adults). Note that due to some technical issues when recording behavioral responses of 4 young participants, we retained 19 subjects in total for reaction time (RT) outcomes of which 10 females and 9 males.

2.2. N -Back Task. We used an object N -Back task as in our previous studies [50, 52]. The task requires to decide whether the current stimulus matches the one presenting N items before (Figure 1). We divided the task into three difficulty levels, 1-Back, 2-Back, and 3-Back, respectively. A given N -Back level consisted of four blocks (or rounds), and each block contains 100 stimuli. For each block of 100 stimuli, 33% of them were targets. The stimuli were presented for 1 s followed by a 1.5 s interstimulus interval (ISI) and 0.5 s of feedback presentation (red frowny/green smiley). Between blocks, the subject could take a break and continue when pressing the space bar. Subjects performed the N -Back task within 60 minutes.

TABLE 1: Participant demographics showing gender division, mean age, and years of education (standard deviation between brackets).

	Young adults ($N = 23$)	Older adults ($N = 16$)
Age	24.78 (3.99)	62.12 (4.69)
Gender	9 M/14 F	8 M/8 F
Education (in years)	14.3 (2.94)	7.5 (1.9)

2.3. Procedure. Prior to the experiment, participants were informed about the experimental procedure and signed an informed consent form. Participants filled out a brief demographic questionnaire. All participants were invited to the laboratory for a single session of the N -Back task with all three task difficulty levels, as described above, repeated four times (further called “rounds”). The study was approved by our university hospital’s ethical committee (UZ KU Leuven, S59475). Participants received a payment of 20 euros.

2.4. Statistical Analyses. All analyses were performed using Matlab (version 2019b). The N -Back task results were analyzed using a two-way ANOVA with accuracy or reaction time (RT) as dependent variable and N -Back level and gender as independent variables. The Shapiro-Wilk test [54] was applied to assess whether normality was violated, considering our relatively small sample size. If normality was violated, a more conservative p value of 0.01 was adopted. Post hoc analyses were calculated using independent samples t -tests with N -Back level or gender as independent variable and accuracy or RT as dependent variable, and correction for multiple comparison (Bonferroni) was used. Furthermore, we separated the first and last round N -Back performance, considering as first round the first 100 stimuli and as last round the last 100 stimuli of the same N -Back level. In order to address the possible effect of imbalanced gender, we ran a bootstrapping procedure (using bootstrap sampling, accessible via Matlab Central), for which we created two equally sized groups ($N = 8/N = 8$) of male and female accuracies (in % correct responses) by sampling the corresponding (with replacement) original $N = 14/N = 9$ male and female accuracies, and determined the p value of the difference in accuracy and its confidence interval (CI). We repeated this procedure 10000 times and used the obtained difference distribution to estimate the p value of the actual difference in mean.

Finally, $\eta^2 = 0.098$ [55] were reported ($\eta^2 = \text{SS}_{\text{effect}} / (\text{SS}_{\text{effect}} + \text{SS}_{\text{error}})$, where SS is sum of squares) to indicate the magnitude of the significant differences for behavioral outcomes.

3. Results

3.1. Mental Fatigue in Reaction Time and Accuracy. We compared reaction times (RTs) of the first and last rounds of the N -Back task to detect a possible effect of mental fatigue (Figure 2). We did not find any significant difference in RT when comparing the first and last rounds in younger adults. When looking at older adults, our results showed a significant main effect for N -Back level ($F(2, 90) = 4.89$, p

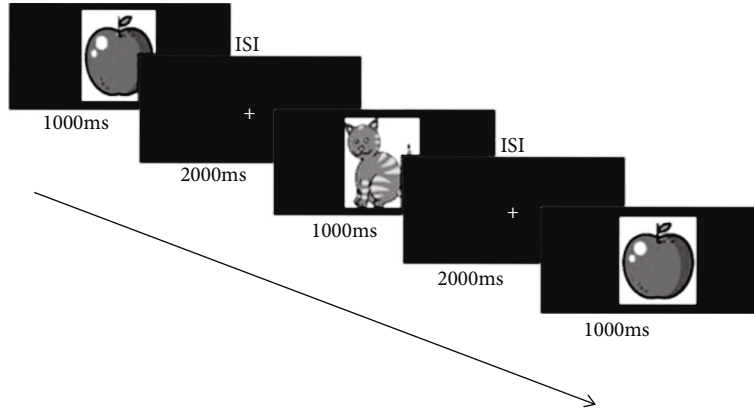


FIGURE 1: Example stimulus sequence during 2-Back task with stimulus durations of 1000 ms and an interstimulus interval (ISI) of 1500 ms.

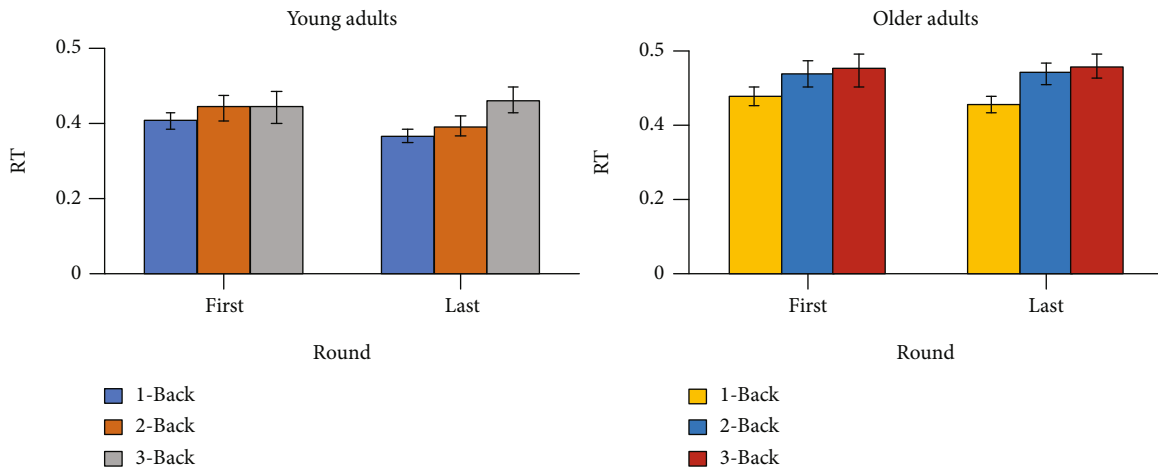


FIGURE 2: Means \pm SD for reaction time (RT) for young (a) and older (b) adults comparing first and last round during N -Back task performance for the three different N -Back levels.

= 0.010, $\eta p^2 = 0.098$), indicating that, overall, older adults exhibited lower RTs for 1-Back compared to 2- and 3-Back. However, no significant differences across levels were found when comparing the first and last rounds.

Similarly, we compared the accuracy of the first and last rounds of the N -Back task (Figure 3). Our results showed a significant difference for N -Back level in younger adults ($F(2,126) = 71.76$, $p < 0.001$, $\eta p^2 = 0.533$) but not for round; for older adults, our results showed a significant difference for both round ($F(1,60) = 7.35$, $p = 0.009$, $\eta p^2 = 0.109$) and N -Back level ($F(2,60) = 15.24$, $p < 0.001$, $\eta p^2 = 0.337$).

3.2. Gender Differences in Reaction Time. To detect further differences, we used a two-way ANOVA (gender \times N -Back level) for the first and last round separately (Figure 4). Our results showed significant differences for young adults in the first round for gender ($F(1,51) = 5.50$, $p = 0.023$, $\eta p^2 = 0.097$) and N -Back level ($F(2,51) = 8.34$, $p = 0.001$, $\eta p^2 = 0.246$). Looking at the fourth round, our results showed a significant main effect for N -Back level ($F(2,51) = 4.13$, $p = 0.022$, $\eta p^2 = 0.139$).

We performed the same analyses for older adults and found significant differences only in the fourth round for

N -Back level ($F(2,42) = 3.94$, $p = 0.027$, $\eta p^2 = 0.158$). For both rounds, no significant differences between males and females were found.

Our post hoc analyses of gender revealed a significantly shorter RT for young females ($M = 0.431$; $SD = 0.019$) compared to young males ($M = 0.497$; $SD = 0.025$) in the first round ($t(55) = 2.1101$, $p = 0.0394$), considering all three N -Back levels.

Furthermore, our post hoc analyses of N -Back level showed significant differences in RT for young adults in the first round between 1-Back ($M = 0.406$; $SD = 0.021$) and 3-Back ($M = 0.542$; $SD = 0.030$) ($t(36) = 3.6426$, $p < 0.001$) and between 2-Back ($M = 0.440$; $SD = 0.022$) and 3-Back ($M = 0.542$; $SD = 0.030$) ($t(36) = 2.7030$, $p = 0.0104$) and in the fourth round between 1-Back ($M = 0.036$; $SD = 0.017$) and 3-Back ($M = 0.459$; $SD = 0.031$) ($t(36) = 2.6634$, $p = 0.0115$). Similarly, our results showed significant RT differences for older adults between 1-Back ($M = 0.461$; $SD = 0.016$) and 2-Back ($M = 0.535$; $SD = 0.021$) ($t(62) = 2.8631$, $p = 0.0057$) and between 1-Back ($M = 0.461$; $SD = 0.016$) and 3-Back ($M = 0.549$; $SD = 0.026$) ($t(62) = 2.9162$, $p = 0.0049$) when taking the first and last rounds together. Furthermore, we found significant differences in RT of older

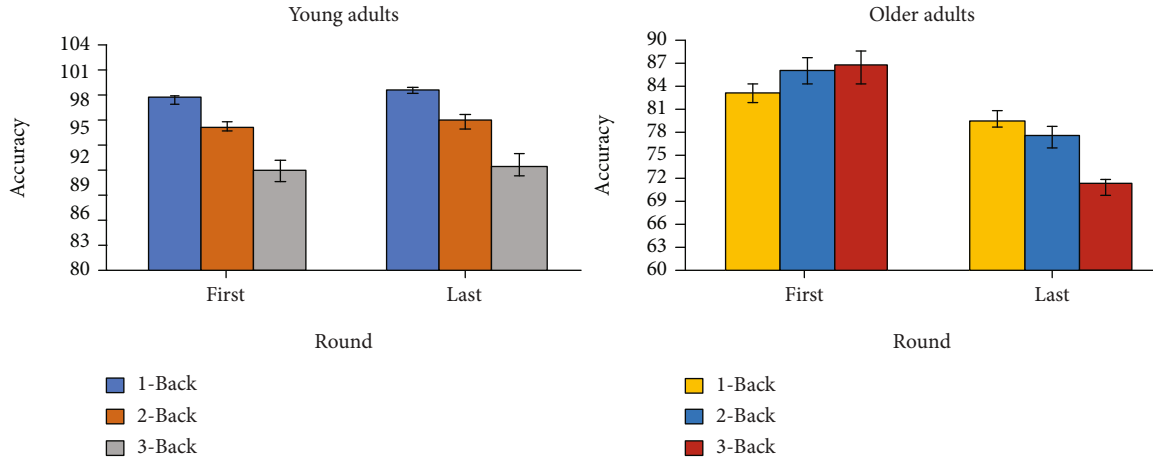


FIGURE 3: Means \pm SD for accuracy for young (a) and older (b) adults comparing first and last round during N -Back task performance for the three different N -Back levels.

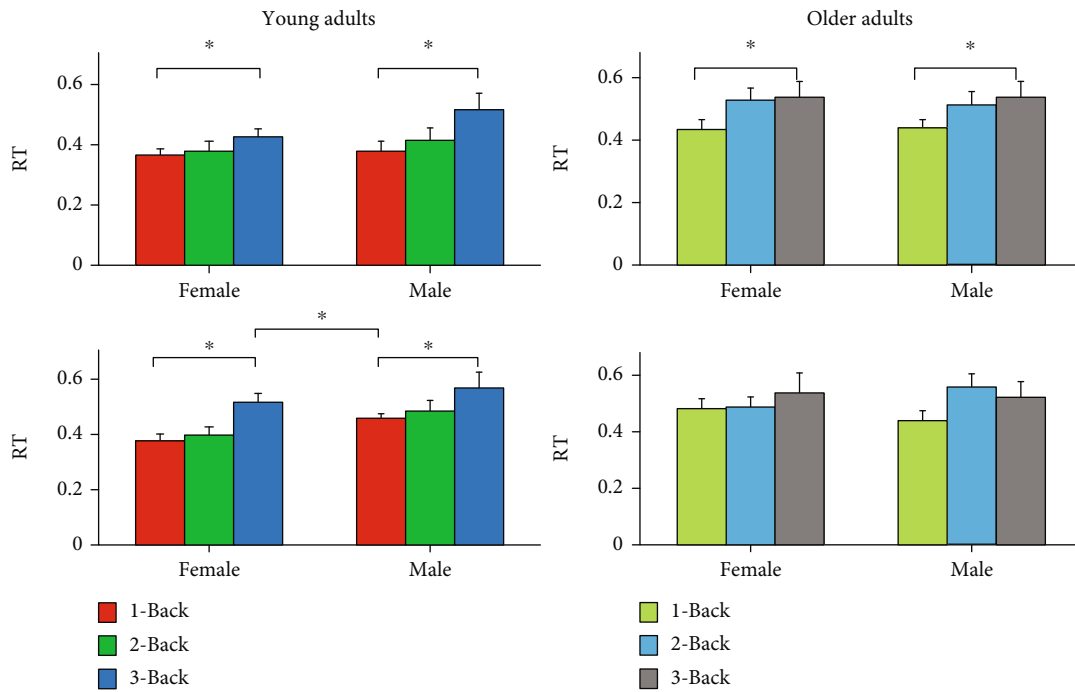


FIGURE 4: Means \pm SD for reaction time (RT) for young (a) and older (b) adults comparing females and males during N -Back task performance for the three different N -Back levels.

adults in the fourth round between 1-Back ($M = 0.448$; $SD = 0.020$) and 2-Back ($M = 0.534$; $SD = 0.027$) ($t(30) = 2.5631$, $p = 0.0156$) and between 1-Back ($M = 0.448$; $SD = 0.020$) and 3-Back ($M = 0.5531$; $SD = 0.033$) ($t(30) = 2.7171$, $p = 0.0108$).

3.3. Gender Differences in Accuracy. To detect significant accuracy differences during the first and last round, we used a two-way ANOVA (gender \times N -Back level). Our results showed significant differences in the first round for young adults for gender ($F(1, 60) = 7.92$, $p = 0.007$, $\eta p^2 = 0.117$) and N -Back level ($F(2, 60) = 50.86$, $p < 0.001$, $\eta p^2 = 0.629$) and in the fourth round for gender ($F(1, 63) = 27.45$, $p <$

0.001 , $\eta p^2 = 0.306$), N -Back level ($F(2, 63) = 51.45$, $p < 0.001$, $\eta p^2 = 0.620$), and gender \times N -Back level ($F(2, 63) = 4.92$, $p = 0.010$, $\eta p^2 = 0.135$). Furthermore, we performed the same analyses for older adults and found significant differences in the fourth round for N -Back level ($F(2, 42) = 3.46$, $p = 0.041$, $\eta p^2 = 0.141$), but not in the first round. No significant differences for gender across N -Back levels were found in both rounds.

Our post hoc analyses of young adults for gender revealed a trend for higher accuracies for females ($M = 95.833$; $SD = 0.621$) compared to males ($M = 93.796$; $SD = 0.0986$) ($t(64) = 1.837$, $p = 0.0707$) in the first round, considering all three N -Back levels.

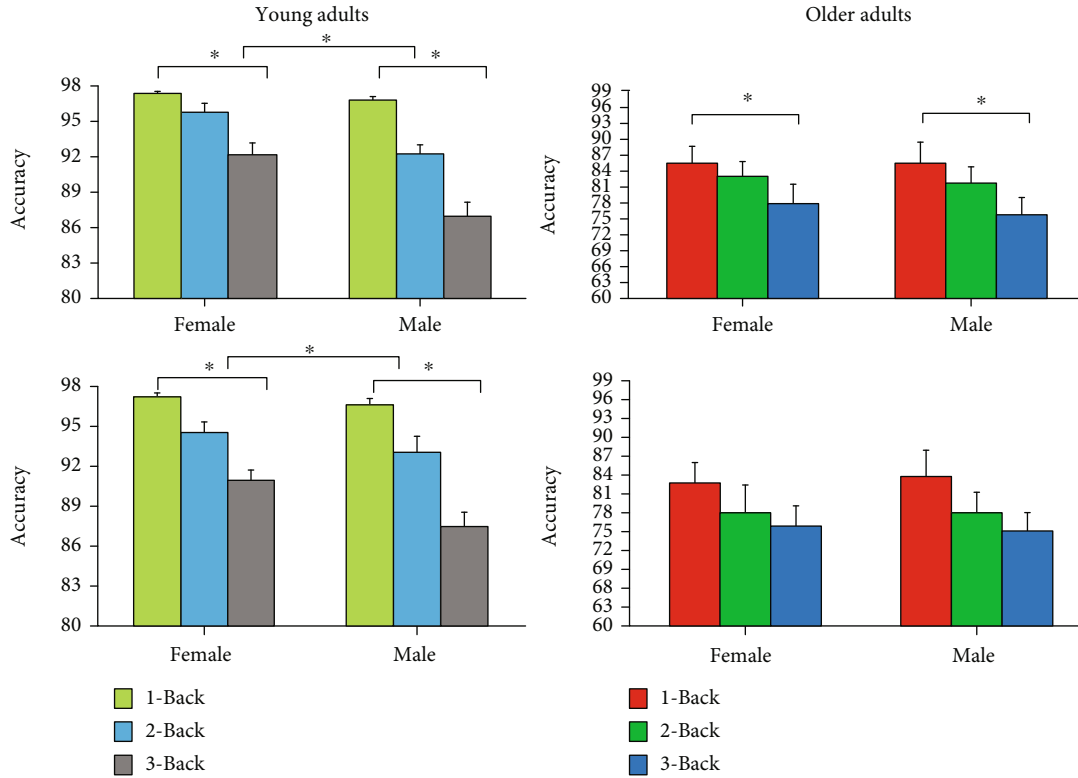


FIGURE 5: Means \pm SD for accuracy for young (a) and older (b) adults comparing females and males during N -Back task performance for the three N -Back levels.

Moreover, our post hoc analyses of young adults for N -Back level showed significant differences in accuracy, when considering the first and last rounds jointly, for 1-Back ($M = 99.130$; $SD = 0.168$) and 2-Back ($M = 95.853$; $SD = 0.513$) performance ($t(86) = 6.2951$, $p < 0.001$), between 1-Back ($M = 99.130$; $SD = 0.168$) and 3-Back ($M = 91.023$; $SD = 0.657$) performance ($t(88) = 12.1853$, $p < 0.001$), and between 2-Back ($M = 95.853$; $SD = 0.513$) and 3-Back ($M = 91.023$; $SD = 0.657$) performance ($t(84) = 5.7557$, $p < 0.001$) for young adults (Figure 5). Furthermore, our results showed significant differences in older adults between 1-Back ($M = 79.091$; $SD = 0.012$) and 2-Back ($M = 75.606$; $SD = 1.220$) performance ($t(42) = 2.1981$, $p = 0.0335$), between 1-Back ($M = 79.091$; $SD = 0.012$) and 3-Back ($M = 71.477$; $SD = 0.824$) performance ($t(42) = 5.8324$, $p < 0.001$), and between 2-Back ($M = 75.606$; $SD = 1.220$) and 3-Back ($M = 71.477$; $SD = 0.824$) performance ($t(42) = 2.8042$, $p = 0.0076$), also when considering the first and last rounds jointly. Moreover, significant differences between rounds were found in older adults. Specifically, accuracy level was higher for the last round ($M = 76.919$; $SD = 1.042$) compared to the first round ($M = 73.864$; $SD = 0.878$) ($t(64) = 2.2427$, $p = 0.0284$).

Furthermore, we found significant differences for young adults in the fourth round between 1-Back ($M = 99.456$; $SD = 0.186$) and 2-Back ($M = 96.341$; $SD = 0.684$) performance ($t(44) = 4.3923$, $p = 0.0143$), between 1-Back ($M = 99.456$; $SD = 0.186$) and 3-Back ($M = 91.594$; $SD = 1.008$) performance ($t(44) = 7.6730$, $p < 0.001$), and between

2-Back ($M = 96.341$; $SD = 0.684$) and 3-Back ($M = 91.594$; $SD = 1.008$) performance ($t(44) = 3.8962$, $p < 0.001$). Also, we found significant differences in the first round between 1-Back ($M = 98.804$; $SD = 0.266$) and 2-Back ($M = 95.476$; $SD = 0.713$) performance ($t(44) = 4.3923$, $p = 0.0143$), between 1-Back ($M = 98.804$; $SD = 0.266$) and 3-Back ($M = 90.568$; $SD = 0.808$) performance ($t(43) = 9.8538$, $p < 0.001$), and between 2-Back ($M = 95.476$; $SD = 0.713$) and 3-Back ($M = 90.568$; $SD = 0.808$) performance ($t(41) = 4.5382$, $p < 0.001$). Similarly, we found significant differences for older adults in the fourth round between 1-Back ($M = 86.198$; $SD = 2.417$) and 2-Back ($M = 82.917$; $SD = 2.160$) performance ($t(42) = 4.5272$, $p < 0.001$).

In order to address the possible effect of imbalanced gender, we ran a bootstrapping procedure (see Materials and Method and Statistical Analyses, for further details). We obtained a p value of 0.0273, with a 95% CI in mean accuracy difference of [0.178572, 1.369047], for young adults in the last (fourth) round for 1-Back, indicating a significant higher accuracy for females compared to males. Moreover, we ran the bootstrapping test which returned a p value of 0.0009, with a 95% CI of [1.574074, 6.296297], for young adults in the last round for 2-Back, indicating a significant higher accuracy for females compared to males. Last, we ran the bootstrapping test that showed a p value of 0.0010, with a 95% CI of [2.407407, 9.259259], for young adults in the last round for 3-Back, indicating a significant higher accuracy for females compared to males. Furthermore, we applied the same method for the first round in young adults.

The bootstrapping test returned a p value of 0.166, with a 95% CI of [-0.462963, 1.851852], for 1-Back, indicating a nonsignificant accuracy difference between females and males. Also, the bootstrapping test showed a p value of 0.1376, with a 95% CI of [-1.296297, 4.537037], for 2-Back, indicating a nonsignificant accuracy difference between females and males. Last, the bootstrapping approach returned a p value of 0.0097, with a 95% CI of [0.648148, 6.574074], for 3-Back, indicating a significant higher accuracy for females compared to males. To summarize, our results showed that females performed better than males in the last round for all three N -Back levels and in the first round for the 3-Back task. No significant differences were found in the first round for the 2 and 3-Back levels between females and males.

4. Discussion

We investigated the accuracy and reaction time of a demanding WM task, called N -Back, for three different difficulty levels, in healthy young and older adults. Based on the studies of Käthner et al. [48] and Kohlmorgen et al. [49], which demonstrated the impact of mental fatigue between the beginning and the end of cognitive performance, especially in older individuals, we split our N -Back data into first (1st) and last (4th) rounds. However, besides mental fatigue, learning might be a moderator of improved performance outcomes [51]. The N -Back task has been commonly shown to depend on individual characteristics such as age and education level. However, the influence of gender has largely been neglected, especially in older individuals, and the few studies that have been conducted returned inconsistent results [17, 19]. Speck et al. [34] observed a female advantage for a verbal N -Back task that uses sequences of displayed words and numbers and reported gender-specific neural patterns. On the other hand, the study of Lejbak et al. [28] revealed a male advantage when using an object-location memory task (N -Back) in young adults. Schmidt et al. [17], using a verbal N -Back task in young adults, observed that WM is not affected by gender at behavioral and neural levels, while Pliatsikas et al. [18], using a digit N -Back task in older adults, showed that with increasing age, males exhibited a steeper WM decline than females, and with increasing education, females exhibited greater WM gains than males. Similarly, Lejbak et al. [24] reported a gender-related difference in favor of females for an object location memory task, in young adults. McCarrey et al. [23] showed that males outperformed females in two visuospatial ability tasks and that females outperformed males in most other tests of cognition. However, gender differences over time revealed steeper rates of age-related decline for men on measures of mental status, perceptuomotor speed and integration, and visuospatial ability. Taking into consideration these studies, we hypothesized an advantage for females compared to males.

Our results showed that mental fatigue did not affect N -Back task performance in both young and older adults, but instead, we observed a higher accuracy level in the last round

compared to the first round for older adults, suggesting that mental fatigue was countered by a learning effect. Indeed, although mental fatigue might be present during the last part of cognitive task performance, learning is another factor that could interfere with performance. It has been shown that repeated practice of a cognitive task involving visual stimuli can improve implicit memory and attention processes [51, 56], resulting in increased performance even within one hour. This process has been defined as fast perceptual learning (PL) [51, 57]. Beside this, a persistent higher accuracy level was observed for the easier 1-Back level compared to the more difficult 2- and 3-Back levels.

Furthermore, according to our expectations regarding gender, the outcomes of young subjects were in line with the results observed by Pliatsikas et al. [18] and Speck et al. [34], as we found a general female advantage for both accuracy and RT for the N -Back task performance. Additionally, our bootstrapping test, which we used to balance the number of females and males in the group of young adults, confirmed that females enjoy higher accuracy levels compared to males in all three N -Back levels, especially in the last round. Further studies should focus on the underlying factors producing these differences, e.g., whether the stimuli used were more or less familiar. One hypothesis is that the male advantage can be explained by a better performance in terms of semantic memory for common objects [28], whereas worse male performance has been seen for other object categories, such as words, numbers, shapes, faces, and less familiar objects. Another possibility can be attributed to gender-related differences in brain structure as demonstrated by Xin et al. [36] which might be reflected by gender-related differences in cognition. An additional hypothesized moderator could be the specific strategy used to perform the task instead of the task per se [58]. From this perspective, tasks using different stimulus types could not be the prime reason for differences in outcome, whereas the cognitive processes used to solve the WM task may instead play an important role in accuracy and RT responses [37]. Moreover, Upadhayay and Guragain [41] suggested that hormones, especially progesterone, could have a significant impact on cognitive performance. While studies have already shown that estrogen and testosterone accentuate cognitive functions, their study indicated that male cognitive functions were comparable to female preovulatory phase cognitive functions, although females, during the postovulatory phase of their cycle, enjoyed advantages in executive tasks and disadvantages in attentional tasks, compared to males.

The present study has important implications in terms of behavioral outcomes. We examined whether accuracy and RT additionally affected WM task performance as well as mental fatigue and learning. Importantly, we showed that gender plays an important role in young, but not in older adults. However, previous studies indicated that gender does not contribute to differences in behavioral outcomes [23], while in contrast, other studies [33, 34] reported differences in neural activation pathways of males and females. Although the N -Back task has been commonly used for assessing WM in a variety of populations, such as young

and older subjects, and individuals with neurological disorders, the role of gender in cognitive performance in older adults is still elusive.

Our findings should be considered in light of some limitations. It has been shown that increasing stimulus duration improves memory performance during retention of visuospatial information [50, 59]. Consequently, it is important to note that, similar to the study of Speck et al. [34], our stimuli were presented for 1 second, whereas most of the studies used a longer stimulus duration [60, 61]. This factor might have affected our RT and accuracy outcomes. Moreover, in this study, we used an object *N*-Back task; however, further studies should compare different *N*-Back task modalities (animals, human faces, geometric shapes, and inanimate/animate stimuli), which have been considered in the study of Lejbak et al. [28] albeit they considered only a 2-Back task (hence, no variation in *N*-Back level). Also, our final outcomes showed gender difference only for young, but not older adults. This pattern could reflect a ceiling effect; however, if true, then we probably would have seen some gender differences in at least one of the three *N*-Back difficulty levels, which is not the case. We suggest more research to address this issue. While here we reported accuracy and RT outcomes, combining neuroimaging data recorded when performing an *N*-Back task could be a worthwhile addition to study gender differences. Last, our test sample was limited to examine both the effect of age and gender. A large test sample size should be used to verify whether the effect of gender in young but not in older adults could be replicated. The reason behind the small groups is due to the fact that most of the participants came from our previous studies ([52] for older adults; [53] for young adults) that investigated the effect of WM training, both in terms of behavioral and EEG responses, requiring several in-person training sessions for each subject.

Data Availability

The data of this study are available via this link: <https://drive.google.com/drive/folders/1eFhk33uOAFdyiHBuF4olwoIUgOxBHdRo?usp=sharing>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

VP and NV collected the data and conducted the experiments; VP designed the experiments, wrote the manuscript, and performed the analyses; NV contributed to the performance of the analyses. All authors discussed the results and contributed to the final manuscript. Valentina Pergher and Nele Vanbilsen share co-first authorship.

Acknowledgments

VP was supported by a Postdoctoral Mandate (PDM) from KU Leuven and by research grants to MMVH from the

European Union's Horizon 2020 research and innovation program under grant agreement No. 857375, the Financing program (PFV/10/008), and a special research fund project (C24/18/098) of the KU Leuven, the Fonds De La Recherche Scientifique - FNRS (Belgian Fund for Scientific Research-Flanders) (G088314N, G0A0914N, and G0A4118N), and the Hercules Foundation (AKUL 043).

References

- [1] A. Baddeley, "Working memory: looking back and looking forward," *Nature Reviews Neuroscience*, vol. 4, no. 10, pp. 829–839, 2003.
- [2] M. J. Kane and R. W. Engle, "The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective," *Psychonomic Bulletin & Review*, vol. 9, no. 4, pp. 637–671, 2002.
- [3] R. H. Logie, "The functional organization and capacity limits of working memory," *Current Directions in Psychological Science*, vol. 20, no. 4, pp. 240–245, 2011.
- [4] L. Cahill, "Why sex matters for neuroscience," *Nature Reviews Neuroscience*, vol. 7, no. 6, pp. 477–484, 2006.
- [5] H. Christensen, "What cognitive changes can be expected with normal ageing?," *Australian & New Zealand Journal of Psychiatry*, vol. 35, no. 6, pp. 768–775, 2001.
- [6] L. L. Drag, L. A. Bieliauskas, S. A. Langenecker, and L. J. Greenfield, "Cognitive functioning, retirement status, and age: results from the Cognitive Changes and Retirement among Senior Surgeons study," *Journal of the American College of Surgeons*, vol. 211, no. 3, pp. 303–307, 2010.
- [7] S. E. MacPherson, L. H. Phillips, and S. Della Sala, "Age, executive function and social decision making: a dorsolateral prefrontal theory of cognitive aging," *Psychology and Aging*, vol. 17, no. 4, pp. 598–609, 2002.
- [8] S. Cansino, E. Hernández-Ramos, C. Estrada-Manilla et al., "The decline of verbal and visuospatial working memory across the adult life span," *Age*, vol. 35, no. 6, pp. 2283–2302, 2013.
- [9] S. Hale, N. S. Rose, J. Myerson et al., "The structure of working memory abilities across the adult life span," *Psychology and Aging*, vol. 26, no. 1, pp. 92–110, 2011.
- [10] G. Cona, C. Scarpazza, G. Sartori, M. Moscovitch, and P. S. Bisiacchi, "Neural bases of prospective memory: a meta-analysis and the "Attention to Delayed Intention" (AtoDI) model," *Neuroscience & Biobehavioral Reviews*, vol. 52, pp. 21–37, 2015.
- [11] D. C. Park, G. Lautenschlager, T. Hedden, N. S. Davidson, A. D. Smith, and P. K. Smith, "Models of visuospatial and verbal memory across the adult life span," *Psychology and Aging*, vol. 17, no. 2, pp. 299–320, 2002.
- [12] F. Schmiedek, S. C. Li, and U. Lindenberger, "Interference and facilitation in spatial working memory: age-associated differences in lure effects in the n-back paradigm," *Psychology and Aging*, vol. 24, no. 1, pp. 203–210, 2009.
- [13] W. A. Meijer, R. H. M. de Groot, M. P. J. van Boxtel, P. W. M. van Gerven, and J. Jolles, "Verbal learning and aging: combined effects of irrelevant speech, interstimulus interval, and education," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 61, no. 5, pp. 285–294, 2006.
- [14] R. M. Ruff and S. B. Parker, "Gender-and age-specific changes in motor speed and eye-hand coordination in adults:

- normative values for the Finger Tapping and Grooved Pegboard Tests," *Perceptual and motor skills*, vol. 76, 3_supplement, pp. 1219–1230, 1993.
- [15] Y. Stern, "What is cognitive reserve? Theory and research application of the reserve concept," *Journal of the international neuropsychological society*, vol. 8, no. 3, pp. 448–460, 2002.
- [16] L. B. Zahodne, Y. Stern, and J. J. Manly, "Differing effects of education on cognitive decline in diverse elders with low versus high educational attainment," *Neuropsychology*, vol. 29, no. 4, pp. 649–657, 2015.
- [17] H. Schmidt, J. Jogia, K. Fast et al., "No gender differences in brain activation during the N-back task: an fMRI study in healthy individuals," *Human Brain Mapping*, vol. 30, no. 11, pp. 3609–3615, 2009.
- [18] C. Pliatsikas, L. Meteyard, J. Verissimo, V. DeLuca, K. Shattuck, and M. T. Ullman, "The effect of bilingualism on brain development from early childhood to young adulthood," *Brain Structure and Function*, vol. 225, no. 7, pp. 2131–2152, 2020.
- [19] A. C. Hill, A. R. Laird, and J. L. Robinson, "Gender differences in working memory networks: a BrainMap meta-analysis," *Biological Psychology*, vol. 102, pp. 18–29, 2014.
- [20] G. Alarcón, A. Cservenka, D. A. Fair, and B. J. Nagel, "Sex differences in the neural substrates of spatial working memory during adolescence are not mediated by endogenous testosterone," *Brain Research*, vol. 1593, pp. 40–54, 2014.
- [21] H. L. Hsu, D. Y. T. Chen, Y. C. Tseng et al., "Sex differences in working memory after mild traumatic brain injury: a functional MR imaging study," *Radiology*, vol. 276, no. 3, pp. 828–835, 2015.
- [22] R. M. Martoni, G. Salgari, E. Galimberti, M. C. Cavallini, and J. O'Neill, "Effects of gender and executive function on visuo-spatial working memory in adult obsessive-compulsive disorder," *European Archives of Psychiatry and Clinical Neuroscience*, vol. 265, no. 8, pp. 707–718, 2015.
- [23] A. C. McCarrey, Y. An, M. H. Kitner-Triolo, L. Ferrucci, and S. M. Resnick, "Sex differences in cognitive trajectories in clinically normal older adults," *Psychology and Aging*, vol. 31, no. 2, pp. 166–175, 2016.
- [24] L. Lejbak, M. Vrbancic, and M. Crossley, "The female advantage in object location memory is robust to verbalizability and mode of presentation of test stimuli," *Brain and Cognition*, vol. 69, no. 1, pp. 148–153, 2009.
- [25] Q. Rahman, S. Abrahams, and F. Jussab, "Sex differences in a human analogue of the radial arm maze: the '17-Box Maze Test'," *Brain and Cognition*, vol. 58, no. 3, pp. 312–317, 2005.
- [26] J. L. Reed, N. M. Gallagher, M. Sullivan, J. H. Callicott, and A. E. Green, "Sex differences in verbal working memory performance emerge at very high loads of common neuroimaging tasks," *Brain and Cognition*, vol. 113, pp. 56–64, 2017.
- [27] D. Voyer, S. D. Voyer, and J. Saint-Aubin, "Sex differences in visual-spatial working memory: a meta-analysis," *Psychonomic Bulletin & Review*, vol. 24, no. 2, pp. 307–334, 2017.
- [28] L. Lejbak, M. Crossley, and M. Vrbancic, "A male advantage for spatial and object but not verbal working memory using the n-back task," *Brain and Cognition*, vol. 76, no. 1, pp. 191–196, 2011.
- [29] L. S. Tottenham, D. Saucier, L. Elias, and C. Gutwin, "Female advantage for spatial location memory in both static and dynamic environments," *Brain and Cognition*, vol. 53, no. 2, pp. 381–383, 2003.
- [30] D. Voyer, A. Postma, B. Brake, and J. Imperato-McGinley, "Gender differences in object location memory: a meta-analysis," *Psychonomic Bulletin & Review*, vol. 14, no. 1, pp. 23–38, 2007.
- [31] M. A. Norman, J. D. Evans, W. S. Miller, and R. K. Heaton, "Demographically corrected norms for the California verbal learning test," *Journal of Clinical and Experimental Neuropsychology*, vol. 22, no. 1, pp. 80–94, 2000.
- [32] E. M. Weiss, J. D. Ragland, C. M. Bressinger, W. B. Bilker, E. A. Deisenhammer, and M. Delazer, "Sex differences in clustering and switching in verbal fluency tasks," *Journal of the International Neuropsychological Society*, vol. 12, no. 4, pp. 502–509, 2006.
- [33] J. M. Goldstein, M. Jerram, R. Poldrack et al., "Sex differences in prefrontal cortical brain activity during fMRI of auditory verbal working memory," *Neuropsychology*, vol. 19, no. 4, pp. 509–519, 2005.
- [34] O. Speck, T. Ernst, J. Braun, C. Koch, E. Miller, and L. Chang, "Gender differences in the functional organization of the brain for working memory," *Neuroreport*, vol. 11, no. 11, pp. 2581–2585, 2000.
- [35] K. P. Cosgrove, C. M. Mazure, and J. K. Staley, "Evolving knowledge of sex differences in brain structure, function, and chemistry," *Biological Psychiatry*, vol. 62, no. 8, pp. 847–855, 2007.
- [36] J. Xin, Y. Zhang, Y. Tang, and Y. Yang, "Brain differences between men and women: evidence from deep learning," *Frontiers in Neuroscience*, vol. 13, p. 185, 2019.
- [37] X. J. Chai and L. F. Jacobs, "Effects of cue types on sex differences in human spatial memory," *Behavioural Brain Research*, vol. 208, no. 2, pp. 336–342, 2010.
- [38] B. Pletzer, T. A. Harris, and T. Ortner, "Sex and menstrual cycle influences on three aspects of attention," *Physiology & Behavior*, vol. 179, pp. 384–390, 2017.
- [39] E. Riley, H. Okabe, L. Germine, J. Wilmer, M. Esterman, and J. DeGutis, "Gender differences in sustained attentional control relate to gender inequality across countries," *PLoS One*, vol. 11, no. 11, article e0165100, 2016.
- [40] L. M. Giambra and R. E. Quilter, "Sex differences in sustained attention across the adult life span," *Journal of Applied Psychology*, vol. 74, no. 1, pp. 91–95, 1989.
- [41] N. Upadhayay and S. Guragain, "Comparison of cognitive functions between male and female medical students: a pilot study," *Journal of Clinical and Diagnostic Research: JCDR*, vol. 8, no. 6, pp. BC12–BC15, 2014.
- [42] A. M. Owen, K. M. McMillan, A. R. Laird, and E. Bullmore, "N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies," *Human Brain Mapping*, vol. 25, no. 1, pp. 46–59, 2005.
- [43] D. J. Moore, E. Keogh, and C. Eccleston, "Identifying experimental methods to determine the effect of pain on attention: a review of pain, caffeine, alcohol and nicotine studies," *Human Psychopharmacology: Clinical and Experimental*, vol. 24, no. 8, pp. 601–618, 2009.
- [44] M. Pesonen, H. Hämäläinen, and C. M. Krause, "Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load," *Brain Research*, vol. 1138, pp. 171–177, 2007.
- [45] Y. N. Chen, S. Mitra, and F. Schlaghecken, "Sub-processes of working memory in the N-back task: an investigation using ERPs," *Clinical Neurophysiology*, vol. 119, no. 7, pp. 1546–1559, 2008.

- [46] T. A. Salthouse, "The processing-speed theory of adult age differences in cognition," *Psychological Review*, vol. 103, no. 3, pp. 403–428, 1996.
- [47] M. A. Boksem, T. F. Meijman, and M. M. Lorist, "Effects of mental fatigue on attention: an ERP study," *Cognitive Brain Research*, vol. 25, no. 1, pp. 107–116, 2005.
- [48] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface," *Biological Psychology*, vol. 102, pp. 118–129, 2014.
- [49] J. Kohlmorgen, G. Dornhege, M. Braun et al., "Improving human performance in a real operating environment through real-time mental workload detection," *Toward Brain-Computer Interfacing*, vol. 409422, pp. 409–422, 2007.
- [50] V. Pergher, B. Wittevrongel, J. Tournoy, B. Schoenmakers, and M. M. Van Hulle, "Mental workload of young and older adults gauged with ERPs and spectral power during N-Back task performance," *Biological Psychology*, vol. 146, article 107726, 2019.
- [51] Y. Yotsumoto, T. Watanabe, and Y. Sasaki, "Different dynamics of performance and brain activation in the time course of perceptual learning," *Neuron*, vol. 57, no. 6, pp. 827–833, 2008.
- [52] V. Pergher, N. Vanbilsen, J. Tournoy, B. Schoenmakers, and M. M. Van Hulle, "Impact of strategy use during N-Back training in older adults," *Journal of Cognitive Psychology*, vol. 32, no. 8, pp. 715–733, 2020.
- [53] V. Pergher, B. Wittevrongel, J. Tournoy, B. Schoenmakers, and M. M. Van Hulle, "N-Back training and transfer effects revealed by behavioral responses and EEG," *Brain and Behavior*, vol. 8, no. 11, article e01136, 2018.
- [54] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for the exponential distribution (complete samples)," *Technometrics*, vol. 14, no. 2, pp. 355–370, 1972.
- [55] J. Cohen, "Eta-squared and partial eta-squared in fixed factor ANOVA designs," *Educational and Psychological Measurement*, vol. 33, no. 1, pp. 107–112, 1973.
- [56] C. Alain, J. S. Snyder, Y. He, and K. S. Reinke, "Changes in auditory cortex parallel rapid perceptual learning," *Cerebral Cortex*, vol. 17, no. 5, pp. 1074–1084, 2007.
- [57] T. Poggio, M. Fahle, and S. Edelman, "Fast perceptual learning in visual hyperacuity," *Science*, vol. 256, no. 5059, pp. 1018–1021, 1992.
- [58] V. Pergher, M. A. Shalchy, A. Pahor, M. M. Van Hulle, S. M. Jaeggi, and A. R. Seitz, "Divergent research methods limit understanding of working memory training," *Journal of Cognitive Enhancement*, vol. 4, no. 1, pp. 100–120, 2020.
- [59] M. Kunimi, "Effects of age, gender, and stimulus presentation period on visual short-term memory," *Journal of Women & Aging*, vol. 28, no. 1, pp. 24–33, 2016.
- [60] A. Costa, A. Peppe, G. Dell'Agnello, C. Caltagirone, and G. A. Carlesimo, "Dopamine and cognitive functioning in de novo subjects with Parkinson's disease: effects of pramipexole and pergolide on working memory," *Neuropsychologia*, vol. 47, no. 5, pp. 1374–1381, 2009.
- [61] B. A. Parmenter, J. L. Shucard, R. H. B. Benedict, and D. W. Shucard, "Working memory deficits in multiple sclerosis: comparison between the *n*-back task and the paced auditory serial addition test," *Journal of the International Neuropsychological Society: JINS*, vol. 12, no. 5, pp. 677–687, 2006.