Aging modulates the effects of scene complexity on visual search and target selection in virtual environments

Isaiah J. Lachica[1], Aniruddha Kalkar[2], and James M. Finley[1,3,4]*

[1]Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, 90033, USA

[2]Department of Computer Science, University of Southern California, Los Angeles, 90089, USA

[3]Neuroscience Graduate Program, University of Southern California, Los Angeles, 90089, USA

[4]Department of Biomedical Engineering, University of Southern California, Los Angeles, 90089, USA

*jmfinley@usc.edu

Abstract

Processing task-relevant visual information is important for successfully completing many everyday tasks. Prior work demonstrated that older adults are more susceptible to distraction by salient stimuli, leading to less efficient visual search. However, these studies often relied on simple stimuli, and little is known about how aging influences visual attention in environments more representative of real-world complexity. Here, we test the hypothesis that aging impacts how the visual complexity of the environment influences visual search. We asked young and older adults to complete a virtual reality-based visual search task with three levels of increasing visual complexity. As visual complexity increased, all participants took longer to complete the task. This increase in time to completion resulted from two factors: an increased time required to transfer gaze from one target to the next as participants re-fixated task-relevant objects and longer lags between when correct targets were fixated and selected. We also found that these changes in gaze trajectory were greater in older adults. In addition, we found that short-term and working memory capacities were positively associated with multiple performance measures in the visual search task. This suggests that visual search performance could be integrated into assessments of working memory in dynamic environments.

Introduction

Selecting and processing relevant visual information from the environment is vital for planning and executing everyday tasks[1–4]. For example, to successfully find one's luggage at a busy airport, it is essential to attend to visual features such as color and size to distinguish one's luggage from another passenger's. Both natural and human-made environments are filled with a broad range of visual stimuli that may be relevant or irrelevant for guiding future actions. The control of visual attention has been prominently described as a balance between top-down and bottom-up processes. Bottom-up processes emphasize saliency, a characteristic of visual stimuli based on low-level visual features such as contrast and luminance[5]. Highly salient visual stimuli have the effect of "popping out" of a scene and capturing attention automatically. In contrast, top-down processes control visual attention based on cognitive control processes that prioritize factors such as task demands and prior knowledge[6]. When selecting a stimulus to direct visual attention towards, top-down cognitive processes modulate bottom-up saliency by enhancing the representation of a desired region or stimuli while inhibiting others[7–9]. The use of top-down cognitive control to modulate bottom-up processes has also been demonstrated at the neural level[10] as brain activity increases for target objects and decreases for distractors. As such, while highly salient stimuli often capture attention automatically, top-down cognitive control can inhibit this process, allowing attention to be focused on relevant visual information. Task-dependent trade-offs between top-down and bottom-up control enable attention to be directed based on top-down factors when performing a task while allowing it to be captured by highly salient stimuli when necessary. However, the balance between bottom-up and top-down control appears to rely heavily on an individual's cognitive capacity, which has been demonstrated to be affected by normal aging[11,12].

There is evidence that older adults are more susceptible to distraction by irrelevant visual stimuli due to aging-related changes in their cognition. In particular, their ability to exert top-down inhibition has been shown to be impaired relative to young adults[13–15]. These effects

prevent older adults from suppressing the automatic capture of attention by salient task-irrelevant stimuli, negatively affecting their performance in visually demanding tasks[16–19]. Impaired capacity for inhibition can also lead to task-irrelevant information being encoded into working memory at the cost of task-relevant information[17,20]. This increased distractibility from task-irrelevant stimuli is problematic as this may prevent older adults from efficiently completing everyday tasks and make them more vulnerable to accidents that may lead to injuries. However, prior studies of age-related changes in visual search used simple visual scenes composed of arbitrarily selected shapes and letters. It is unclear whether these aging-related effects on visual attention extend to performing tasks in scenes more closely replicating the visual complexity in real-world environments.

Assessing the cognitive capacity of older adults, particularly their ability to direct visual attention appropriately, is important for developing interventions to improve their function in daily activities. The Trail Making Test[21], especially part B, is a visual search task commonly used to assess cognitive domains such as visual attention, working memory, and inhibition[22–24]. It is typically implemented in a pen-and-paper format with numbers and letters as the search targets and time to completion as the outcome measure. Recent attempts have been made to increase the ecological validity of the Trail Making Test by transforming it into a three-dimensional reaching task in virtual reality (VR)[25]. However, the context in which people search for targets in both implementations of the Trail Making Test does not include distractors as would be present in real-world environments, leaving questions about how these results generalize to the real-world. In addition, neither version of the Trail Making Test provides information on gaze behaviors such as fixation location which reflects the deployment of visual attention[4,5,26].

We aimed to determine how aging-related differences in cognition influence visual search and attention in increasingly complex virtual environments. Participants completed a custom-designed VR-based visual search task in three visual complexity levels that simulated real-world environments[27]. To determine the influence of visual attention on task performance,

we tracked and recorded eye movements from which we calculated search times between correct targets, fixation time on task-relevant objects, fixation time on task-irrelevant objects, and the saliency of fixated regions of the visual scene. We also determined how the actions involved in target selection influenced task completion time by measuring the number of selection errors and the lag between fixating and selecting a correct target. Additionally, all participants completed assessments of global cognition, short-term memory, working memory, and inhibitory capacity. We hypothesized that as the complexity of the visual scene increased, participants would take longer to complete the task because attention would be attracted by task-irrelevant stimuli, leading to a reduced capacity to encode correct targets in working memory. In particular, all participants would be more prone to re-fixating search targets and directing their attention to salient, task-irrelevant distractors, leading to longer search times. In addition, we hypothesized that this performance decline would be greater in older adults due to aging-related impairments in their cognition.

Results

Task completion times. All participants took longer to complete the task as the visual complexity of the environment increased, and this effect was greatest in older adults (Fig. 1). We found main effects of age group ($F(1,26) = 48.17$, $p = 2.28e\text{-}7$) and visual complexity level ($F(2,52) = 74.82$, $p = 4.98e\text{-}16$) on task completion time. We also found an interaction between age group and visual complexity level ($F(2,52) = 9.27$, $p = 3.60e\text{-}4$). Compared to older adults, young adults completed the task 9.4 seconds faster in the low complexity level (Bonferroni-corrected $p = 0.0036$), 11.9 seconds faster in the medium complexity level (Bonferroni-corrected $p < 0.0001$), and 19.3 seconds faster in the high complexity level (Bonferroni-corrected $p < 0.0001$). In addition, the young adults were 8.8 seconds slower in the low versus high visual complexity (Bonferroni-corrected $p < 0.0001$) and were 7.6 seconds slower in the medium versus high visual complexity (Bonferroni-corrected $p = 0.0004$) while the older adults were 18.6 seconds slower in the low versus high visual complexity (Bonferroni-corrected $p < 0.0001$) and

were15 seconds slower in the medium versus high visual complexity (Bonferroni-corrected $p < 0.0001$).

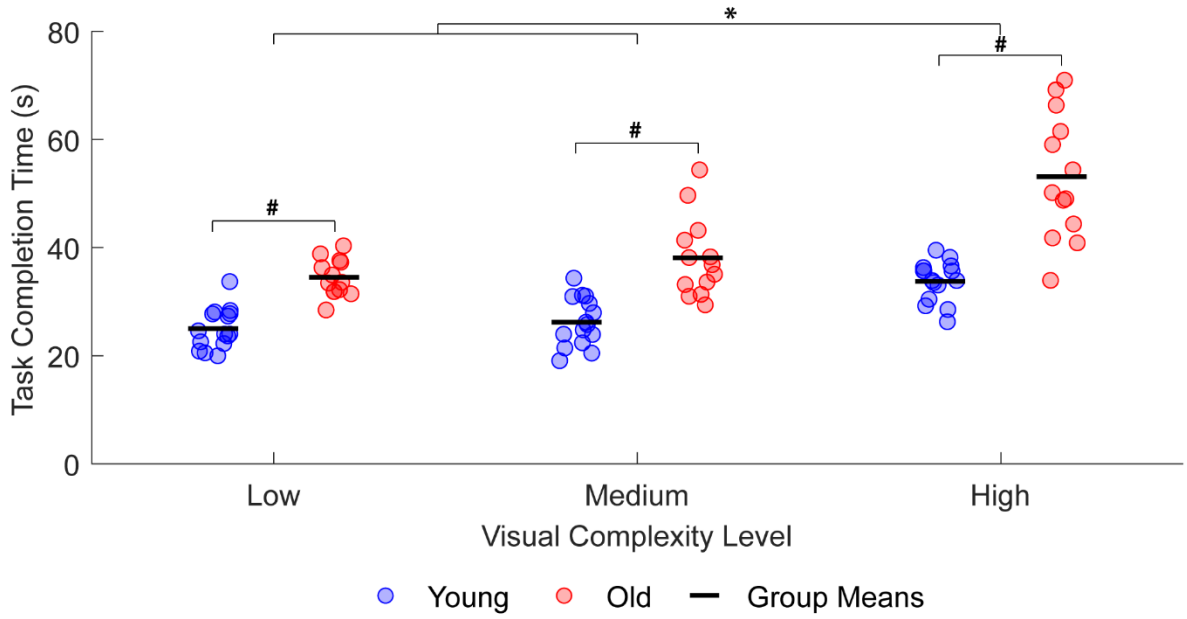

Figure 1. Task completion time across age groups and visual complexity levels. *Indicates significant differences between visual complexity levels. #Indicates significant differences between age groups.

Search times between correct targets. There are several performance-related factors that could explain the changes in task completion time with complexity level and differences in completion time between groups. For example, participants may have taken longer to complete the task in levels with higher visual complexity because they took longer to shift their gaze from one correct target to another. Consistent with this idea, we found main effects of age group ($F(1,26) = 40.08$, $p = 1.06e\text{-}6$) and visual complexity ($F(2,52) = 68.24$, $p = 2.88e\text{-}15$) on the lag between fixations on correct targets (Fig 2a). We also found an interaction between age group and visual complexity ($F(2,52) = 9.655$, $p = 2.72e\text{-}40.0002718$). Compared to young adults, older adults searched for correct targets 0.5 seconds slower in the low complexity level (Bonferroni-corrected $p = 0.0428$), 0.6 seconds slower in the medium complexity level(Bonferroni-corrected $p = 0.0007$), and 1.1 seconds slower in the high visual complexity level (Bonferroni-corrected $p < 0.0001$). In addition, young adults searched for correct targets

0.5 seconds slower in the low versus high complexity level (Bonferroni-corrected $p = 0.0002$) and 0.5 seconds going in the medium versus high complexity level (Bonferroni-corrected p = 0.0004) while older adults searched for correct targets 1.1 seconds slower in the low versus high complexity level (Bonferroni-corrected $p < 0.0001$) and 0.9 seconds slower in the medium versus high complexity level (Bonferroni-corrected $p < 0.0001$).

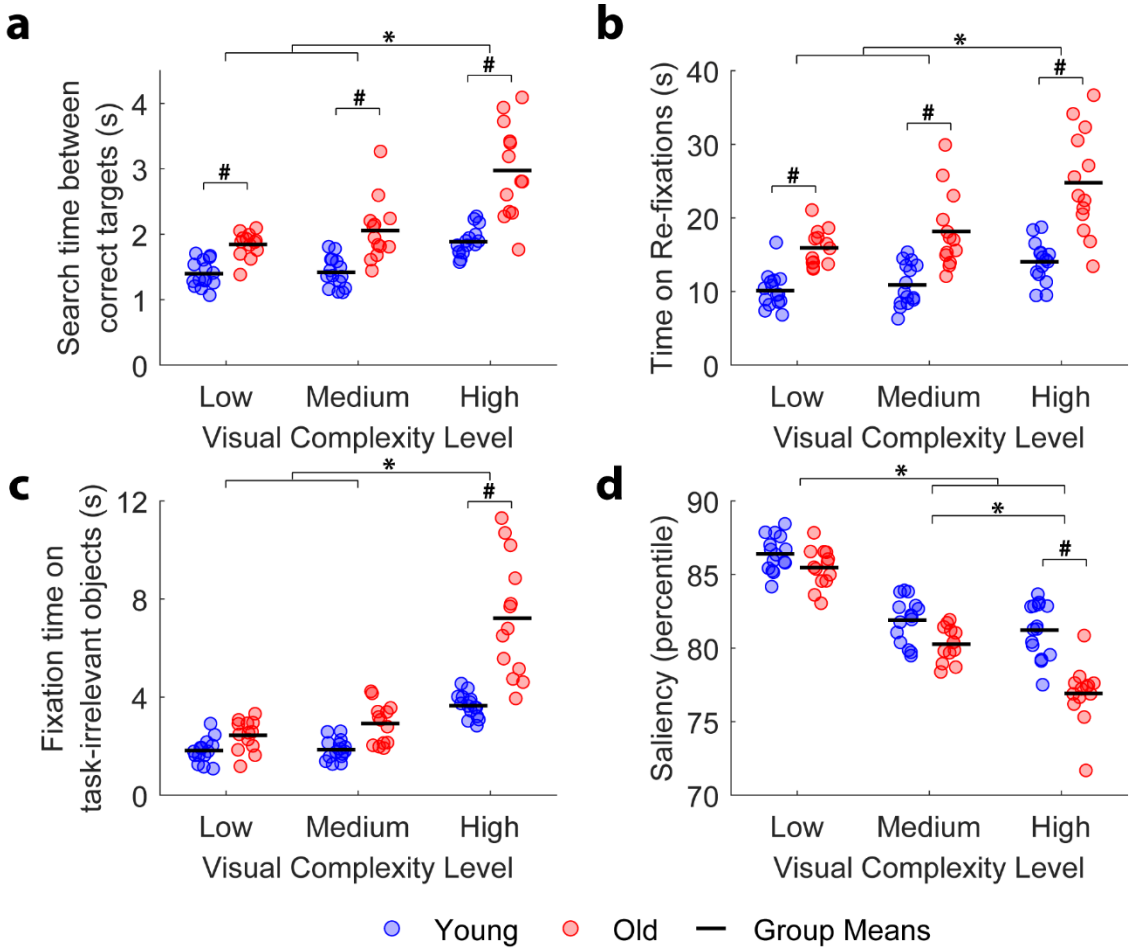

Figure 2. Effect of age group and visual complexity level on gaze behavior. (a) Search time between correct targets, (b) re-fixation time on task-relevant objects, (c) fixation time on task-irrelevant objects, and (d) saliency of fixated regions. *Indicates significant difference between visual complexity levels. #Indicates significant difference between age groups.

Re-fixation time on task-relevant objects. Participants may have also increased task completion time with increasing visual complexity because they spent time re-fixating objects that were already fixated at an earlier time. We separately analyzed how frequently this occurred for task-relevant and task-irrelevant objects. We found main effects of age group

($F(1,26) = 37.16$, $p = 1.92e-6$) and visual complexity ($F(2,52) = 42.37$, $p = 1.21e-11$) on the time spent re-fixating task-relevant objects (Fig. 2b). We also found an interaction effect between age group and visual complexity ($F(2,52) = 6.08$, $p = 0.004223$) on time spent re-fixating task-relevant objects. Compared to young adults, older adults re-fixated task-relevant objects 5.8 seconds longer in the low complexity level (Bonferroni-corrected $p = 0.0073$), 7.3 seconds longer in the medium complexity level (Bonferroni-corrected $p = 0.0003$), and 10.7 seconds longer in the high complexity level (Bonferroni-corrected $p < 0.0001$). In addition, young adults re-fixated task-relevant objects 3.9 seconds longer in the low versus high complexity level (Bonferroni-corrected $p = 0.0034$) and 3.2 seconds longer in the medium versus high complexity level (Bonferroni-corrected $p = 0.0349$) while older adults re-fixated task-relevant objects 8.8 seconds longer in the low versus high complexity level (Bonferroni-corrected $p < 0.0001$) and 6.6 seconds longer in the medium versus high complexity level (Bonferroni-corrected $p < 0.0001$).

Fixation time on task-irrelevant objects. We found main effects of age group ($F(1,26) = 27.96$, $p = 1.58e-5$) and visual complexity ($F(2,52) = 142.34$, $p < 2.2e-16$) on the amount of time fixating task-irrelevant objects (Fig. 2c). We also found an interaction effect between age group and visual complexity level on the amount of time fixating task-irrelevant objects ($F(2,52) = 26.598$, $p = 1.11e-8$). Compared to young adults, older adults fixated task-irrelevant objects 3.6 seconds longer than young adults in the high visual complexity level (Bonferroni-corrected $p < 0.0001$). In addition, young adults fixated task-irrelevant objects 1.8 seconds longer in the low versus high complexity level (Bonferroni-corrected $p < 0.0001$) and 1.8 seconds longer in the medium versus high complexity level (Bonferroni-corrected $p < 0.0001$) while older adults fixated task-irrelevant objects 4.8 seconds longer in the low versus high complexity level (Bonferroni-corrected $p < 0.0001$) and 4.3 seconds longer in the medium versus high complexity level (Bonferroni-corrected $p < 0.0001$).

Saliency of fixated regions. Although both young and older adults fixated on task-irrelevant objects longer as the visual complexity of the environment increased, it does not appear to be caused by increased fixations to salient distractors (Fig. 2d). We found main effects of age group ($F(1,26) = 32.55$, $p = 5.28e\text{-}6$) and visual complexity ($F(2,52) = 186.37$, $p < 2.2e\text{-}16$) on the saliency of fixated regions. We also found an interaction between age group and visual complexity ($F(2,52) = 11.83$, $p = 5.85e\text{-}5$). Young adults only fixated more salient regions than older adults in the highest visual complexity level (Bonferroni-corrected $p < 0.0001$). In addition, young adults fixated more salient regions in the low versus in the medium (Bonferroni-corrected $p < 0.0001$) and high (Bonferroni-corrected $p < 0.0001$) complexity levels while older adults fixated more salient regions in the low versus in the medium (Bonferroni-corrected $p < 0.0001$) and high (Bonferroni-corrected $p < 0.0001$) complexity levels as well as in the medium versus the high (Bonferroni-corrected $p < 0.0001$) complexity level.

Number of selection errors. Participants may have also taken longer to complete the task with increasing visual complexity due to changes in their target selection strategy. To this end, we quantified target selection errors, defined as selecting an incorrect target or misaiming (laser pointed off-center) when selecting a correct target (Fig. 3a). We found a main effect of age group on the number of selection errors ($F(1,26) = 6.36$, $p = 0.01810$), with older adults committing more selection errors than younger adults. We also found a main effect of visual complexity ($F(2,52) = 3.24$, $p = 0.04727$) on the number of selection errors; however, post-hoc tests found no differences between visual complexity levels. We did not find an interaction effect between age group and visual complexity ($F(2,52) = 1.77$, $p = 0.18055$) on the number of selection errors.

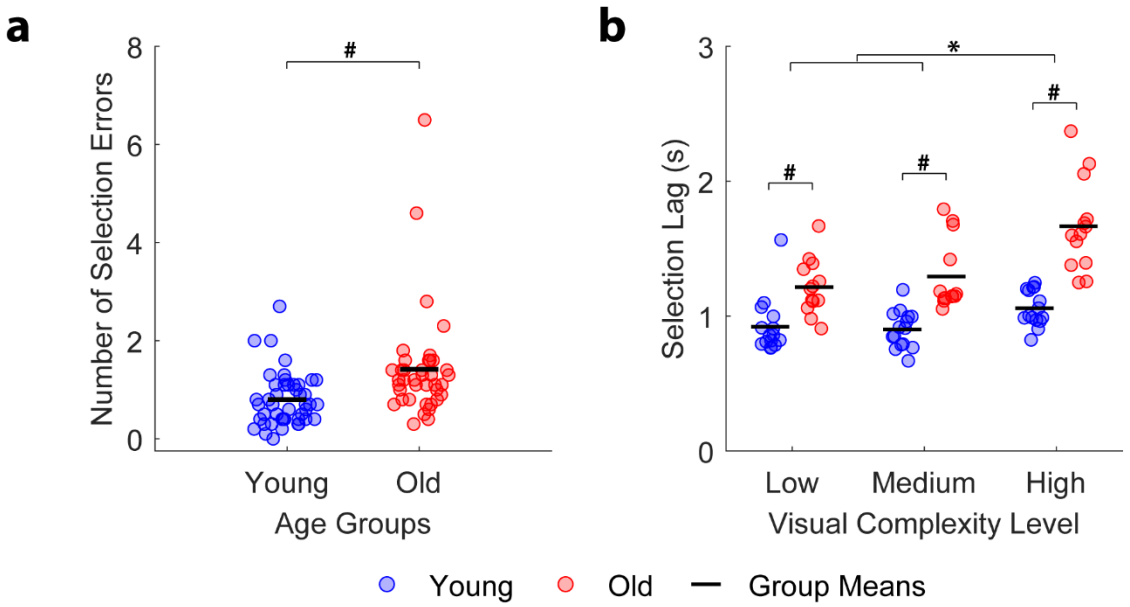

Figure 3. Effect of age group and visual complexity level on motor behavior. (a) The number of selection errors and (b) the time between fixating and selecting a correct target. *Indicates significant difference between visual complexity levels. #Indicates significant difference between age groups.

Selection lag. Another way to investigate possible influences of changes in target selectionmotor function on task completion time is to quantify selection lag, defined as the time between fixating a correct target and successfully selecting it (Fig. 3b). We found main effects of age group ($F(1,26) = 38.92$, $p = 1.34e\text{-}6$) and visual complexity ($F(2,52) = 31.70$, $p = 9.98e\text{-}10$) on selection lag. We also found an interaction effect between age group and visual complexity ($F(2,52) = 7.81$, $p = 0.001083$) on selection lag. Compared to young adults, older adults selected the correct target 0.3 seconds slower in the low complexity level (Bonferroni-corrected $p = 0.0143$), 0.4 seconds slower in the medium complexity level (Bonferroni-corrected $p = 0.0003$), and 0.6 seconds slower in the high complexity level (Bonferroni-corrected $p < 0.0001$). In addition, older adults selected the correct target 0.5 seconds slower in the low versus high complexity level (Bonferroni-corrected $p < 0.0001$) and 0.4 seconds slower in the medium versus high complexity level (Bonferroni-corrected $p < 0.0001$).

Cognitive assessment scores. Age-related differences were observed in several standard cognitive assessments (Table 1). In particular, older adults had lower global cognition according

to their Montreal Cognitive Assessment scores (Wilcoxon rank sum test p = 0.0440) and poorer short-term memory according to their Corsi Block task results (Wilcoxon rank sum test p = 0.0014). However, no differences were found in executive function, task switching, or inhibition.

Cognitive Assessment	**Cognitive Domain**	**Young Adults (Mean ± SD)**	**Older Adults (Mean ± SD)**	**P-value**
Montreal Cognitive Assessment	Global Cognition	28 ± 2	27 ± 3	0.0440 *
Trail Making Test-B	Executive Function, Task-Switching	63 ± 20 s	65 ± 23 s	0.8012
Corsi Block Test (Forwards)	Short Term Memory	6 ± 1	5 ± 1	0.0014 *
Corsi Block Test (Backwards)	Working Memory	5 ± 1	5 ± 2	0.229
Stroop Test	Inhibition	58.87 ± 110.17 ms	110.69 ± 144.78 ms	0.293
Flanker Test	Inhibition	5 ± 70.17 ms	-10.77 ± 95.82 ms	0.6203

Table 1. Cognitive assessment scores (Mean ± SD) of young and older adults with the corresponding p-value from their comparisons using two-sample t-tests and Wilcoxon rank sum tests. * Indicates significant differences between age groups.

Associations between task completion time and cognition. Longer task completion times were associated with decreased short-term and working memory capacity (Fig. 4). A multiple linear regression model was used to test if short-term memory (Corsi Block task), working memory (Backwards Corsi Block task), and inhibitory capacity (Stroop and Flanker tasks) were associated with task completion time. The final regression model included the Corsi Block task and Backward Corsi Block task (Adjusted R^2 = 0.48, $F(4,23)$ = 7.219, p = 6.42e-4) with higher scores on the Corsi Block task (Fig. 3a; β = -4.27, p = 0.0299) and Backwards Corsi Block task (Fig 3b; β = -3.46, p = 0.0221) being associated with faster task completion time.

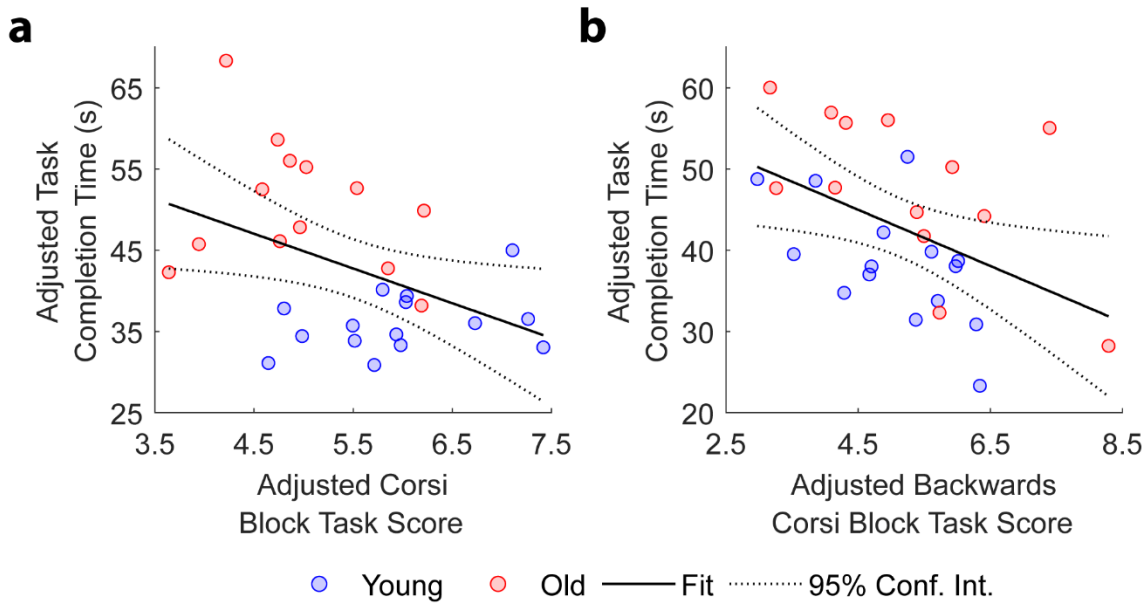

Figure 4. Added variable plots showing relationships between task completion time and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task, and (b) working memory capacity assessed using the Backwards Corsi Block task.

Re-fixation time and cognition. Longer re-fixation times on task-relevant objects were associated with decreased short-term and working memory capacity across all participants (Fig. 5). A multiple linear regression model was fit to determine if short-term (Corsi Block task) and working (Backwards Corsi Block task) memory capacities were associated with re-fixation time on task-relevant objects. The final regression model included both assessments of memory (Adjusted R^2 = 0.45, $F(2,25)$ = 12.07, p = 2.15e-4) with higher scores on the Corsi Block task (Fig. 4a; β = -2.36, p = 0.0396) and Backwards Corsi Block task (Fig. 4b; β = -2.11, p = 0.0189) being associated with shorter re-fixation times on task-relevant objects.

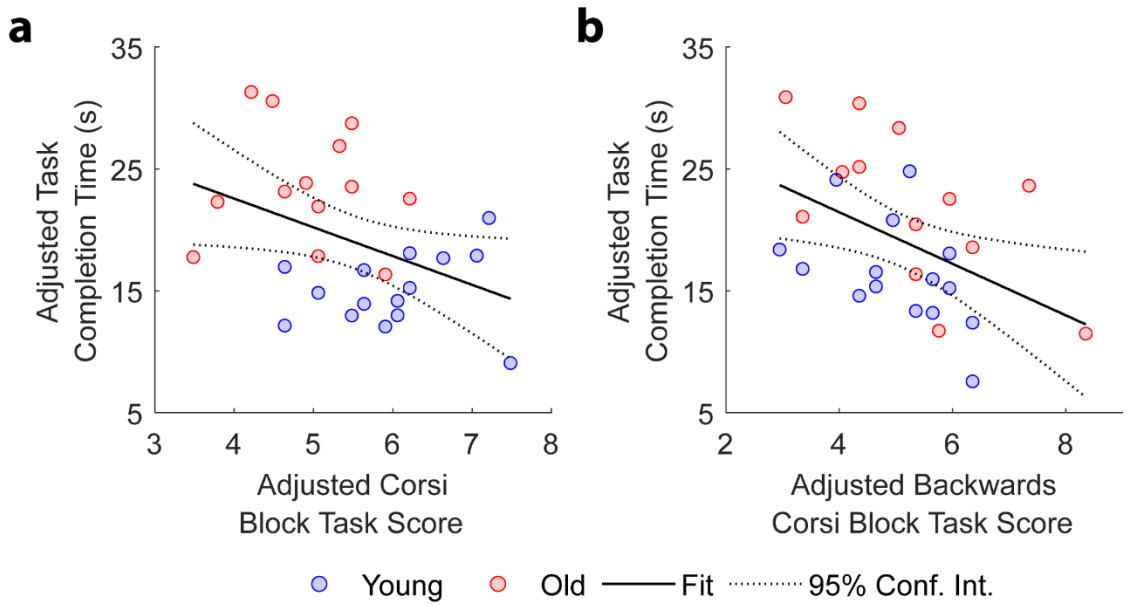

Figure 5. Added variable plots showing relationships between re-fixation time on task-relevant objects and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task and (b) working memory capacity assessed using the Backwards Corsi Block task.

Discussion

We used a VR-based visual search task to determine if age-dependent differences in the control of visual attention are observed in environments that simulate real-world visual complexity. In addition, we also sought to understand which cognitive domains were associated with performance in the task. For both age groups, increasing the visual complexity of the virtual environments led to longer task completion times. This was due, in part, to the fact that all participants took more time to search for correct targets as they re-fixated task-relevant and fixated task-irrelevant objects longer. We also found that all participants took more time to select a correct target after fixating them in levels with higher visual complexity, which likely contributed to the longer task completion times. Each of these effects was greater in older than younger adults. However, increased completion time did not appear to be influenced by saliency in both groups, as the saliency of the fixated regions decreased as the visual complexity level increased. Lastly, we found that the variability in performance across participants was explained by short-term and working memory capacities.

As the complexity of the virtual environments increased, all participants spent more time completing the search task and this was because they took longer to move their gaze between correct targets. These results are consistent with those classically found in simple conjunction paradigms with simple stimuli composed of arbitrarily selected shapes and colors. In particular, when participants were asked to search for stimuli that shared features with distractors, search times increased as the number of distractors in the visual display increased[28]. We also found that the effect of visual complexity on search performance was greater in older adults than young adults. Similar age-related differences in search performance have been demonstrated using simple conjunction search paradigms[29–31] and in paradigms that used more complex stimuli. For example, when participants were tasked with searching for and responding to a star-shaped stimulus in a vehicle's digital dashboard, search times increased as the visual complexity of the display increased, with the effect being greater in older adults[32]. Similarly, when participants searched for a specific face configuration among distractor faces, search times increased with the number of distractors, with the effect being similarly greater in older adults[33]. As such, it can be said that age-related changes in cognition, visual complexity, and their combination on search performance generalizes from two-dimensional displays to more complex, three-dimensional virtual environments.

Another factor contributing to the increase in completion time in all participants was the increase in fixation times on task-irrelevant objects This increase in fixation time on task-irrelevant objects could be due to a decreased capacity for top-down suppression of irrelevant stimuli with increasing visual complexity of search displays[34], making all participants more susceptible to fixating task-irrelevant objects. Additionally, the effect of visual complexity on fixation times on task-irrelevant objects was greater in older adults. Increased susceptibility to distraction by task-irrelevant information in older adults has been attributed to cognitive impairments, particularly with deficiencies in inhibitory capacity[13,14]. This increased distractibility has been previously shown across a range of two-dimensional visual-perceptual tasks[16–19,35].

Together, these findings support the notion that visual attention, and by extension, performance in a visual search task, is affected not only by the visual complexity of the environment but also by aging. Interestingly, while fixation times on task-irrelevant objects increased with visual complexity, these times only accounted for a small portion of the task completion time. Thus, visual attention was primarily allocated toward task-relevant stimuli in our task.

While all participants became more prone to distraction by task-irrelevant stimuli as visual complexity increased, this effect was not driven by the saliency of these stimuli as the saliency of fixated regions decreased as the visual complexity of the virtual environments increased. This finding is inconsistent with previous studies showing that the susceptibility to distraction is greater in the presence of salient visual stimuli, particularly in older adults. In an onset distractor task, while young and older adults were similarly distracted when the search targets and the distractors had the same luminance, older adults became more susceptible to distraction when the distractors were more salient by making them brighter[16]. Similarly, older adults were more distracted when distractors' saliency increased in selective attention tasks with task-relevant and task-irrelevant stimuli superimposed on each other[18,35]. The ability of visual stimuli to automatically capture visual attention may not be contingent upon the stimuli being highly salient. Instead, people may fixate more on task-irrelevant stimuli because greater demands are placed on information processing and memory as the complexity of the environment increases. Future studies should consider manipulating the saliency of the task-relevant stimuli and task-irrelevant distractors in the search environments to better determine the role of saliency in visual information processing.

Proper allocation of visual attention is important for successfully completing visual search tasks, and it appears to rely on intact cognition. Performance in our task was related to participants' short-term and working memory capacity, as those who scored higher in the Corsi block task and the Backwards Corsi block task demonstrated faster task completion times. In addition, we found a relationship between re-fixation times on task-relevant objects and both

short-term and working memory. Participants with shorter re-fixation times on task-relevant objects also scored higher in the Corsi and Backwards Corsi block tasks. These findings echo the importance of short-term and working memory capacity on visual attention[36–38], which has been demonstrated in simple selective attention[39] and visual search[33] tasks. Beyond memory, inhibitory capacity is another cognitive domain that influences the allocation of visual attention and the successful completion of visual search tasks. As previously discussed, inhibitory capacity allows for the proper allocation of visual attention by selecting for and processing information relevant to the task while suppressing those that are irrelevant. Scores on cognitive assessments that specifically targeted inhibition were not found to be associated with task completion times, which is consistent with the fact that the time participants spent fixating task-irrelevant objects only accounted for a small portion of the task completion time.

The increase in task completion times at higher visual complexity levels can be explained not only by the time required to search for the correct targets, as exhibited using the calculated fixation times, but also potentially by changes that influence the speed with which targets are selected, such as by moving slower. We quantified selection lag as the time between fixating and selecting a correct target which is influenced by a combination of the time it takes to move the laser pointer to the correct target and successfully select it. All participants demonstrated longer selection lag as the visual complexity of the search environments increased, with this effect being greater in older adults. However, target selection seems to only have a minor impact on the observed increase in task completion times. While selection lag increased with the visual complexity of the virtual environments for all participants, its magnitude was less than the various calculated fixation times. Interestingly, the number of selection errors was unaffected by increasing visual complexity for all participants. In conjunction with the changes in selection lag, it appears that all participants traded off moving towards and selecting the correct targets faster to maintain accuracy at higher visual complexity levels, consistent with the well-established speed-accuracy trade-off phenomenon[40–43]. However, the selection errors

were greater in older adults than young adults, contradicting previously demonstrated parity between the two groups resulting from older adults using movement slowing as a compensatory mechanism to maintain accuracy[44–46].

In conclusion, susceptibility to distraction by task-irrelevant information during visual search and target selection increases with the environment's complexity. This negative effect was greater in older adults, likely driven by age-related changes in short-term and working memory. Increased susceptibility to distraction is potentially problematic as suboptimal allocation of visual attention may lead to difficulties completing everyday tasks and potentially can lead to injuries, such as when driving or navigating complex environments. Our results highlight the importance of designing cognitive assessments that use tasks that closely simulate the visual complexity and dynamicity in the real world. In addition, our study also emphasizes the need for a more comprehensive analysis of task performance beyond completion time to accurately assess cognitive capacity, particularly in older adults.

Methods

Participants. 15 young (9 female, age: 27.7 ± 3.33 years) and 15 older adults (9 female, age: 71.8 ± 4.46 years) with normal or corrected-to-normal vision participated in the study. Data from two older adults were excluded due to hardware/software issues. We conducted a sample size calculation for mixed-effect regression models using GLIMMPSE[47] and used pilot to specify the estimated variance in our dependent (task completion time, fixation time on task-relevant targets, and target saliency) and independent variables (visual complexity and age group). We found the largest sample size estimate across our three dependent variables and determined that a sample size of 26 (13 for each age group) was needed for a target power of 0.80 assuming a medium effect size of 0.45. All study procedures were reviewed and approved by the University of Southern California's Institutional Review Board. All participants provided written informed consent before participating and were provided monetary compensation for

their time. All aspects of the study conformed to the principles described in the Declaration of Helsinki.

Experimental protocol. After informed consent was obtained, all participants completed a battery of tests to assess different domains of cognition. Two of these assessments, the Montreal Cognitive Assessment[48] and the Trail Making Test-B[21], were administered on paper and were used to assess global cognition and executive function, respectively while the rest were provided on a computer through Psytoolkit[49,50]. Of these computer-based assessments, the Corsi task[51] was used to assess short-term memory, the backward Corsi task[52] for working memory, and the Stroop[53] and Flanker[54] tasks for inhibition.

After completing the cognitive assessments, the participants performed three familiarization trials for the visual search task. The first trial consisted of the experimenter guiding the participant through a single round of the task to provide a clear description of the goal of the task, the visual and auditory feedback that they will encounter, and the search target sequence. The guided familiarization trial was followed by two more trials where participants completed the task independently and were only provided feedback by the experimenter at the end. After the familiarization trials, the participants completed 30 trials of the visual search task presented in three sets of 10 trials, each corresponding to one of three visual complexity levels.

VR task. We based our VR visual search task (Fig. 6a-c) on the principles of the Trail Making Test-B[21]. We selected the Trail Making Test-B based on its ability to test various cognitive domains, including visual attention, working memory, and inhibition[22–24]. Participants were instructed to search for and select targets alternating between letters of the alphabet and objects whose names start with those letters in ascending order (A, Ant, B, Butterfly, C, Chickens, D, Dog, E, Eagle, F, and Frogs). The search targets were positioned randomly for each trial in regions of the virtual environments that will make them less salient and allow them to blend in, especially at higher visual complexity levels, such as across the sky for the letters and at ground level for the objects. The participants viewed the virtual environments using the

HTC Vive Pro Eye (HTC, New Taipei, Taiwan) head-mounted display and selected targets using an HTC Vive Controller. The head-mounted display had a 110-degree field of view with a resolution of 1440x1600 pixels per eye (2880x1600 pixels combined). A laser pointer extended from the top of the controller in the virtual environment, and participants selected targets by aiming the laser at the center of the target and pulling the trigger on the controller. All participants were informed that they could freely move their heads in all directions when completing the search task. Visual and auditory feedback were provided to indicate whether the selected target was correct. Specifically, a bell sound was played, and a “Correct!” text appeared in the environment when the correct target was selected. In contrast, a buzzer sound was played, and a “Try Again!” text appeared when the selected target was incorrect or when the correct target was not selected properly (laser not aimed at the center of the target).

The visual search task was performed in virtual environments with three levels of increasing visual complexity. The sequence of complexity levels experienced by participants was pseudo-randomized such that five participants in both young and older adult groups started with the low visual complexity level (Fig. 6a), five participants started with the medium visual complexity level (Fig. 6b), and the remaining five started with the high visual complexity level (Fig. 6c). The visual complexity of the environment was manipulated by increasing the number of visual distractors in the foreground and background and was measured using feature congestion[55]. Feature congestion quantifies the distribution of low-level visual features of an image, which includes color, edge orientation, and luminance, as a single scalar measure with higher values indicating a more complex image. To determine if the modifications to the virtual environments produced increasing levels of visual complexity, we measured the feature congestion method in every frame of representative gameplay recording from each visual complexity level (using Rosenholtz et al.’s Matlab implementation: https://dspace.mit.edu/handle/1721.1/37593). Feature congestion increased monotonically across visual complexity levels (Fig. 6d; $F(2,1203.7) = 28690$, $p < 2.2e\text{-}16$). Post-hoc tests

indicate that the high visual complexity level had greater feature congestion than the medium (Bonferroni-corrected $p < 2e\text{-}16$) and low (Bonferroni-corrected $p < 2e\text{-}16$) visual complexity levels. In addition, feature congestion was higher in the medium visual complexity level than in the low ($p < 2e\text{-}16$).

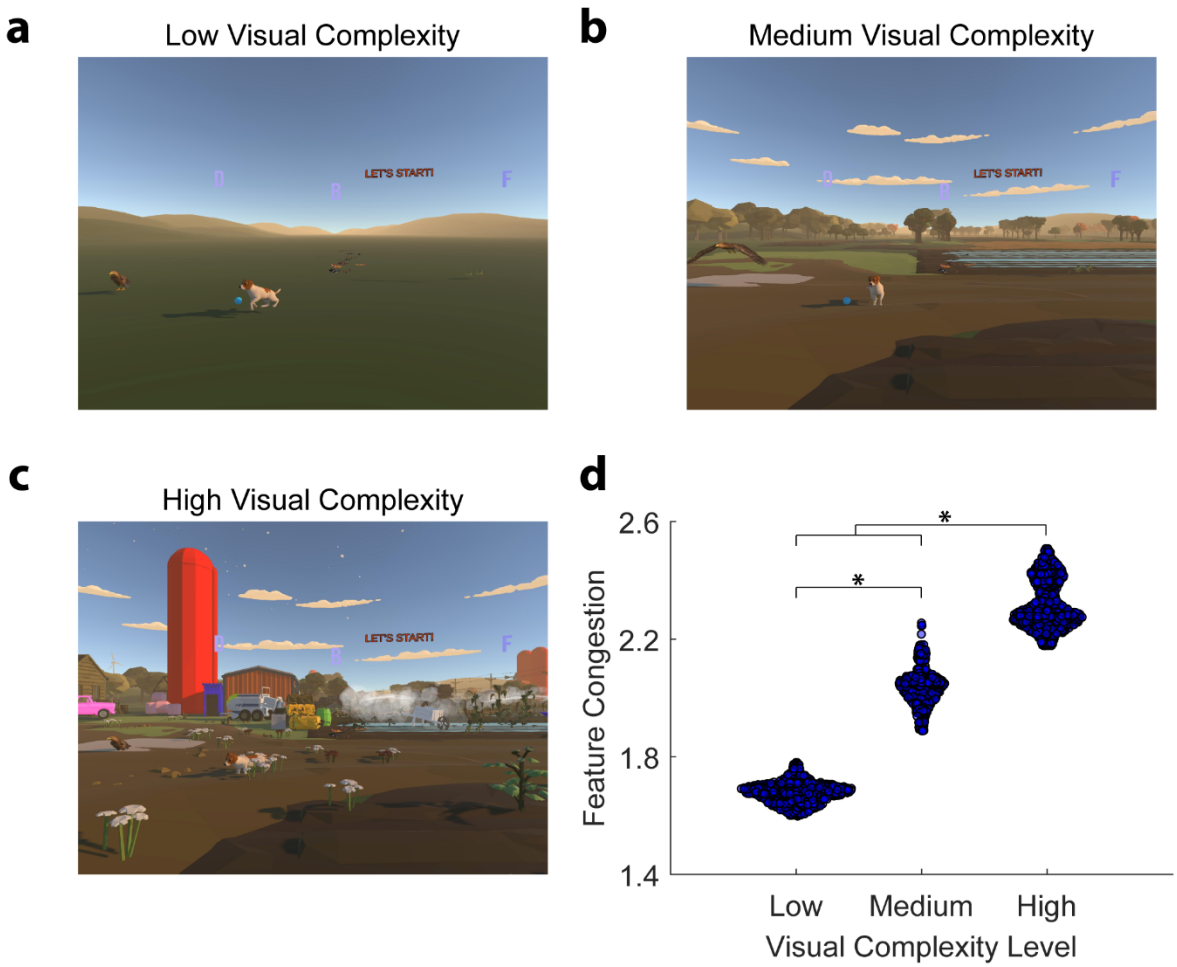

Figure 6. Variation in feature congestion across the three visual complexity levels used in the visual search task. (a-c) Single frames from representative gameplay recordings of each visual complexity level. (d) Scatter plot comparing feature congestion of each visual complexity level with each data point representing a frame from representative gameplay recordings. *Indicates significant differences between visual complexity levels.

Data collection and processing. We collected eye and head movement data at 90 Hz throughout each trial using eye trackers and an inertial measurement unit built into the HTC Vive Pro Eye. The eye trackers were calibrated following a 5-point calibration process provided by HTC. Additionally, video recordings of the first-person point of view of the participants were

also recorded at 60 Hz for each trial using NVIDIA Shadowplay. All data processing was performed using a custom script in MATLAB R2022a (Mathworks, Natick, MA). Horizontal and vertical eye-in-head angles were calculated from the raw eye movement data and were combined with the horizontal and vertical rotations of the head to compute gaze angles. The gaze angles were then time-synchronized with the video recordings. Since the video recordings and the gaze angles did not have the same sampling rate, the gaze angles were then interpolated at each video recording frame.

As visual processing only occurs during periods of fixation[56,57], it was important to identify which samples qualified as such to determine how visual attention was deployed throughout each trial. To do so, we first computed horizontal and vertical gaze angular velocities by differentiating the interpolated gaze angles with respect to the time between video frames. We then categorized a sample as a fixation if its horizontal or vertical gaze angular velocity was less than 100 degrees per second[58]. Fixations were then classified as task-relevant or task-irrelevant based on the identity of the object being fixated. We computed gaze vectors in the virtual environment with their origins based on head position in the virtual environment and their direction determined by the gaze angles. If the gaze vectors intersected a search target, the name of the fixated search target was recorded and identified as a task-relevant fixation. In contrast, if the gaze vector did not intersect a search target, it was identified as a task-irrelevant fixation. We calculated the search times between correct targets by averaging the time interval between fixations on correct targets in a trial. We calculated re-fixation times on task-relevant objects by determining which fixations were on task-relevant targets that participants previously fixated and then calculating the sum of those fixation times within a trial. We also calculated the fixation times on task-irrelevant objects by determining which fixations were on task-irrelevant objects or regions and then calculating the sum of those fixation times in a trial. Search times between correct targets, re-fixation times on task-relevant objects, and the fixation times on

task-irrelevant objects were calculated for every trial and reported as an average of all trials at each visual complexity level for each participant.

Saliency maps were created from each frame of every recording based on the original Itti, Koch, and Neibur[59] algorithm as implemented by Harel (J. Harel, A Saliency Implementation in MATLAB: http://vision.caltech.edu/~harel/share/gbvs.php). Each video frame was decomposed into feature maps of its low-level visual features, which included color, orientation, and intensity. These feature maps were then combined into a saliency map with corresponding scalar values for each pixel. These scalar saliency values were then converted to percentile ranks[60,61], with 100% indicating the most salient pixel while 0% the least salient pixel in the video frame. The saliency of the fixated regions were then determined for every trial and reported as an average of all trials at each visual complexity level for each participant.

At every trigger press, we recorded features of selected objects which included the identity of the selected object, the location where the laser pointer was being aimed in the virtual environment at the time of the trigger press, and if the selection was correct. The number of selection errors, defined as selecting an incorrect target or misaiming (controller aimed outside the bounds of the correct target) when selecting a correct target, and selection lag, which is the time between fixating a correct target and successfully selecting it, were calculated from the object selection data. The number of selection errors and selection lag were calculated for every trial and reported as an average of all trials at each visual complexity level for each participant.

To quantify each participant's overall performance, task completion time was calculated at every trial, defined as the difference between the time when the trial started and when the last correct target was selected. This measure of overall task performance is equivalent to the primary output of the paper-based versions of the Trail Making Test. Task completion time was calculated for every trial and reported as an average of all trials at each visual complexity level for each participant.

Statistical analysis. All statistical analyses were performed in R (R Project for Statistical Computing) with the alpha value set at $p < 0.05$. Tests for normality and equal variances were performed using the Shapiro-Wilk test and Levene's test with functions from the *stats* and *car* packages, respectively. Welch's analysis of variance and pairwise Wilcoxon rank sum tests were used to compare the feature congestion of each visual complexity level using the *stats* package as assumptions of equal variance and normality were violated.

We fit linear mixed-effects models using the *lme4* package to test for the effects of age group, visual complexity, and their interaction on the following dependent variables: VR task completion time, search times between correct targets, re-fixation time on task-relevant objects, fixation time on task-irrelevant objects, saliency of fixated regions, number of selection errors, and selection lag. We used the *lmerTest* package, which uses Satterthwaite approximations for the degrees of freedom, to test the null hypothesis that our model coefficients were zero. All models included random intercepts for each participant to account for repeated measures. Lastly, we performed post-hoc Bonferroni-corrected pairwise comparisons using the *emmeans* package when significant main effects or interactions, particularly differences between age groups within visual complexity levels, were found.

We compared cognitive assessment scores between age groups to determine the effect of age on various domains of cognition using functions from the *stats* package. Assumptions of normality and equal variance were tested using the Shapiro-Wilk and Levene's tests, respectively. The Wilcoxon rank sum test was used when the normality assumption was violated; otherwise, a two-sample t-test was used.

To test for associations between cognition and task performance measures, we used multiple linear regression. In particular, task completion time was used as the response, while scores on the Corsi Block task, Backwards Corsi Block task, Stroop task, and Flanker task were used as predictors. A separate multiple linear regression was used to test for associations between re-fixation times on task-relevant objects and measures of memory capacity.

Specifically, re-fixation time on task-relevant objects was used as the response, while scores on the Corsi Block task and the Backwards Corsi Block task were used as predictors.

Data Availability

The analyzed data from this study are available on the Open Science Framework at https://osf.io/r9q3k/.

References

1. Land, M. F. & Hayhoe, M. In what ways do eye movements contribute to everyday activities? *Vision Res.* **41**, 3559–3565 (2001).

2. Land, M. F. Eye movements and the control of actions in everyday life. *Prog. Retin. Eye Res.* **25**, 296–324 (2006).

3. Hayhoe, M. & Ballard, D. Eye movements in natural behavior. *Trends Cogn. Neurosci.* **9**, 188–194 (2005).

4. Hayhoe, M. M. & Rothkopf, C. A. Vision in the natural world. *Wiley Interdiscip. Rev. Cogn. Sci.* **2**, 158–166 (2011).

5. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203 (2001).

6. Tatler, B. W., Hayhoe, M. M., Land, M. F. & Ballard, D. H. Eye guidance in natural vision: Reinterpreting salience. *J. Vis.* **11**, (2011).

7. Cosman, J. D., Lowe, K. A., Zinke, W., Woodman, G. F. & Schall, J. D. Prefrontal Control of Visual Distraction. *Curr. Biol.* **28**, 414-420.e3 (2018).

8. Gaspelin, N. & Luck, S. J. The Role of Inhibition in Avoiding Distraction by Salient Stimuli. *Trends Cogn. Sci.* **22**, 79–92 (2018).

9. Gaspelin, N. & Luck, S. J. Inhibition as a potential resolution to the attentional capture debate. *Curr. Opin. Psychol.* **29**, 12–18 (2019).

10. Seidl, K. N., Peelen, M. V. & Kastner, S. Neural Evidence for Distracter Suppression during Visual Search in Real-World Scenes. *J. Neurosci.* **32**, 11812–11819 (2012).

11. Mather, M. Aging and cognition. *WIREs Cogn. Sci.* **1**, 346–362 (2010).

12. Harada, C. N., Natelson Love, M. C. & Triebel, K. Normal Cognitive Aging. *Clin. Geriatr. Med.* **29**, 737–752 (2013).

13. Dempster, F. N. The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Dev. Rev.* **12**, 45–75 (1992).

14. Lustig, C., Hasher, L. & Zacks, R. T. Inhibitory deficit theory: Recent developments in a 'new view'. in *Inhibition in cognition* (eds. Gorfein, D. S. & Macleod, C. M.) 145–162 (American Psychological Association, 2007).

15. Gazzaley, A. & Mark D'esposito, A. Top-Down Modulation and Normal Aging. *Ann. N. Y. Acad. Sci.* **1097**, 67–83 (2007).

16. Kramer, A. F., Hahn, S., Irwin, D. E. & Theeuwes, J. Age differences in the control of looking behavior: Do You Know Where Your Eyes Have Been? *Psychol. Sci.* **11**, 210–217 (2000).

17. Schmitz, T. W., Cheng, F. H. T. & De Rosa, E. Failing to ignore: paradoxical neural effects of perceptual load on early attentional selection in normal aging. *J. Neurosci. Off. J. Soc. Neurosci.* **30**, 14750–14758 (2010).

18. Schmitz, T. W., Dixon, M. L., Anderson, A. K. & De Rosa, E. Distinguishing attentional gain and tuning in young and older adults. *Neurobiol. Aging* **35**, 2514–2525 (2014).

19. Mertes, C., Wascher, E. & Schneider, D. Compliance instead of flexibility? On age-related differences in cognitive control during visual search. *Neurobiol. Aging* **53**, 169–180 (2017).

20. Gazzaley, A. *et al.* Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl. Acad. Sci.* **105**, 13122–13126 (2008).

21. Reitan, R. M. Validity of the Trail Making Test as an indicator of organic brain damage. *Percept. Mot. Skills* **8**, 271 (1958).

22. Sánchez-Cubillo, I. *et al.* Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J. Int. Neuropsychol. Soc.* **15**, 438–450 (2009).

23. Salthouse, T. A. What cognitive abilities are involved in trail-making performance? *Intelligence* **39**, 222 (2011).

24. Llinàs-Reglà, J. *et al.* The Trail Making Test. *Assessment* **24**, 183–196 (2017).

25. Plotnik, M. *et al.* Multimodal immersive trail making-virtual reality paradigm to study cognitive-motor interactions. *J. NeuroEngineering Rehabil.* **18**, 1–16 (2021).

26. Goldberg, M. E. & Wurtz, R. H. Visual Processing and Action. in *Principles of Neural Science* (eds. Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A. & Hudspeth, A. J.) 638–653 (McGraw-Hill, 2013).

27. Adlakha, G. *et al.* Development of a virtual reality assessment of visuospatial function and oculomotor control. in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* 753–754 (Institute of Electrical and Electronics Engineers Inc., 2021). doi:10.1109/VRW52623.2021.00259.

28. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cognit. Psychol.* **12**, 97–136 (1980).

29. Folk, C. L. & Lincourt, A. E. The effects of age on guided conjunction search. *Exp. Aging Res.* **22**, 99–118 (1996).

30. Hommel, B., Li, K. Z. H. & Li, S. C. Visual search across the life span. *Dev. Psychol.* **40**, 545–558 (2004).

31. Müller-Oehring, E. M., Schulte, T., Rohlfing, T., Pfefferbaum, A. & Sullivan, E. V. Visual Search and the Aging Brain: Discerning the Effects of Age-related Brain Volume Shrinkage on Alertness, Feature Binding, and Attentional Control. *Neuropsychology* **27**, 48 (2013).

32. Lee, S. C., Kim, Y. W. & Ji, Y. G. Effects of visual complexity of in-vehicle information display: Age-related differences in visual search task in the driving context. *Appl. Ergon.* **81**, 102888 (2019).

33. Aziz, J. R., Good, S. R., Klein, R. M. & Eskes, G. A. Role of aging and working memory in performance on a naturalistic visual search task. *Cortex* **136**, 28–40 (2021).

34. Bonacci, L. M., Bressler, S., Kwasa, J. A. C., Noyce, A. L. & Shinn-Cunningham, B. G. Effects of Visual Scene Complexity on Neural Signatures of Spatial Attention. *Front. Hum. Neurosci.* **14**, 91 (2020).

35. Tsvetanov, K. A., Mevorach, C., Allen, H. & Humphreys, G. W. Age-related differences in selection by visual saliency. *Atten. Percept. Psychophys.* **75**, 1382–1394 (2013).

36. Hayhoe, M. M. Vision and Action. *Annu. Rev. Vis. Sci.* **3**, 389–413 (2017).

37. Madden, D. J. Aging and Visual Attention. *Curr. Dir. Psychol. Sci.* **16**, 70 (2007).

38. Diamond, A. Executive Functions. *Annu. Rev. Psychol.* **64**, 135–168 (2013).

39. Shipstead, Z., Harrison, T. L. & Engle, R. W. Working memory capacity and visual attention: Top-down and bottom-up guidance. *Q. J. Exp. Psychol.* **65**, 401–407 (2012).

40. Fitts, P. M. COGNITIVE ASPECTS OF INFORMATION PROCESSING: III. SET FOR SPEED VERSUS ACCURACY x. *J. Exp. Psychol.* **71**, 849–857 (1966).

41. Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S. & Quinn, J. T. Motor-Output Variability: A Theory for the Accuracy of Rapid Motor Acts. *Psychol. Rev.* **86**, (1979).

42. Schmidt, R., Lee, T., Winstein, C., Wulf, G. & Zelaznkik, H. Principles of Speed and Accuracy. in *Motor Control and Learning: A Behavioral Emphasis* 213–248 (Human Kinetics, 2019).

43. Heitz, R. P. The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Front. Neurosci.* **8**, (2014).

44. Zhang, L., Yang, J., Inai, Y., Huang, Q. & Wu, J. Effects of aging on pointing movements under restricted visual feedback conditions. *Hum. Mov. Sci.* **40**, 1–13 (2015).

45. Lamb, D. G. *et al.* The aging brain: Movement speed and spatial control. *Brain Cogn.* **109**, 105–111 (2016).

46. Norheim, K. L., Samani, A. & Madeleine, P. The effects of age on response time, accuracy, and shoulder/arm kinematics during hammering. *Appl. Ergon.* **90**, 103157 (2021).

47. Kreidler, S. M. *et al.* GLIMMPSE: Online Power Computation for Linear Models with and without a Baseline Covariate. *J. Stat. Softw.* **54**, i10 (2013).

48. Nasreddine, Z. S. *et al.* The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).

49. Stoet, G. PsyToolkit: A software package for programming psychological experiments using Linux. *Behav. Res. Methods* **42**, 1096–1104 (2010).

50. Stoet, G. PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teach. Psychol.* **44**, 24–31 (2017).

51. Corsi, P. M. Memory and the medial temporal region of the brain. (McGill University, 1972).

52. Isaacs, E. B. & Vargha-Khadem, F. Differential course of development of spatial and verbal memory span: A normative study. *Br. J. Dev. Psychol.* **7**, 377–380 (1989).

53. Stroop, J. R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643–662 (1935).

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

55. Rosenholtz, R., Li, Y. & Nakano, L. Measuring visual clutter. *J. Vis.* **7**, 17–17 (2007).

56. Matin, E. Saccadic suppression: a review and an analysis. *Psychol. Bull.* **81**, 899–917 (1974).

57. Rayner, K. Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol. 2006* **62**, 1457–1506 (2009).

58. Salvucci, D. D. & Goldberg, J. H. Identifying Fixations and Saccades in Eye-Tracking Protocols. in *Proceedings of the EyeTracking Research and Applications Symposium* 71–78 (Association for Computing Machinery, 2000). doi:https://doi.org/10.1145/355017.355028.

59. Itti, L., Koch, C. & Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 1254–1259 (1998).

60. Le Meur, O. & Baccino, T. Methods for comparing scanpaths and saliency maps: Strengths and weaknesses. *Behav. Res. Methods* **45**, 251–266 (2013).

61. Kretch, K. S. & Adolph, K. E. Active vision in passive locomotion: Real-world free viewing in infants and adults. *Dev. Sci.* **18**, 736–750 (2015).

Acknowledgements

This work was supported by NSF Award 2043637.

Author Contributions

IJL and JMF conceived and designed the experiment. IJL and AK developed the VR-based visual search task. IJL collected the data. IJL and JMF analyzed the data and interpreted the results. IJL and JMF drafted, edited, and revised the manuscript. IJL, AK, and JMF approved the final version of the manuscript.

Additional Information

Competing interests: The authors declare no competing interests.

Figure 1. Task completion time across age groups and visual complexity levels. *Indicates significant differences between visual complexity levels. #Indicates significant difference between age groups.

Figure 2. Effect of age group and visual complexity level on gaze behavior. (a) Search time between correct targets, (b)re-fixation time on task-relevant objects, (c) fixation time on task-irrelevant objects, and (d) saliency of fixated regions. *Indicates significant difference between visual complexity levels. #Indicates significant difference between age groups.

Figure 3. Effect of age group and visual complexity level on motor behavior. (a) The number of selection errors and (b) the time between fixating and selecting a correct target. *Indicates significant difference between visual complexity levels. #Indicates significant difference between age groups.

Figure 4. Added variable plots showing relationships between task completion time and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task, and (b) working memory capacity assessed using the Backwards Corsi Block task.

Figure 5. Added variable plots showing relationships between re-fixation time on task-relevant objects and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task and (b) working memory capacity assessed using the Backwards Corsi Block task.

Figure 6. Variation in feature congestion across the three visual complexity levels used in the visual search task. (a-c) Single frames from representative gameplay recordings of each visual complexity level. (d) Scatter plot comparing feature congestion of each visual complexity level with each data point representing a frame from representative gameplay recordings. *Indicates significant differences between visual complexity levels.

Table 1. Cognitive assessment scores (Mean ± SD) of young and older adults with the corresponding p-value from their comparisons using two-sample t-tests and Wilcoxon rank sum tests. * Indicates significant differences between age groups.