Aging modulates the effects of scene complexity on visual search in naturalistic 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 the successful completion of many everyday tasks. Prior work demonstrated that aging is associated with increased susceptibility to distraction by salient stimuli. However, these studies often use simple stimuli and little is known about how aging influences visual attention in 3D environments that are more representative of real-world visual complexity. We asked young and older adults to complete a virtual reality-based visual search task with three levels of increasing visual complexity. As the visual complexity of the environment increased, all participants took longer to complete the task, in part because they increased the time spent re-fixating task-relevant objects and the time spent fixating task-irrelevant objects. We also found that older adults took longer to complete the task and spent more time re-fixating task-relevant objects and fixating task-irrelevant objects. In addition, we found that short-term and working memory capacities were related to multiple measures of performance in the visual search task. These results demonstrate the importance of assessing the effects of aging on the control of visual attention using tasks and environments that better capture features of the real world.

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. Processing all stimuli is impractical as the capacity for visual information processing is limited[5]. What is ultimately perceived is dictated by what is selected for using visual attention[6].

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[7]. 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[8]. 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[9–11]. The use of top-down cognitive control to modulate bottom-up processes has also been demonstrated at the neural level during visual search in naturalistic photographs[12]. As measured using fMRI, neural activity was increased for search objects and decreased for distractors. As such, while highly salient stimuli attempt to 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 freely directed based on top-down factors, such as when performing a task while allowing it to be captured by highly salient stimuli when necessary. However, this interaction appears to rely heavily on an individual's cognitive capacity, which has been demonstrated to be affected by normal aging.

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 inhibitory mechanisms has been shown to be impaired[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 completing everyday tasks and make them more vulnerable to accidents that may lead to injuries. However, these studies 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 search targets, leaving questions about its ability to interrogate visual attention in environments that more closely resemble those encountered in the real world. Recent attempts have been made to increase the ecological validity of the Trail Making Test by transforming the task into a three-dimensional reaching-like task in virtual reality (VR)[25]. However, the generalizability of this three-dimensional adaptation to real-world experiences remains as the search targets do not simulate stimuli present in natural environments. In addition, both implementations of the Trail Making Test cannot provide moment-to-moment information on visual attention or visual search strategies as they only measure time to completion.

We aimed to determine how aging-related differences in cognition influence visual search and attention in naturalistic virtual environments. Participants performed a custom-designed VR-based visual search task in three visual complexity levels that closely simulated real-world environments[26]. Performance was quantified by measuring time to completion, and allocation of visual attention was analyzed by quantifying re-fixation time on task-relevant objects, fixation time on task-irrelevant objects, and the saliency of fixated regions of the visual scene. 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 exhibit longer task completion times caused by a decreased capacity to encode task-relevant stimuli in working memory and to inhibit task-irrelevant stimuli. In particular, all participants would be more prone to re-fixating search targets and directing their attention to salient, task-irrelevant distractors. In addition, we hypothesized that this performance decline would be greater in older adults due to aging-related impairments in their cognition.

Results

Feature congestion. The visual complexity of the three virtual environments used in the visual search task was quantified using feature congestion[27]. To determine whether manipulations of the virtual environments led to increasing levels of visual complexity, we performed Welch's analysis of variance to compare the feature congestion of the three visual complexity levels (Fig. 1b). Feature congestion increased monotonically across visual complexity levels ($F(2,1203.7) = 28690$, $p < 2.2e\text{-}16$). Post-hoc Bonferroni-corrected pairwise comparisons using Wilcoxon rank sum tests indicate that the high visual complexity level had greater feature congestion than the medium ($p < 2e\text{-}16$) and low ($p < 2e\text{-}16$) visual complexity levels. In addition, feature congestion was higher in the medium visual complexity level than in the low ($p < 2e\text{-}16$).

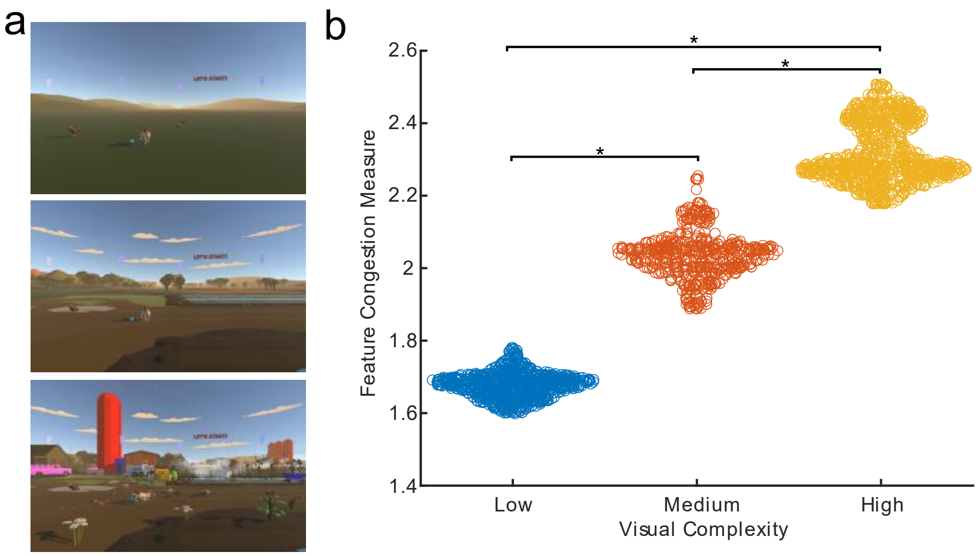

Figure 1. Variation in feature congestion across the three visual complexity levels used in the visual search task. (a) Single frames from representative gameplay recordings of each visual complexity level. (b) Scatter plot comparing feature congestion of each visual complexity level with each data point representing a frame from representative gameplay recordings. * indicates statistically significant differences between visual complexity levels at $p < 0.05$.

Task completion times. While all participants took longer to complete the task as the visual complexity of the environment increased, this effect was greater in older adults (Fig. 2a). There was a main effect of age group on task completion time ($F(1,26) = 48.17$, $p = 2.277e\text{-}7$), with the older adults spending more time completing the task than young adults ($p < 0.0001$). There was also a main effect of visual complexity level ($F(2,52) = 74.82$, $p = 4.977e\text{-}16$), with all participants spending more time completing the task in the high visual complexity level compared to the low (Bonferroni-corrected $p < 0.0001$) and medium (Bonferroni-corrected $p < 0.0001$) visual complexity levels. We also found an interaction between age group and visual complexity level ($F(2,52) = 9.27$, $p = 3.596e\text{-}4$), with the difference in completion between age groups increasing with visual complexity level.

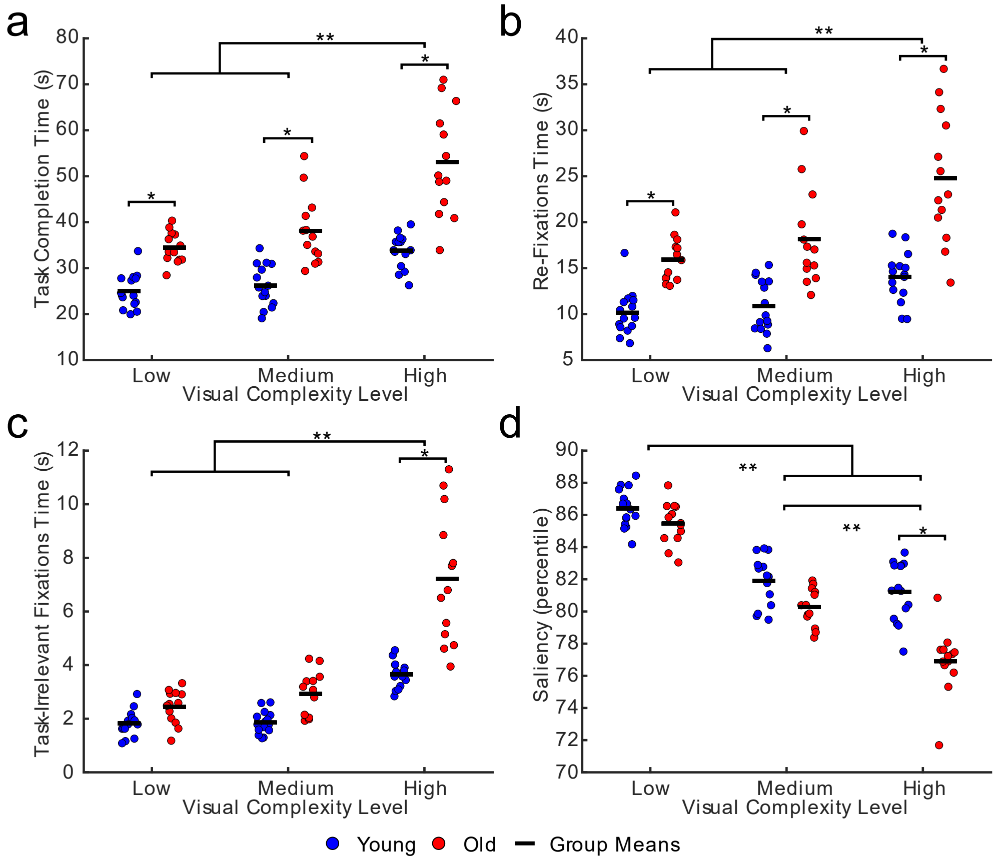

Figure 2. Effect of age group and visual complexity level on task performance. (a) Task completion time, (b) re-fixation time on task-relevant objects, (c) fixation time on task-irrelevant objects, and (d) saliency of fixated regions. ** indicates statistically significant differences between visual complexity levels at $p < 0.05$. * indicates statistically significant differences between age groups within a visual complexity level ($p < 0.05$).

Re-fixation time on task-relevant objects. The increase in task completion time with visual complexity could be due to suboptimal search strategies, such as re-fixating task-relevant objects that were already fixated at an earlier time. As such, we next tested if longer task completion times for levels with higher visual complexity could be due to longer re-fixation times on task-relevant objects (Fig. 2b). A main effect of age group was found ($F(1,26) = 37.16$, $p =$ 1.924e-6), with the older adults exhibiting longer re-fixation times on task-relevant objects than young adults ($p <$ 1e-4). A main effect of visual complexity was also found ($F(2,52) = 42.37$, $p =$ 1.208e-11), with all participants demonstrating longer re-fixation times on task-relevant objects longer in the high visual complexity level as compared to the low (Bonferroni-corrected $p <$ 0.0001) and medium (Bonferroni-corrected $p <$ 1e-4) visual complexity levels. An interaction between age group and visual complexity level was also found ($F(2,52) = 6.07$, $p = 0.004223$), such that the difference in average re-fixation time between age groups increased with visual complexity level.

Fixation time on task-irrelevant objects. Participants may have also taken longer to complete the task as visual complexity increased because they were more susceptible to distraction and spent more time fixating task-irrelevant objects (Fig. 2c). A main effect of age group was found on the amount of time fixating task-irrelevant objects ($F(1,26) = 27.96$, $p = 1.575e-5$), with the older adults fixating task-irrelevant objects longer than young adults ($p < 1e-4$). A main effect of visual complexity was also found ($F(2,52) = 142.434$, $p < 2.2e-16$), with all participants spending more time fixating task-irrelevant objects in the high visual complexity level as compared to the low (Bonferroni-corrected $p < 0.0001$) and medium (Bonferroni-corrected $p < 1e-4$) visual complexity levels. There was also an interaction between age group and visual complexity level ($F(2,52) = 26.598$, $p = 1.107e-8$), such that the difference in the time that young and older adults spent fixating task-irrelevant objects increased with visual complexity level.

Saliency of fixated regions. Although both young and older adults fixated longer on task-irrelevant objects as the visual complexity of the environment increased, it does not appear to be caused by increased fixations to salient distractors (Fig. 2d). A main effect of age group on the saliency of fixated regions was found ($F(1,26) = 32.55$, $p = 5.279e-6$), with older adults fixating regions that are less salient than those fixated by young adults ($p < 1e-4$). A main effect of visual complexity was also found ($F(2,52) = 186.37$, $p < 2.2e-16$). All participants fixated more salient regions in the low visual complexity level compared to the medium (Bonferroni-corrected $p < 1e-4$) and hard (Bonferroni-corrected $p < 1e-4$) visual complexity levels. In addition, all participants also fixated more salient regions in the medium visual complexity level than the high visual complexity level (Bonferroni-corrected $p < 1e-4$). There was also an interaction between age group and visual complexity ($F(2,52) = 11.83$, $p = 5.846e-5$), such that between-group differences in the saliency of fixated regions increased with the level of visual complexity.

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 statistically significant differences between age groups at $p < 0.05$.

Task performance and cognition. Longer task completion times were associated with decreased short-term and working memory capacity across all participants (Fig. 3). 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* = 6.416e-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.

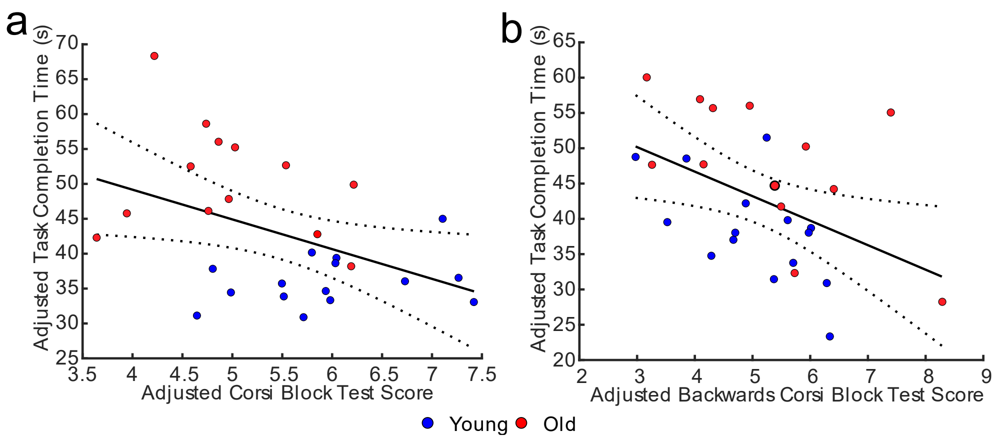

Figure 3. Added variable plots showing relationships between task performance and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task, and (b) working memory capacity assessed using the Backwards Corsi Block task.

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. 4). 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 $R^2 = 0.45$, $F(2,25) = 12.07$, $p = 2.146e\text{-}4$) with higher scores on the Corsi Block task (Fig. 4a; $\beta = -2.36$, $p = 0.0396$) and Backwards Corsi Block task (Fig. 4b; $\beta = -2.11$, $p = 0.0189$) being associated with shorter re-fixation times on task-relevant objects.

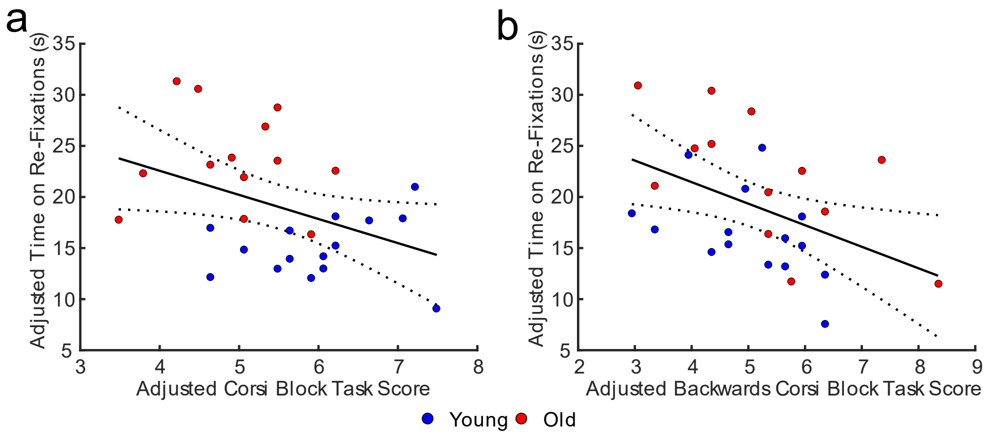

Figure 4. Added variable plots showing relationships between re-fixation time on task-relevant objects and memory capacities. (a) Short-term memory capacity assessed using the Corsi Block task and (b) working memory capacity assessed using the Backwards Corsi Block task.

Discussion

We used a custom-designed 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 are associated with decreased performance in the task. As the visual complexity of the virtual environments increased, participants spent more time completing the task and had longer re-fixation times on task-relevant and task-irrelevant objects. Each of these effects was greater in older than younger adults. However, increased distraction 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. In addition, we found that the variability in performance across participants was explained, in part, by short-term and working memory capacities.

As the complexity of the virtual environments increased, we found that all participants spent more time completing the search task. 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]. Our results indicate that the effect of visual complexity on task performance was greater in older adults. Similar age-related differences in performance have been previously demonstrated using simple conjunction search paradigms[29–31]. These results have also been replicated in paradigms that used more complex stimuli. 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]. In a different study where participants were instructed to search 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 the negative effect of aging, visual complexity, and their combination on search performance generalizes from two-dimensional displays to more naturalistic, three-dimensional virtual environments.

Decreased performance in the search task as the visual complexity of the virtual environments increased coincided with longer fixation times on task-irrelevant objects in all participants. 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. This finding is consistent with previous studies, which found that visual attention is allocated primarily toward task-relevant stimuli when completing various tasks[3,4,36,37].

While all participants became more prone to distraction by task-irrelevant stimuli as visual complexity increased, this effect does not seem to be driven by the saliency of these stimuli. In particular, 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 both the visual complexity of the search environments and the saliency of the stimuli present to better determine the role of saliency in visual information processing in naturalistic environments.

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 block and Backwards Corsi block tasks. These findings echo the importance of short-term and working memory capacity on visual attention[36,38,39], which has been demonstrated in simple selective attention[40] 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. Surprisingly, scores on cognitive assessments that specifically targeted inhibitory capacity were not found to be related to task completion times, despite all participants spending more time fixating task-irrelevant objects as task completion time increased. It is possible that the Stroop and Flanker tasks failed to capture the domain of inhibitory capacity that is particular to the visuospatial nature of the task. Future studies would benefit from including assessments that target multiple domains of inhibitory capacity.

In conclusion, susceptibility to distraction by task-irrelevant information increases with the environment's complexity. This negative effect was greater in older adults, likely driven by age-related changes in their cognition. 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 navigating complex terrain. Results from this study highlight the importance of using tasks with designs that closely simulate real-world conditions and everyday behaviors 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. 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[41] 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[42,43]. Of these computer-based assessments, the Corsi task[44] was used to assess short-term memory, the backward Corsi task[45] for working memory, and both the Stroop[46] and Flanker[47] 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 a different visual complexity level.

VR task. Participants performed a VR-based visual search task (Fig. 1a) designed using principles of the Trail Making Test-B[21]. The Trail Making Test-B was selected based on its ability to test various cognitive domains, including visual attention, working memory, and inhibition[22–24]. They 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. The search targets were positioned randomly in the virtual environment, and their position changed for each trial. The participants viewed the virtual environments using the HTC Vive Pro Eye (HTC, New Taipei, Taiwan) head-mounted display. To interact with the environment and select search targets, the participants used an HTC Vive Controller (HTC, New Taipei, Taiwan). 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. Visual and auditory feedback was 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 (Fig. 1a). 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, five other participants started with the medium visual complexity level, and the rest started with the high visual complexity level. 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[27]. 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).

Data collection and processing. Eye and head movement data were collected at 90Hz 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[48,49], it was important to identify which samples qualify as such to determine how visual attention was deployed throughout each trial. A simple velocity threshold was used to identify whether a sample is a fixation or a saccade. First, horizontal and vertical gaze angular velocities were calculated by differentiating the interpolated gaze angles with respect to the time between video frames. A sample was categorized as a fixation if its horizontal or vertical gaze angular velocity was less than 100 degrees per second[50].

Fixations were then classified as task-relevant or task-irrelevant based on the identity of the object being fixated. Gaze vectors were projected 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. Fixation times on task-irrelevant objects and re-fixation times on task-relevant objects were then calculated.

Saliency maps were created from each frame of every recording based on the original Itti, Koch, and Neibur[51] 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[52,53], with 100% indicating the most salient pixel while 0% the least salient pixel in the video frame.

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.

Linear mixed-effects models were fit using the *lme4* package to test for the effects of age group, visual complexity, and their interaction on VR task completion time, re-fixation time on task-relevant objects, fixation time on task-irrelevant objects, and saliency of fixated regions. Fixed effects p-values were calculated using the *lmerTest* package, which uses Satterthwaite approximations for the degrees of freedom. For all models, random intercepts for each participant were added to account for repeated measures. Post-hoc Bonferroni-corrected pairwise comparisons were performed using the *emmeans* package when significant main effects or interactions, particularly differences between age groups within visual complexity levels, were found.

Cognitive assessment scores were compared 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.

Multiple linear regression was used to test for associations between cognition and task performance measures. 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 from the corresponding author upon reasonable request.

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 in Cognitive Neuroscience* **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. Gazzaniga, M. S., Ivry, R. B. & Mangun, G. R. (George R. *Cognitive neuroscience: the biology of the mind*. (W.W. Norton & Company, 2019).
6. 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).
7. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat Rev Neurosci* **2**, 194–203 (2001).
8. Tatler, B. W., Hayhoe, M. M., Land, M. F. & Ballard, D. H. Eye guidance in natural vision: Reinterpreting salience. *J Vis* **11**, (2011).
9. Cosman, J. D., Lowe, K. A., Zinke, W., Woodman, G. F. & Schall, J. D. Prefrontal Control of Visual Distraction. *Current Biology* **28**, 414-420.e3 (2018).
10. Gaspelin, N. & Luck, S. J. The Role of Inhibition in Avoiding Distraction by Salient Stimuli. *Trends Cogn Sci* **22**, 79–92 (2018).
11. Gaspelin, N. & Luck, S. J. Inhibition as a potential resolution to the attentional capture debate. *Curr Opin Psychol* **29**, 12–18 (2019).
12. Seidl, K. N., Peelen, M. V & Kastner, S. Neural Evidence for Distracter Suppression during Visual Search in Real-World Scenes. *The Journal of Neuroscience* **32**, 11812–11819 (2012).
13. Dempster, F. N. The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review* **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* **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. *Proceedings of the National Academy of Sciences* **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. *Journal of the International Neuropsychological Society* vol. 15 438–450 Preprint at https://doi.org/10.1017/S1355617709090626 (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 Neuroeng Rehabil* **18**, 1–16 (2021).

26. 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.

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

28. Treisman, A. M. & Gelade, G. A feature-integration theory of attention. *Cogn 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. Rothkopf, C. A., Ballard, D. H. & Hayhoe, M. M. Task and context determine where you look. *J Vis* **7**, (2007).

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

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

40. Shipstead, Z., Harrison, T. L. & Engle, R. W. Working memory capacity and visual attention: Top-down and bottom-up guidance. *Quarterly Journal of Experimental Psychology* **65**, 401–407 (2012).

41. 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).

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

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

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

45. Isaacs, E. B. & Vargha-Khadem, F. Differential course of development of spatial and verbal memory span: A normative study. *British Journal of Developmental Psychology* **7**, 377–380 (1989).

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

47. 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).

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

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

50. 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.

51. 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).

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

53. 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. Variation in feature congestion across the three visual complexity levels used in the visual search task. (a) Single frames from representative gameplay recordings of each visual complexity level. (b) Scatter plot comparing feature congestion of each visual complexity level with each data point representing a frame from representative gameplay recordings. * indicates statistically significant differences between visual complexity levels at $p < 0.05$.

Figure 2. Effect of age group and visual complexity level on task performance. (a) Task completion time, (b) re-fixation time on task-relevant objects, (c) fixation time on task-irrelevant objects, and (d) saliency of fixated regions. ** indicates statistically significant differences between visual complexity levels at $p < 0.05$. * indicates statistically significant differences between age groups within a visual complexity level ($p < 0.05$).

Figure 3. Added variable plots showing relationships between task performance 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 4. 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.

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 statistically significant differences between age groups at $p < 0.05$.