Title

Visual Attention During Seeing for Speaking in Healthy Aging

Permalink

https://escholarship.org/uc/item/48v067m1

Journal

Psychology and Aging, 38(1)

ISSN

0882-7974

Authors

Rehrig, Gwendolyn
Hayes, Taylor R
Henderson, John M
et al.

Publication Date

2023-02-01

DOI

10.1037/pag0000718

Peer reviewed

Published in final edited form as:
Psychol Aging. 2023 February ; 38(1): 49–66. doi:10.1037/pag0000718.

Visual Attention during Seeing for Speaking in Healthy Aging

Gwendolyn Rehrig[1], **Taylor R. Hayes**[2], **John M. Henderson**[1,2], **Fernanda Ferreira**[1]
[1]Department of Psychology, University of California, Davis

[2]Center for Mind and Brain, University of California, Davis

Abstract

As we age, we accumulate a wealth of information about the surrounding world. Evidence from visual search suggests that older adults retain intact knowledge for where objects tend to occur in everyday environments (semantic information) that allows them to successfully locate objects in scenes, but may over-rely on semantic guidance. We investigated age differences in the allocation of attention to semantically informative and visually salient information in a task in which the eye movements of younger (N=30, aged 18–24) and older (N=30, aged 66–82) adults were tracked as they described real-world scenes. We measured the semantic information in scenes based on "meaning map" ratings from a norming sample of young and older adults, and image salience as Graph-Based Visual Saliency. Logistic mixed-effects modeling was used to determine whether, controlling for center bias, fixated scene locations differed in semantic informativeness and visual salience from locations that were not fixated, and whether these effects differed for young and older adults. Semantic informativeness predicted fixated locations well overall, as did image salience, although unique variance in the model was better explained by semantic informativeness than image salience. Older adults were less likely to fixate informative locations in scenes than young adults were, though the locations older adults' fixated were independently predicted well by informativeness. These results suggest young and older adults both use semantic information to guide attention in scenes, and that older adults do not over-rely on semantic information across the board.

Keywords

language production; visual attention; cognitive aging

The world around us is visually complex, yet we are able to rapidly and efficiently orient visual attention to relevant visual information. As cognition changes throughout the lifespan, it is likely that the manner and degree to which we process visual information also changes later in adulthood. In the current study, we investigated whether young and older adults use scene information to guide visual attention in the same way when describing real-world scenes. Specifically, we quantified semantic information and image salience for scenes in

---

Correspondence concerning this article should be addressed to Gwendolyn Rehrig, University of California, Davis, Department of Psychology, Davis, CA 95616. glrehrig@ucdavis.edu.

Supporting materials are available on the Open Science Framework (Rehrig et al., 2022): osf.io/shdbz/.

a comparable fashion and determined which best explained where young and older adults fixated as they described scenes.

Healthy aging is associated with various declines in visual cognition. For example, older adults struggle to adequately ignore irrelevant visual information despite receiving explicit instructions to do so (Rabbitt, 1965; Kramer et al., 2000; Williams et al., 2005; Ryan et al., 2007; Williams et al., 2009). Additionally, older adults have greater difficulty inhibiting saccades as compared to young adults when tasked to execute an eye movement to the opposite side of the display as the target (Olincy et al., 1997; Butler et al., 1999; Kramer et al., 1999; 2000) and are less inhibited and less accurate than young adults at deploying saccades in a cueing task (Ryan et al., 2006). When viewing real-world scenes, older adults show less sensitivity overall to local feature contrasts captured by image salience than young adults do, consistent with age-related deficits in bottom-up perceptual processing (Açık et al., 2010; see also Deng et al., 2021), though older adults benefit more from highly salient targets in visual search tasks than their young adult counterparts (Ramzaoui et al., 2021).

When searching for a target object in real-world scenes, older adults search scenes less efficiently than young adults do, requiring longer search times and more fixations before identifying the target (Wynn et al., 2016; Borges et al., 2020; Ramzaoui et al., 2021), though both young and older adults search more efficiently after repeated exposure to the same scene (Wynn et al., 2016; 2019). Age-related declines in cognitive control do not appear to explain the oculomotor age differences: Borges et al. (2020) showed that both young and older adults with better cognitive control were more accurate when searching for objects that were inconsistent with a scene’s semantics (e.g., a clothes iron in a restaurant), but they found no relationship between cognitive control and attention allocation more broadly. The latter finding appears to contradict the observation that inhibitory control predicts performance on antisaccade tasks (Olincy et al., 1997; Butler et al., 1999; Kramer et al., 1999; 2000), which may suggest either that observers do not need to actively inhibit saccades to consistent object locations during visual search in order to successfully locate targets, or that the drive to look at consistent locations for inconsistent objects is less strong than the corresponding drive to look at the directional cue in an antisaccade task. Wynn et al. (2020) posited a connection between older adults’ eye movement patterns in scenes and memory deficits such that changes in eye movement behavior associated with healthy aging may contribute to memory deficits, which in turn may impact subsequent eye movements.

Despite evidence of oculomotor changes associated with healthy aging in visual search tasks, the wealth of world knowledge accumulated with age can facilitate search for target objects in real-world scenes. In real-world visual search tasks, older adults exploit semantic information in the scene to locate target objects effectively (albeit less efficiently), as evidenced by comparable search performance to young adults when searching for targets occurring in locations that are consistent with the scene category (e.g., a hairdryer in a bathroom; Wynn et al., 2019; Borges et al., 2020; Ramzaoui et al., 2021), and older adults who have been diagnosed with age-related macular degeneration use semantic information to guide attention as well as age-matched controls do when searching for target objects in scenes (Pollmann et al., 2020). In a repeated search task, both young and older adults used scene semantics successfully to search for objects, with older adults showing a stronger influence of scene semantics on eye movements (Wynn et al., 2019). The results suggest older adults may be able to use knowledge of scenes to search successfully despite experiencing age-related declines to other aspects of visual cognition.

There is general agreement that we accumulate a wealth of world knowledge as we grow older, and that world knowledge facilitates performance on a range of tasks (Steyvers et al., 2006). Older adults perform as well as young adults, if not better, on tasks that exploit world knowledge (Umanath & Marsh, 2014), and underperform relative to young adults when task performance requires them to contradict their world knowledge. For example, older adults are less accurate than young adults when recalling unrealistic grocery prices in a memorization task (Amer et al., 2018). Similarly, older adults sometimes rely too heavily on semantic information when searching scenes, underperforming relative to young adults when the task requires them to contradict their world knowledge. For example, when asked to search drawings of scenes for anomalies, older adults detected fewer visual errors than young adults did, suggesting they may have difficulty encoding new or contradictory information (James & Kooy, 2011). When searching a scene for a target object, older adults struggle to find targets placed in semantically incongruent locations (Wynn et al., 2019; Borges et al., 2020; but not when scenes were sparse—see Ramzaoui et al., 2021). Both young and older adults showed poorer performance when searching for targets that were inconsistent with scene category (Wynn et al., 2019; Borges et al., 2020), but with practice young adults were able to overcome the difficulty, whereas older adults showed a greater search penalty that did not diminish with repeated searches (Wynn et al., 2019). Furthermore, when older adults searched for targets that were inconsistent with a scene's semantics, search accuracy was worse when the scene to be searched was preceded by a congruent prime (a scene of the same category, or the category name e.g., "kitchen"), showing that older adults had greater difficulty overriding the influence of semantic information when it was reinforced prior to initiating search (Borges et al., 2020). In sum, older adults appear to lean on semantic information more than their young adult counterparts do, which may be either helpful or detrimental depending on the specific search task.

Taken together, the literature suggests that semantic information acquired through experience with scenes may help preserve attentional guidance in older adults (Wynn et al., 2019; Borges et al., 2020; Ramzaoui et al., 2021), but it does not entirely compensate for the deficits observed in oculomotor tasks (Olincy et al., 1997; Butler et al., 1999; Kramer et al., 1999; 2000; Ryan et al., 2006; Wynn et al., 2016; Wynn et al., 2020) including declines in bottom-up processing (Açık et al., 2010). Furthermore, older adults rely more on semantic information than young adults do, even when it is detrimental to their performance (Wynn et al., 2019; Borges et al., 2020), and while there is a preponderance of evidence showing older adults' use of semantic information is preserved (see Umanath & Marsh, 2014), a recent study suggests semantic networks involved in language production tasks may become less flexible and robust in older adults (Cosgrove et al., 2021). Older adults appear to use similar semantic information to young adults, but may use the information in less adaptive ways.

Much of what we know about how semantic guidance of visual attention changes over the lifespan comes from visual search tasks, including the above-mentioned studies (Wynn et al., 2019; Borges et al., 2020; Ramzaoui et al., 2021), which employ target-scene congruence manipulations. Such studies show that older adults use semantic guidance effectively (but less efficiently) when target objects occur in expected locations, suggesting that older adults retain semantic knowledge for scenes but accessing said knowledge carries a cost. However, it is worth noting that visual search is a specialized task in which behavior is influenced by a variety of factors—such as the features of the target object (Malcolm & Henderson, 2009; Zelinsky, 2008), expectations for target locations in scene (Castelhano & Witherspoon, 2016; Neider & Zelinsky, 2006; Peacock et al., 2021), and memory for previous target locations (Võ & Wolfe, 2013; Wynn et al., 2019)—that do not generalize to most other visual tasks (see Peacock et al., 2021 for discussion). To determine whether the aforementioned findings regarding older adults' use of semantic information in scenes generalizes beyond visual search tasks, we examined what factors influence young and older adults' overt visual attention during a scene description task.

In the current study, we used the meaning map paradigm developed by Henderson and Hayes (2017) to investigate the impact of healthy aging on visual attention in real-world scenes. Meaning maps allow researchers to study the influence of scene semantics on attention without the need to manipulate the semantic consistency of objects in scenes. Henderson and Hayes (2017) used crowdsourced ratings of meaning (based on informativeness and recognizability) for isolated scene patches to construct meaning maps, which capture the spatial distribution of semantic information across a scene, and quantified image salience using the Graph-Based Visual Saliency model to capture the degree to which scene regions contrast from their surroundings on the basis of bottom-up image-computable features, such as luminance and orientation (Harel et al., 2006). The resulting meaning and saliency maps were correlated with one another, but meaning maps explained variance in attention better than saliency maps did, especially when only the unique variance separately explained by meaning and saliency maps was considered. Related work using the same paradigm replicated the advantage of semantic information (as measured by meaning maps) over image salience across various tasks, including a free-viewing task (Peacock et al., 2019a), when describing the scene aloud (Henderson et al., 2018), when describing the actions that can be carried out in a scene (Henderson et al., 2018; Rehrig et al., 2020), when engaged in an articulatory suppression task (Rehrig et al., 2020), when semantic information was not task-relevant (Hayes & Henderson, 2019a), and even when image salience was task-relevant and semantic information was not (Peacock et al., 2019b). Taken together, the aforementioned studies suggest that semantic information accounts for the allocation of visual attention in scenes better than image salience does, at least among college-aged adults. The current study uses the meaning map paradigm to investigate whether young and older adults use semantic information in scenes in the same way to guide visual attention during a scene description task.

In the current study, young and older adults described full color real-world scenes aloud, during which time subjects' eye movements were recorded. To determine how semantic information and image salience influence visual attention for both young and older adults, we constructed meaning maps using the method introduced by Henderson and Hayes (2017). Meaning maps differ from semantic congruence manipulations in visual search—the former captures explicit semantic judgments about local scene regions, and the latter implicitly captures expectations about the relationships between objects in scenes—and yet both tap into the semantic representations for objects and scene categories that constitute semantic knowledge for scenes. In the current study, we consider meaning map values to capture semantic informativeness: we use the term “meaning” to refer to meaning map values, and “semantic informativeness” to refer to the local semantic features in a scene that meaning maps were designed to estimate. Because older adults’ knowledge for scenes likely differs from that of young adults with less life experience, we constructed two meaning maps for each scene: one from ratings provided by young adults (aged 18–24; meaning map-Y) and one from older adult raters (aged 55 or over; meaning map-O). Image-computable saliency maps were created using Graph-Based Visual Saliency (GBVS; Harel et al., 2006). To determine what factors influence whether a location in a scene was fixated or not, for each location in a trial that was fixated we sampled a location in the image that was not fixated. For each location (fixated or randomly sampled), we computed the average saliency and age-matched meaning map values for a 3° visual angle window around the coordinate. Because we observed center bias—a tendency for fixations to occur near the center of the screen and not in the periphery (Tatler, 2007; Hayes & Henderson, 2019)—in similar studies (e.g., Henderson et al., 2018; Rehrig et al., 2020), we added center proximity as a fixed effect in our model to account for center bias statistically, using the method developed by Hayes and Henderson (2021a). We computed the Euclidean distance between each pixel location and the center coordinate of the image to determine whether center bias influenced which locations were fixated. To create a more intuitive measure such that higher values correspond to locations closer to the center, the distance was inverted and z-scored to produce center proximity values. We then constructed a logistic mixed-effects model to examine what factors—age group, average meaning map value, average saliency map value, and center proximity—predicted whether a location was fixated.

Based on our prior work (Henderson & Hayes, 2017; Henderson et al., 2018), we expect fixated locations to be more semantically informative (as captured by meaning maps) than randomly sampled locations that were not fixated. If older adults rely more on automatic semantic processing, exploiting a vast accumulation of world knowledge, then we predict an interaction such that the locations older adults fixate will be more semantically informative than the locations young adults fixate—or, at the very least, semantic guidance will be intact for older adults, consistent with a compensatory influence of knowledge for scenes (Wynn et al., 2019; Borges et al., 2020; Ramzaoui et al., 2021). Conversely, because our task is a scene description task, it is possible we may instead observe poorer use of semantic information for older adults, as shown in Cosgrove et al. (2021), in which case we may expect locations older adults fixate to be less informative than those fixated by young adults. Because semantic information and image salience were correlated in our prior work, we predict a main effect of salience by association. Based on Açık et al. (2010), we predict older adults will fixate less salient locations than young adults. In addition, for both groups we anticipate that locations close to the center of the screen would be more likely to be fixated based on the center bias observed in the literature and in our prior work (Tatler, 2007; Hayes & Henderson, 2019), an effect that is robust, but not directly relevant to the theoretical questions addressed here.

Methods

Transparency and Openness

The study hypotheses, design, and analysis plan were not pre-registered. Experimental stimuli, de-identified data, and analysis code are available on the Open Science Framework (see link in the author note).

Subjects

The sample size for each age group ($N = 30$) was based on samples used in previous work using similar tasks and the same scenes (Henderson et al., 2018; Rehrig et al., 2020a; Rehrig et al., 2022) and is comparable to the sample sizes used to test for age differences in previous studies (e.g., Wynn et al., 2019; Borges et al., 2020; Ramzaoui et al., 2021; Açık et al., 2010). An observed power calculation conducted using the R package ‘simr’ (Green & MacLeod, 2016), using the observed effect size (−0.11) and an alpha level of .05, revealed the design was adequately powered to detect the critical interaction of interest between meaning map values and age group (84.00% power, 95% CI = [78.17%, 88.79%]).

Young Adults—Thirty-four undergraduates enrolled at the University of California, Davis participated for course credit. Data were collected during the COVID-19 pandemic (2022). All subjects indicated by self-report that they spoke English as a first language, had not learned a language other than English before the age of 5, were between 18 and 25 years old ($M = 19.27$ years, $SD = 1.44$ years), had completed high school, and had normal or corrected-to-normal vision with no known color blindness. Subjects who elected to share their year in school ($n = 29$) had approximately 13.72 years of education on average ($SD = 1.03$ years) at the time of testing. They were naive to the purpose of the experiment and provided informed consent as approved by the UC Davis IRB Administration (study title: Multi-Utterance Language Production, protocol: 1621276). Three subjects were excluded from analysis because their eyes could not be accurately tracked, and an additional subject was excluded due to previous participation in an experiment involving the same scenes. Data from the remaining 30 subjects (21 female, 7 male, 2 nonbinary) were analyzed. Of the 20 subjects analyzed who disclosed their race, 13 reported their race as white, 3 as Asian, and 3 as mixed-race; one of the aforementioned subjects additionally reported their ethnicity as Latinx.

Older Adults—Thirty-six older adult volunteers from Davis and surrounding areas participated and received $20/hour for completing the study. All indicated by self-report that they spoke English as a first language, had not learned a language other than English before the age of 5, had no known history of dementia, had normal or corrected-to-normal vision, and had no known color blindness. Educational data were collected for all but two of the subjects. Those for whom we have data indicated that they had completed high school, and most ($n = 27$) had some college education as well ($M = 16.68$ education years, $SD = 1.76$ years). Participants provided informed consent as approved by the UC Davis IRB Administration (study title: Behavioral, Electrophysiological and Neuroimaging Studies of Language, protocol: 263396). Data from five subjects who could not be accurately eye tracked were excluded, as well as an additional subject who did not fit the study criteria.

Data from the remaining 30 subjects (22 female, 8 male; 66–82 years old, *M* = 72.63 years, *SD* = 4.94 years) were analyzed. Of the 17 subjects analyzed who disclosed their race, 16 reported their race as white, and one as mixed-race.

Stimuli

Scenes were 30 digitized (1024×768) and luminance-matched photographs of real-world scenes used in Henderson et al. (2018). People were not present in any scenes.

Meaning Maps—We constructed two meaning maps per scene using the context-free mapping procedure described in Henderson and Hayes (2017). The first maps were generated from ratings provided by young adults; the maps will be referred to as meaning map-Y hereafter, and map values will be referred to as meaning-Y. The second set of maps were collected using separate ratings from older adults; maps generated from older adult ratings will be referred to as meaning map-O (and meaning-O to refer to map values). Refer to the Online Supplement for methodological details on the meaning map rating procedure.

Maps were generated from the ratings by averaging, smoothing, and combining the fine and coarse scale maps from the corresponding patch ratings. First, the ratings for each pixel at each scale in each scene were averaged, producing average fine and coarse scale maps, which were then averaged for each scene across scales [(fine map + coarse map)/2]. The final maps were blurred using a Gaussian filter via the MATLAB function ‘imgaussfilt’ with a sigma of 10. On average, meaning-O values (*M* = 3.48, *SD* = 0.38) were higher than meaning-Y values (*M* = 2.96, *SD* = 0.35; Figure 2), and a paired *t*-test revealed that the difference in means was significant ($t(29) = -32.20$, $p < .0001$).

After generating age-specific maps from meaning map ratings (see Online Supplement), the maps were scaled from 0 to 1 prior to analysis (see Figure 2E&F, as well as the appendix, for example meaning maps).

Saliency Maps—Image-based saliency maps were constructed using the Graph-Based Visual Saliency (GBVS) toolbox in Matlab with default parameters (Harel et al., 2006). We used GBVS because it is a model that computes salience using only semantically uninterpreted image-computable information (relative to deep saliency models; see Hayes & Henderson, 2021b). A 2-step whitening procedure was used to remove the center bias included in the GBVS model (Rahman & Bruce, 2015). A standardized version of each saliency map was created. Each standardized map had a mean of 0 and a standard deviation of 1. Then a pixel-wise standardization procedure was performed across all standardized maps so that each pixel location had a mean of 0 and standard deviation of 1. This process served to remove the saliency map activation that was shared by all the scenes (the peripheral downweighting that introduces center bias), and preserved the variance that was scene-dependent (see Hayes & Henderson, 2019). The maps were then scaled from 0 to 1 (see Figure 2D for an example saliency map).

The resulting maps correlated with one another (Table 1). Meaning maps-Y and meaning maps-O showed a high degree of overlap (*M* = 0.85, *SD* = 0.05), indicating that both young and older adults appraised the informativeness of scenes similarly. Saliency maps overlapped less with both meaning maps-Y (*M* = 0.13, *SD* = 0.10) and meaning maps-O (*M* = 0.13, *SD* = 0.09).

Apparatus

Eye movements were recorded with an SR Research EyeLink 1000+ tower mount eyetracker (spatial resolution 0.01) at a 1000 Hz sampling rate. Head movements were minimized using a chin and forehead rest integrated with the eyetracker's tower mount. Subjects were instructed to lean against the forehead rest to reduce head movement while allowing them to speak. Although viewing was binocular, eye movements were recorded from the right eye. The experiment was controlled using SR Research Experiment Builder software. Scenes were displayed at 1024×768 pixel resolution. Participants sat 83 cm away from a monitor such that scenes subtended approximately 26° × 19° visual angle, presented in 4:3 aspect ratio. Audio was digitally recorded using a Shure SM86 cardioid condenser microphone. Recorded speech was preamplified using an InnoGear IG101 phantom power preamplifier.

Procedure

A calibration procedure was conducted at the beginning of each experimental session to map eye position to screen coordinates. Successful calibration required an average error of less than 0.49° and a maximum error below 0.99°. Fixations and saccades were parsed with EyeLink's standard algorithm using velocity and acceleration thresholds (30°/s and $9500°/s^2$; SR Research, 2017).

Following successful calibration, subjects were told they would see a series of scenes presented individually, and were instructed to describe each scene aloud. The instruction was followed by three practice trials to familiarize subjects with the task and the response window duration. Subjects used a button box to advance throughout the task.

Each subject received a unique trial order that was pseudorandomized to prevent two scenes of the same scene category (e.g., kitchen) from occurring consecutively. A trial proceeded as follows. A five-point fixation array was displayed to check calibration, during which the subject fixated on the central fixation point and the experimenter pressed a key to begin the trial if the fixation was stable, and reran the calibration procedure if not. The scene was then shown for a period of 30 seconds, during which time eye movements and audio were recorded simultaneously. After 30 seconds elapsed, subjects pressed any button on the button box to begin the next trial. The trial procedure repeated until all 30 trials were complete.

Eye movement data were imported offline into Matlab using the Visual EDF2ASC tool packaged with SR Research DataViewer software. The first fixation was excluded from analysis, as were fixation duration (<50ms, >1500ms) and saccade amplitude outliers (>20°). Fixations that fell within 5 pixels of the image border, consistent with momentary track loss, were also excluded.

Following the eye-tracking task, older adults completed a task unrelated to the current study.

Analysis

Fixated Locations—We examined which features influenced visual attention by comparing the saliency and meaning map values at locations in the scene that were fixated to map values for locations that were not fixated, operating on the assumption that differences between regions of the scene that were and were not fixated can speak to what information influences attention, following Nuthmann et al. (2017). Rather than dividing the scene into a grid (as Nuthmann et al., 2017 did), we used the procedure developed by Hayes and Henderson (2021) to measure saliency and meaning map values in a window around each fixated location, approximating the fovea, and compared the average map values for fixated locations to those of sampled locations that were not fixated in the scene. We constructed a logistic mixed-effects model in which the dependent variable was whether a location was fixated (1) or not (0). The mixed-effects model allowed us to determine whether age group (young or older adults), meaning map values, or saliency map values, or a bias to look at the center of the image (Tatler, 2007; Hayes & Henderson, 2019), predicted where subjects looked in the scene, while simultaneously controlling for random effects of subjects and unique scenes.

The dependent variable was defined as follows. For each subject and each trial, the x,y coordinates corresponding to the subject’s fixations were assigned a value of 1 (fixated). A number of locations that were not fixated equal to the number that were fixated were then randomly sampled from all possible coordinates in the 1024×768 image—excluding those that the subject fixated on during that trial, or locations that fell within 1.5° visual angle (56 pixel) radius around fixations—using the ‘sample’ function from the ‘random’ module in Python 3. The randomly sampled coordinates were assigned a value of 0 (not fixated)(Figure 2A).

For each x,y coordinate pair, we computed the mean meaning, saliency, and center proximity map values corresponding to a 3° (113 pixel) diameter window around the coordinate (Figure 2C&D). We defined a mask for the region around the fixation using a 1.5° (56) pixel radius. The mask was then used to extract an array of map values for the meaning, saliency, and center proximity maps, and the mean of each array was stored as the average meaning, saliency, or center proximity map values corresponding to the x,y coordinate under consideration. Because the high correlation between meaning-Y and meaning-O maps caused high collinearity in the data set[1], we instead constructed an age-matched meaning variable (rather than including both map values as separate predictors) such that x,y coordinate pairs corresponding to young adult subjects were assigned meaning map-Y values and older adults’ x,y coordinates were assigned meaning map-O values.

A logistic mixed-effects model was constructed using the ‘glmer’ function of the ‘lme4’ package in R using the default optimizer (bobyqa). Each predictor (meaning, saliency, and center proximity) was placed on a common scale by subtracting the mean and dividing by the standard deviation using the ‘scale’ function in base R. The first model included age group (young or older, using treatment coding with young as the reference level), average

---

[1] Variance inflation factors were 9.61 for meaning map-O and 9.33 for meaning map-Y values (values above 5 indicate the presence of worrying collinearity; see James et al., 2013).

meaning map value, average saliency map value, and center proximity as fixed effects, and interactions between all fixed effects. Two additional models were constructed using only data from older adults or young adults to determine which of the aforementioned feature variables (average meaning, saliency, and center proximity map values) predicted locations that participants in either age group fixated independently. All predictors were standardized prior to analysis. Random effects were included for subjects and items. We included random slopes and intercepts for all fixed effects and their interactions in both subject and item random effect structures. Random slopes corresponding to the fixed effect of age group were not included in the subject random effect because it is a between-subjects variable. The models failed to converge when random intercepts and slopes were correlated, and converged successfully when they were uncorrelated using the double vertical bar operator. Odds ratios were estimated from each model to facilitate interpretation: Odds ratios below one indicate that an increase in a predictor corresponds to a decrease in the odds the location was fixated; conversely, odds ratios above one indicate higher odds that the location was fixated given an increase in a predictor (Tenny & Hoffman, 2021).

Results

Fixated Locations

Locations that were fixated in the scene had higher meaning map-O values ($M = 0.68$, $SD = 0.17$) than meaning map-Y values ($M = 0.62$, $SD = 0.18$) on average (Figure 3). For both map types, fixated locations had higher map values than randomly sampled locations that had not been fixated ($M_Y = 0.43$, $SD_Y = 0.19$, $M_O = 0.50$, $SD_O = 0.21$), reflected in the simple main effect of meaning in the model ($\beta = 1.17$, $z = 14.79$, $p < .0001$, OR = 3.22, 95% CI = [2.76 3.76]; Table 2). Fixated locations were also higher in average image salience ($M = 0.39$, $SD = 0.18$) than locations that were not fixated ($M = 0.32$, $SD = 0.17$), reflected in the simple main effect of salience ($\beta = 0.19$, $z = 2.34$, $p = .02$, OR = 1.21, 95% CI = [1.03 1.41]). Fixated locations were closer to the center of the image on average ($M = 0.49$, $SD = 0.95$) than sampled locations that were not fixated ($M = -0.28$, $SD = 0.90$), reflected in the simple main effect of center proximity, consistent with center bias ($\beta = 0.50$, $z = 5.94$, $p < .0001$, OR = 1.65, 95% CI = [1.40 1.95]). Slopes for meaning differed from those for salience such that salient locations were more likely to be fixated if they were also semantically informative ($\beta = 0.14$, $z = 2.44$, $p = 0.01$, OR = 1.16, 95% CI = [1.03 1.30]; Figure 4A). Consistent with previous findings (Henderson & Hayes, 2017; Henderson et al., 2018), semipartial correlations computed using the R package ‘partR2’ revealed that meaning accounted for more unique variance in the model ($R^2 = 0.25$) than image salience did ($R^2 = 0.02$).

Older adults made more fixations per second ($M = 3.21$, $SD = 0.51$) than young adults did ($M = 2.87$, $SD = 0.34$): $t(1574.9) = -16.25$, $p < .0001$, and older adults’ fixations were shorter in duration ($M = 257.44$ ms, $SD = 143.45$ ms) than those of young adults ($M = 277.82$ ms, $SD = 150.99$ ms): $t(152{,}824) = 27.08$, $p < .0001$. There was a simple main effect of age group such that older adults were less likely to fixate locations than young adults ($\beta = -0.15$, $z = -4.63$, $p < .0001$, OR = 0.86, 95% CI = [0.81 0.92]). Saliency map values for fixated locations were similar for young ($M = 0.39$, $SD = 0.18$) and older adults ($M = 0.38$, *SD* = 0.18), but there was a marginal interaction between image salience and age group (β = −0.06, *z* = −1.84, *p* = 0.07, OR = 0.94, 95% CI = [0.89 1.00]; Figure 4B). Although meaning map-Y values for the locations young adults fixated were lower (*M* = 0.63, *SD* = 0.17) than meaning map-O values for locations fixated by older adults (*M* = 0.67, *SD* = 0.17), there was a reliable interaction in the model such that, despite the higher meaning map values associated with meaning maps-O overall, older adults were less likely than young adults to fixate locations in the scene that were informative (β = −0.17, *z* = −3.57, *p* = 0.0004, OR = 0.84, 95% CI = [0.77 0.93]). While older adults fixated locations that were further from the center of the screen on average (*M* = 0.45, *SD* = 0.95) than the locations young adults fixated (*M* = 0.53, *SD* = 0.95), there was no corresponding interaction between age group and center proximity (β = −0.04, *z* = −0.94, *p* = 0.35, OR = 0.96, 95% CI = [0.88 1.05]). The model revealed no other main effects or reliable interactions.

To determine whether the locations that older adults fixated were predicted by semantic information and image salience independently (outside of comparison to young adults' fixations), we constructed a second model to analyze only the older adult fixation data. The model was identical to the model that tested for group differences, except that the age group variable was absent.

In the older adult only model, there was a main effect of meaning such that older adults were more likely to fixate locations that were higher in meaning (β = 0.98, z = 11.37, p < .0001, OR = 2.65, 95% CI = [2.24 3.15]) and a marginal effect of salience (β = 0.15, z = 1.77, p = 0.08, OR = 1.16, 95% CI = [0.98 1.38]). There was a reliable interaction between meaning and salience indicating that older adults were more likely to fixate locations that were both meaningful and salient (β = 0.13, z = 2.29, p = 0.02, OR = 1.14, 95% CI = [1.02 1.27]). Semipartial correlations showed meaning accounted for more unique variance (R^2 = 0.21) than image salience (R^2 = 0.02), consistent with the previous model. There was a main effect of center proximity such that older adults were more likely to fixate locations with higher center proximity (β = 0.46, z = 5.11, p < .0001, OR = 1.58, 95% CI = [1.32 1.89]). No other predictors or interactions were significant.

In a comparable model with the same variable structure that was constructed using only data from the young adult sample, there were main effects of meaning (β = 1.34, z = 13.57, p < .0001, OR = 3.83, 95% CI = [3.15 4.64]) and salience (β = 0.23, z = 2.55, p = 0.01, OR = 1.25, 95% CI = [1.05 1.49]) such that young adults were more likely to fixate locations that were higher in either meaning or salience. There was a reliable interaction between meaning and salience (β = 0.16, z = 0.08, p = 0.049, OR = 1.17, 95% CI = [1.00 1.37]). Semipartial correlations showed meaning accounted for more unique variance (R^2 = 0.30) than image salience (R^2 = 0.03), consistent with the other models. Note that meaning accounted for more unique variance in the young adult only model (R^2 = 0.30) than in the older adult only model (R^2 = 0.21), consistent with the interaction between meaning and age reported in the combined model. There was a main effect of center proximity suggesting young adults preferentially fixated locations with higher center proximity (β = 0.54, z = 5.31, p < .0001, OR = 1.71, 95% CI = [1.40 2.09]). No other predictors or interactions were significant.

In sum, locations that were higher in semantic information, higher in image salience, and closer to the center of the screen were more likely to be fixated, and salient locations were more likely to be fixated when they were also informative. There were age group differences such that the locations older adults fixated were predicted less well by semantic informativeness than the locations that young adults fixated, despite older adult fixations being predicted well independently by semantic informativeness in a separate model.

Discussion

In the current study, we examined the locations older and young adults fixated while describing real-world scenes aloud for differences in visual attention across age groups. Consistent with our prior work, we predicted that fixated locations would be more semantically informative (as captured by meaning maps-Y and -O) than locations that were not fixated for both age groups (Henderson & Hayes, 2017; Henderson et al., 2018; Rehrig et al., 2020). We anticipated the effect of semantic guidance would be stronger for older adults if older adults indeed rely more on semantic information than young adults do to navigate scenes generally (Madden et al., 2004; Wynn et al., 2019; Borges et al., 2020). We predicted that fixated locations would also be higher in image salience than locations that were not fixated, due to the correlation between saliency and meaning maps observed in prior work (Henderson & Hayes, 2017; Henderson et al., 2018; Rehrig et al., 2020). Because older adults show deficits in early visual processing (Açık et al., 2010; Deng et al., 2021), however, we predicted an age difference such that older adults would be less likely to orient attention to salient regions in the scene than young adults. We expected fixated locations to be closer to the center of the screen for both age groups, consistent with the center bias observed in prior work (Tatler, 2007; Hayes & Henderson, 2019), and we made no specific prediction about age differences in center bias.

In the current study, we used a logistic mixed-effects model to determine what factors predict the locations observers fixated in the scene. Fixated locations had higher saliency map values. Consistent with our predictions, and our prior work (Henderson & Hayes, 2017; Henderson et al., 2018), fixated locations had higher meaning map values than locations that were not fixated; furthermore, semipartial correlations revealed that image salience explained little unique variance once the relationship with semantic informativeness (as measured by meaning-O and meaning-Y values) was accounted for, suggesting that scene informativeness was a stronger influence on overt attention. Locations were more likely to be fixated when they were both more semantically informative and more salient. Taken together, the results suggest that visual attention, as parameterized in our model, was guided primarily by scene informativeness.

Based on claims in the literature that older adults rely more on semantic information than young adults do, we expected higher meaning map values to predict which scene locations older adults fixated better than young adults. Instead, we found the opposite pattern in the first model: The locations older adults fixated were less semantically informative—and marginally less visually salient—than the locations young adults fixated. Crucially, it was not the case that older adults did not demonstrate semantic guidance of visual attention: a separate analysis (the second model) in which older adults were not compared against young adults showed that high meaning map values independently predicted where older adults fixated, suggesting the difference was not due to an insensitivity of older adults to scene informativeness, or an inability to understand the scenes. Counter to our expectations, our results are not consistent with the literature showing that older adults rely more on semantic information than young adults do (Madden et al., 2004; Umanath & Marsh, 2014; Wynn et al., 2019; Borges et al., 2020).

Why do our results differ from those of Wynn et al. (2019), Borges et al. (2020), Ramzaoui et al. (2021), and from our own predictions? While the aforementioned studies provide important context for our research question, it is difficult to draw direct comparisons between the current study and experiments that employed semantic congruence manipulations to study semantic guidance of attention due to differences in the paradigms used. It is possible that either the method of investigating semantic guidance in the current study (using meaning maps to quantify semantic informativeness), the scene description task, or the analysis method used may be responsible for the difference in results between our study and the visual search literature. Our methods have several advantages over those used in previous studies. First, instead of limiting a critical semantic manipulation to select objects or locations in the scene, we were able to measure local semantics across the entirety of the scene. Second, we used a logistic mixed-effects model that allowed us to account for the influence of critical factors known to influence visual attention, such as image salience and center bias, and to control for differences attributable to individual subjects and scenes through the inclusion of random effects. Third, while visual search studies capture observers' expectations about the relationships between objects and scenes, the scene description task may be better able to engage scene semantics than visual search tasks are. A promising direction for future work would be to compare semantic informativeness for the locations where young and older adults look during both a scene description task and a visual search task using the analysis approach in the current study.

Another possibility is that the task constraints unique to visual search push older adults to rely on visual scene information more than young adults do, both because search for a target object makes information specific to that object highly task-relevant, and because the task requires a response that is either correct or incorrect—in other words, the costs associated with the strategy of using stored knowledge to complete task goals may be worth paying when accuracy matters, but such a strategy would likely not be worth the cost in a relatively unconstrained task like the one used in the current study. If the latter explanation accounts for our results, it would suggest older adults use semantic information strategically, only paying the cost to access it when doing so improves task performance. Such a strategy shift—in which careful processing is deprioritized over other factors, such as speed—might also explain the observation that older adults made more fixations per second during scene viewing and showed shorter fixation durations. We expect that any such task-related strategy shifts occur without conscious awareness. The above proposed future directions, in addition to addressing methodological differences between the current study and visual search studies, would also be able to address whether age-related differences in semantic guidance are influenced by task goals.

Another possible explanation for the finding that older adults' fixated locations were less informative than young adults' is that the description task may have been more cognitively taxing for older adults, and cognitive load has been shown to impact oculomotor behavior. Children look away from informative stimuli while answering difficult questions to reduce cognitive load associated with bottom-up stimulus processing (Doherty-Sneddon & Phelps, 2005), and irrelevant bottom-up stimulus information interferes with top-down processing during long-term memory retrieval (Wais et al., 2010). Similarly, performance in a recall task was higher when subjects' eyes were closed (Vredeveldt & Hitch, 2011). Additionally, observers make more eye movements when engaged in long-term memory search tasks as opposed to tasks with no memory search component (Ehrlichman & Micic, 2012) and are less sensitive to bottom-up stimulus information when engaged in a task that poses cognitive load (Buetti & Lleras, 2016). It is possible that older adults similarly looked away from informative regions of the scene to manage the cognitive load associated with incremental description planning more than young adults did. A passive viewing task would be able to speak to whether managing cognitive load contributed to older adults' differential use of semantic information relative to young adults in our task. In a similar vein, it is possible that age-related declines in inhibitory control could explain our findings. Although Borges et al. (2020) did not find an influence of cognitive control on attention in a search task, it is unclear whether age-related declines in inhibitory control could have influenced attention allocation in our task. Future work could assess inhibitory control in both the young and older adult groups to determine what role cognitive control might play in attentional guidance using a similar task.

Older adults rated isolated scene patches as higher in meaning (defined as informativeness and recognizability) than young adult raters did, and their ratings produced more information-dense meaning maps. One possible explanation for the difference in ratings is that older adults may have been better able to recognize objects in the scene (such as a television antenna or VCR) that young adults would likely have had less experience with, and older adults have had more life experience with objects familiar to raters of both age groups, rendering those objects more recognizable and informative. Future work could compare ratings for patches depicting relatively new objects that both groups would likely have comparable experience with, such as smart home speakers, to determine whether differences in experience with objects accounted for the difference in patch ratings observed in the current study.

Based on evidence in the literature that older adults rely more on semantic information than young adults do, we expected higher meaning map values—to the extent that meaning maps capture information acquired through experience with real-world scenes—to predict which scene locations older adults fixated better than young adults. Older adults in our study rated scene patches as higher in meaning than young adults did, and produced denser meaning maps, consistent with the idea that older adults have richer semantic representations for scenes. However, the locations older adults fixated were less semantically informative than the locations young adults fixated, despite the use of age-matched meaning map values in our analysis. Our results are not consistent with the finding that older adults rely more on semantic information than young adults do (Madden et al., 2004; Wynn et al., 2019)—or that older adults use such information as well as young adults (Ramzaoui et al., 2021)—in visual search tasks. Given that our study used a language production task, our results may be consistent with poorer use of semantic information in older adults relative to young adults recently reported in a language production task (Cosgrove et al., 2021), with the caveat that the findings of the present study, as well as those of Cosgrove et al. (2021), are inconsistent with studies demonstrating that older adults use semantic information well. Future work might investigate whether there is any relationship between the age group difference for visual attention that we observed and the richness or accuracy of the subjects' verbal descriptions. For example, perhaps older adults described less of the scene's content, or struggled to retrieve accurate terms for objects in the scene, relative to their young adult counterparts. Alternatively, given that older adults tend to have larger vocabularies than younger adults, it is possible the older adults described the scenes in more detail and using a richer set of linguistic expressions. Given how little is known at this stage, a priori hypotheses are difficult to formulate.

Although our older adult subjects indicated by self-report that they had no known history of dementia, the lack of a cognitive status measure in our study is a limitation. In hindsight we would have collected such a measure in each experimental session, but unfortunately we are now not in a position to do so. However, our older adult sample's characteristics are consistent with participants in other studies whose cognitive status measures indicated normal cognitive function (e.g., Wynn et al., 2019; Ramzoui et al., 2021; Pereira et al., 2020; McLaughlin et al., 2010) with respect to average age (72.63 years, which is on the younger side), gender distribution (majority female, as were the subjects in the young adult sample), and years of education attained (16.68 years). More importantly, it is unclear how our results would have been affected if our older adults had undetected mild cognitive impairment (MCI), given that MCI does not appear to influence the spatial deployment of attention (Yang et al., 2011,2012; Seligam & Giovannetti, 2015; Molitor et al., 2015; Coco et al., 2021). In sum, our older adult sample is demographically similar to older adult samples in the literature that do not show signs of cognitive impairment, and it is unclear whether or not the presence of MCI would have impacted the spatial distribution of attention as captured by our dependent variable.

Another limitation of the current study is that the young adult sample was more racially and ethnically diverse than the older adult sample, which introduced cultural background as a potential confound in our study. This difference could be relevant given that many indoor environments (as well as some outdoor environments) differ across cultures. At the same time, it is important to note that meaning maps were generated separately for each group based on ratings obtained from the same age cohorts. A final limitation is that the use of a cross-sectional between-subjects design in our study to examine age differences limited our ability to draw inferences about the effects of aging compared to other differences between our young and older adult samples (Nyberg et al., 2010). Funding and resources permitting, we could bring the same older adults back to the lab to repeat the description experiment and evaluate changes in visual attention over time for those individuals.

Conclusions

The current study investigated whether scene meaning predicts where older adults look in scenes better than young adults, given that older adults have a larger semantic knowledge base and have been shown to over-rely on semantic information in visual search tasks (e.g., Wynn et al., 2019; Borges et al., 2020). We replicated our previous map-level analyses of visual attention (Henderson et al., 2018; Rehrig et al., 2020) using a logistic mixed-effects model and meaning maps that were generated separately for young and older adults. We found that semantic information predicted fixated scene locations less well for older adults than young adults, although older adults' fixations were independently predicted well by semantic information, suggesting that older adults did not over-rely on semantic information as captured by meaning maps to guide visual attention while describing scenes.

Supplementary Material

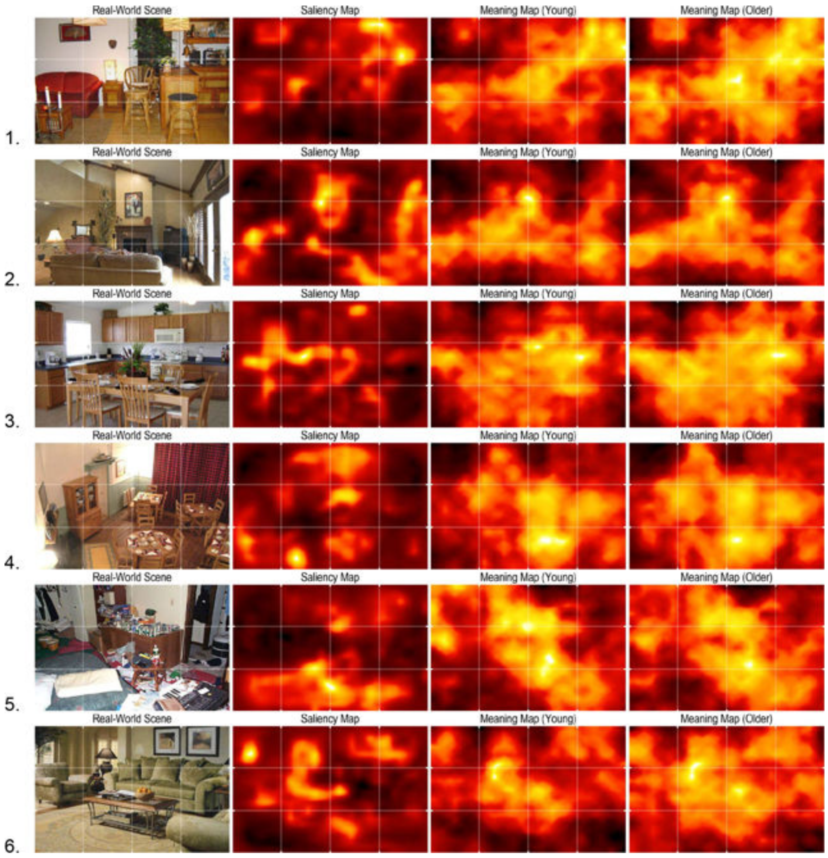
Refer to Web version on PubMed Central for supplementary material.

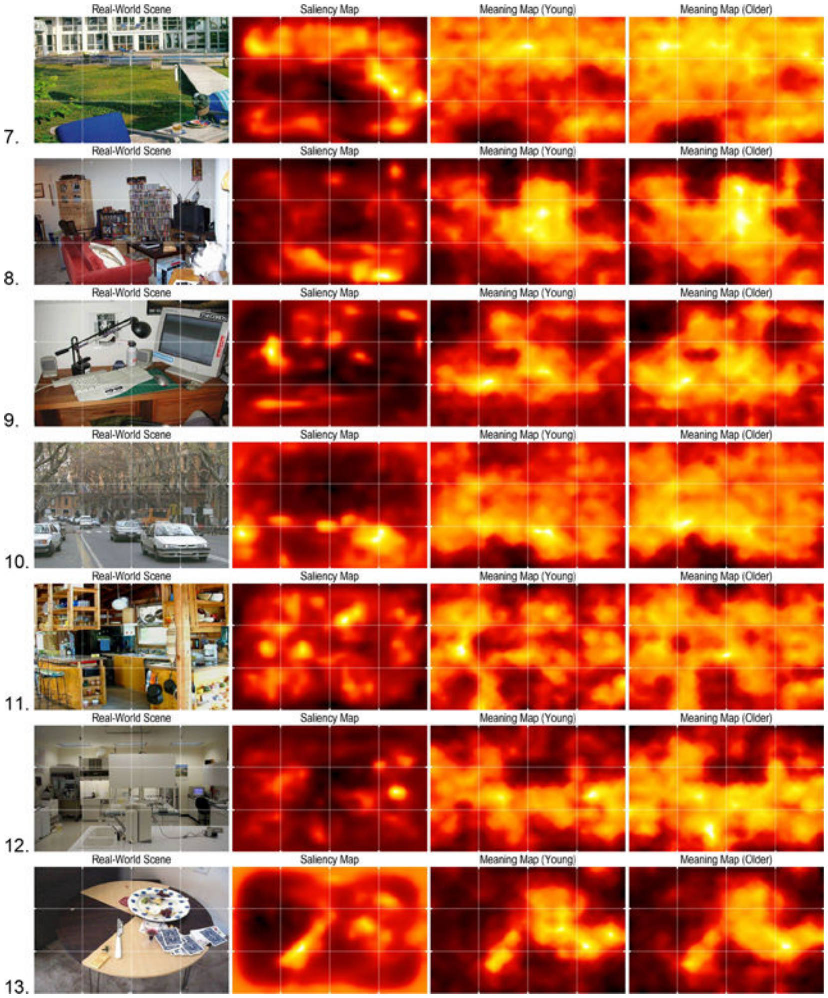
Acknowledgments

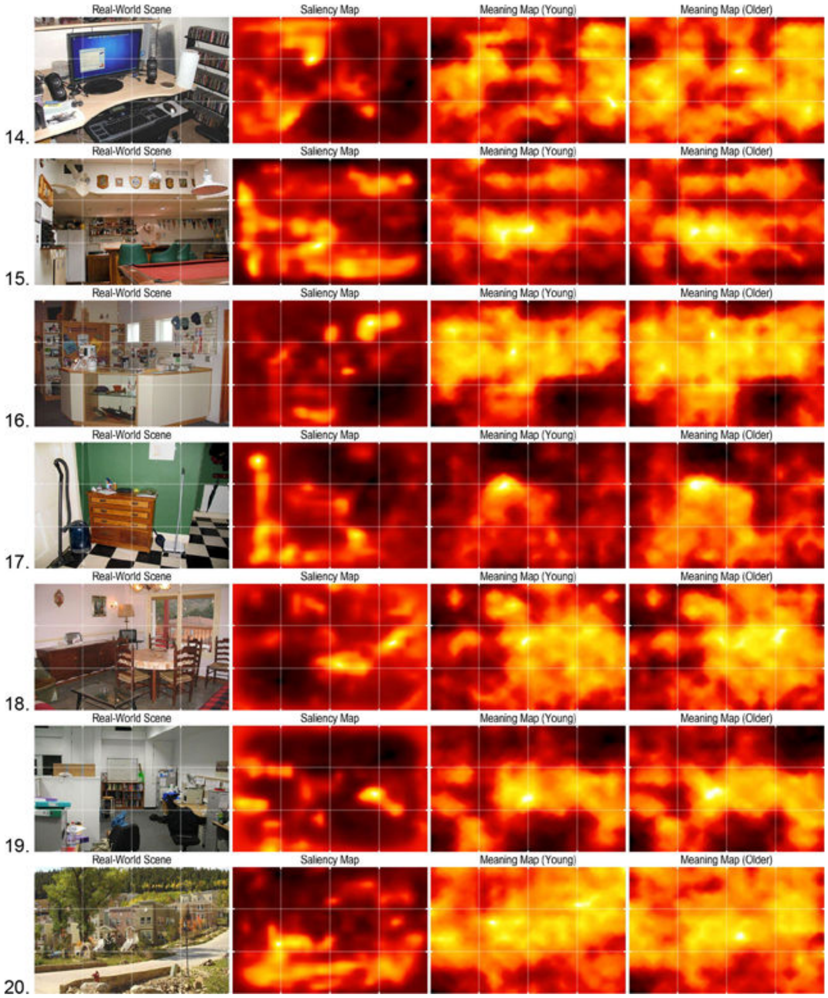
This work was funded by the National Institute of Health grants R56 AG053346 awarded to Fernanda Ferreira, John Henderson, and Tamara Swaab, R01 HD100516 awarded to Fernanda Ferreira, and R01 EY027792 awarded to John Henderson, and by National Science Foundation grants BCS1650888 awarded to Fernanda Ferreira and BCS2019445 awarded to John Henderson. The young adult dataset was used in a prior publication (Henderson, Hayes, Rehrig, & Ferreira, 2018: "Meaning Guides Attention during Real-World Scene Description", published in Scientific Reports, DOI:10.1038/s41598-018-31894-5). A preliminary version of this work was presented as a poster at the Annual Meeting of the Psychonomic Society, November 21, 2020.

Appendix

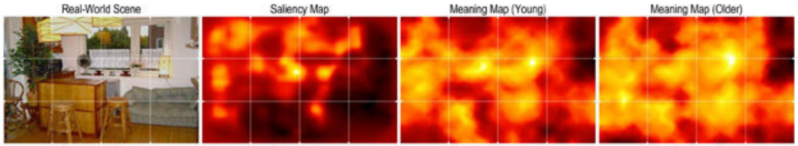
Real-world scenes presented in the eyetracking task, alongside visualizations of the corresponding saliency, meaning-Y, and meaning-O maps for each scene. Numbers next to each scene correspond to the scene numbers shown on the x-axis in Figure 1.


Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
1.
2.
3.
4.
5.
6.

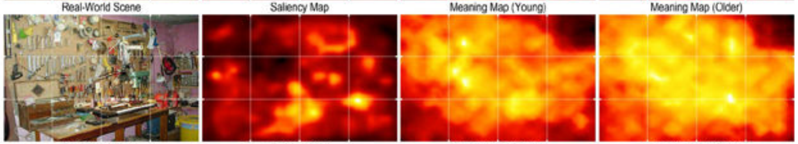

Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
7.
8.
9.
10.
11.
12.
13.


Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
14.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
15.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
16.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
17.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
18.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
19.
Real-World Scene
Saliency Map
Meaning Map (Young)
Meaning Map (Older)
20.

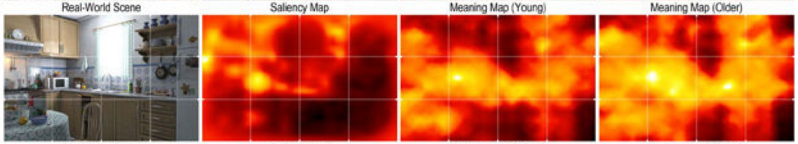
21.



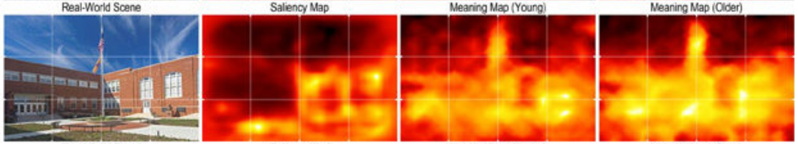
22.



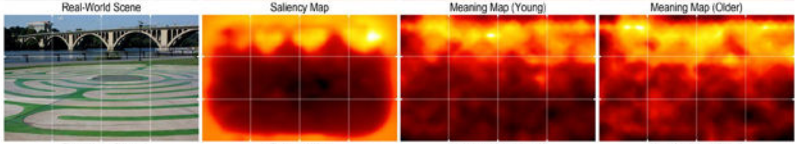
23.



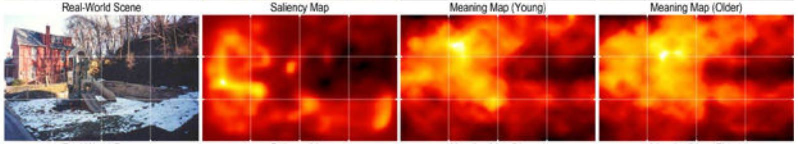
24.



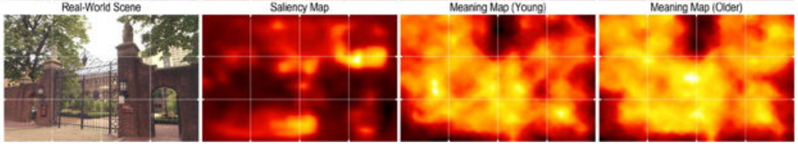
25.



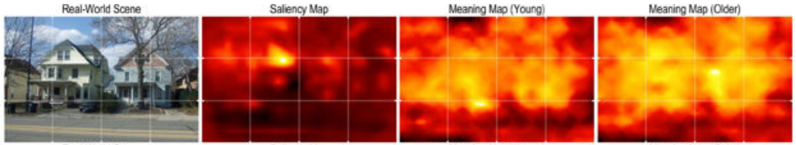
26.



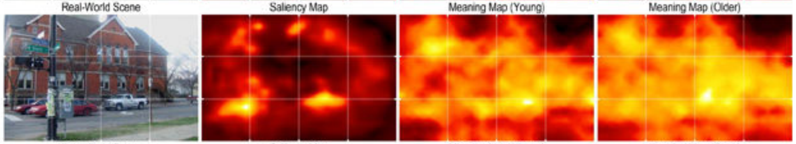
27.



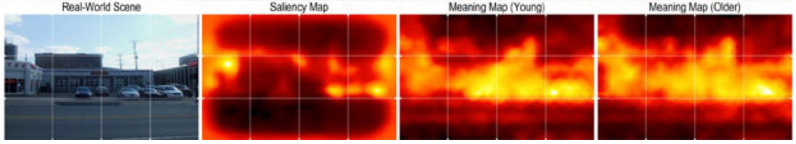
28.



29.



30.



References

Açık A, Sarwary A, Schultze-Kraft R, Onat S, & König P (2010). Developmental changes in natural viewing behavior: bottom-up and top-down differences between children, young adults and older adults. Frontiers in Psychology, 1, 207. [PubMed: 21833263]

Amer T, Giovanello KS, Grady CL, & Hasher L (2018). Age differences in memory for meaningful and arbitrary associations: A memory retrieval account. Psychology and Aging, 33(1), 74. [PubMed: 29494179]

Antes JR (1974). The time course of picture viewing. Journal of Experimental Psychology, 103(1), 62–70. [PubMed: 4424680]

Borges MT, Fernandes EG, & Coco MI (2020). Age-related differences during visual search: the role of contextual expectations and cognitive control mechanisms. Aging, Neuropsychology, and Cognition, 27(4), 489–516.

Buetti S, & Lleras A (2016). Distractibility is a function of engagement, not task difficulty: Evidence from a new oculomotor capture paradigm. Journal of Experimental Psychology: General, 145(10), 1382–1405. [PubMed: 27690512]

Butler KM, Zacks RT, & Henderson JM (1999). Suppression of reflexive saccades in younger and older adults: Age comparisons on an antisaccade task. Memory & Cognition, 27, 584–591. [PubMed: 10479817]

Coco MI, Merendino G, Zappalà G, & Della Sala S (2021). Semantic interference mechanisms on long-term visual memory and their eye-movement signatures in mild cognitive impairment. Neuropsychology, 35(5), 498–513. [PubMed: 34166040]

Cosgrove AL, Kenett YN, Beaty RE, & Diaz MT (2021). Quantifying flexibility in thought: The resiliency of semantic networks differs across the lifespan. Cognition, 211, 104631. [PubMed: 33639378]

Dave S, Brothers T, Hoversten LJ, Traxler MJ, & Swaab TY (2021). Cognitive control mediates age-related changes in flexible anticipatory processing during listening comprehension. Brain Research, 147573.

Deng L, Davis SW, Monge ZA, Wing EA, Geib BR, Raghunandan A, & Cabeza R (2021). Age-related dedifferentiation and hyperdifferentiation of perceptual and mnemonic representations. Neurobiology of Aging.

Doherty-Sneddon G, & Phelps FG (2005). Gaze aversion: A response to cognitive or social difficulty? Memory & Cognition, 33(4), 727–733. [PubMed: 16248336]

Ehrlichman H, & Micic D (2012). Why do people move their eyes when they think? Current Directions in Psychological Science, 21(2), 96–100.

Green P, MacLeod CJ (2016). simr: an R package for power analysis of generalised linear mixed models by simulation. Methods in Ecology and Evolution, 7(4), 493–498. doi: 10.1111/2041-210X.12504, https://CRAN.R-project.org/package=simr.

Harel J, Koch C, & Perona P (2006). Graph-based visual saliency. Proceedings of Neural Information Processing Systems (NIPS), 19, 545–552.

Hayes TR,& Henderson JM (2019). Center bias outperforms image salience but not semantics in accounting for attention during scene viewing. Attention, Perception, & Psychophysics, 1–10.

Hayes TR, & Henderson JM (2021a). Looking for semantic similarity: what a vector-space model of semantics can tell us about attention in real-world scenes. Psychological Science, 32(8), 1262–1270. 10.1177/0956797621994768 [PubMed: 34252325]

Hayes TR, & Henderson JM (2021b). Deep saliency models learn low-, mid-, and high-level features to predict scene attention. Scientific Reports, 11.

Henderson JM, & Hayes TR (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. Nature Human Behaviour, 1(10), 743.

Henderson JM, Hayes TR, Rehrig G, & Ferreira F (2018). Meaning guides attention during real-world scene description. Scientific Reports, 8, 13504. [PubMed: 30202075]

James EL, & Kooy TM (2011). Aging and the detection of visual errors in scenes. Journal of Aging Research, 2011, 984694. [PubMed: 22007300]

James G, Witten D, Hastie T, & Tibshirani R (eds.). (2013). An introduction to statistical learning: with applications in R. New York: Springer.

Kramer AF, Hahn S, Irwin DE, & Theeuwes J (2000). Age differences in the control of looking behavior: Do you know where your eyes have been? Psychological Science, 11, 210–217. [PubMed: 11273405]

Kramer AF, Irwin DE, Theeuwes J, & Hahn S (1999). Oculomotor capture by abrupt onsets reveals concurrent programming of voluntary and involuntary saccades. Behavioral and Brain Sciences, 22(4), 689–690.

Mackworth NH, & Morandi AJ (1967). The gaze selects informative details within pictures. Perception & Psychophysics, 2, 547–552.

Madden DJ, Whiting WL, Cabeza R, & Huettel SA (2004). Age-related preservation of top-down attentional guidance during visual search. Psychology and aging, 19(2), 304. [PubMed: 15222823]

McLaughlin PM, Borrie MJ, & Murtha SJ (2010). Shifting efficacy, distribution of attention and controlled processing in two subtypes of mild cognitive impairment: Response time performance and intraindividual variability on a visual search task. Neurocase, 16(5), 408–417. [PubMed: 20401803]

Molitor RJ, Ko PC, & Ally BA (2015). Eye movements in Alzheimer's disease. Journal of Alzheimer's Disease, 44(1), 1–12.

Nuthmann A, Einhäuser W, & Schültz I (2017). How well can saliency models predict fixation selection in scenes beyond central bias? A new approach to model evaluation using generalized linear mixed models. Frontiers in Human Neuroscience, 11, 491. [PubMed: 29163092]

Nyberg L, Salami A, Andersson M, Eriksson J, Kalpouzos G, Kauppi K, Lind J, Pudas S, Persson J, & Nilsson LG (2010). Longitudinal evidence for diminished frontal cortex function in aging. Proceedings of the National Academy of Sciences, 107(52), 22682–22686.

Ogletree AM, & Katz B (2020). How do older adults recruited using MTurk differ from those in a national probability sample?. The International Journal of Aging and Human Development, 0091415020940197.

Olincy A, Ross RG, Youngd DA, & Freedman R (1997). Age diminishes performance on an antisaccade eye movement task. Neurobiology of Aging, 18(5), 483–489. [PubMed: 9390774]

Peacock CE, Hayes TR, & Henderson JM (2019a). The role of meaning in attentional guidance during free viewing of real-world scenes. Acta Psychologica, 198, 102889. [PubMed: 31302302]

Peacock CE, Hayes TR, & Henderson JM (2019b). Meaning guides attention during scene viewing, even when it is irrelevant. Attention, Perception, & Psychophysics, 81, 20–34.

Peacock CE, Cronin DA, Hayes TR, & Henderson JM (2021). Meaning and expected surfaces combine to guide attention during visual search in scenes. Journal of Vision, 21(11):1, 1–18.

Pereira MLGDF, Camargo MVZDA, Bellan AFR, Tahira AC, Dos Santos B, Dos Santos J, … & Forlenza OV (2020). Visual search efficiency in mild cognitive impairment and Alzheimer's disease: an eye movement study. Journal of Alzheimer's Disease, 75(1), 261–275.

Pollmann S, Rosenblum L, Linnhoff S, Porracin E, Geringswald F, Herbik A, Renner K, et al. (2020). Preserved contextual cueing in realistic scenes in patients with age-related macular degeneration. Brain Sciences, 10(12), 941. [PubMed: 33297319]

Rabbitt P (1965). Age and discrimination between complex stimuli. Behavior, aging, and the nervous system, 35–53.

Ramzaoui H, Faure S, & Spotorno S (2021). Top-down and bottom-up guidance in normal aging during scene search. Psychology and Aging, 36(4), 433–451. [PubMed: 34124920]

Rehrig G, Hayes TR, Henderson JM, & Ferreira F (2020). When scenes speak louder than words: Verbal encoding does not mediate the relationship between scene meaning and visual attention. Memory & Cognition, 48(7), 1181–1195. [PubMed: 32430889]

Rehrig G, Peacock CE, Hayes TR, Henderson JM, & Ferreira F (2020). Where the action could be: Speakers look at graspable objects and meaningful scene regions when describing potential actions. Journal of Experimental Psychology: Learning, Memory and Cognition, 46(9), 1659–1681. [PubMed: 32271065]

Rehrig G, Hayes TR, Henderson JM, & Ferreira F (2022, October 26). Visual attention during seeing for speaking in healthy aging. Retrieved from https://osf.io/shdbz/.

Ryan JD, Shen J, & Reingold EM (2006). Modulation of distraction in ageing. British Journal of Psychology, 97(3), 339–351. [PubMed: 16848947]

Ryan JD, Leung G, Turk-Browne NB, & Hasher L (2007). Assessment of age-related changes in inhibition and binding using eye movement monitoring. Psychology and Aging, 22(2), 239. [PubMed: 17563180]

Steyvers M, Griffiths TL, & Dennis S (2006). Probabilistic inference in human semantic memory. Trends in Cognitive Science, 10(7), 327–334.

Research SR (2017). EyeLink 1000 Plus User Manual, Version 1.0.2. Mississauga, ON: SR Research Ltd.

Tatler BW (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. Journal of Vision, 7(14):4, 1–17.

Tenny S, Hoffman MR. Odds ratio. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2021.

Umanath S, & Marsh EJ (2014). Understanding how prior knowledge influences memory in older adults. Perspectives on Psychological Science, 9(4), 408–426. [PubMed: 26173273]

Wais PE, Rubens MT, Boccanfuso J, & Gazzaley A (2010). Neural mechanisms underlying the impact of visual distraction on retrieval of long-term memory. The Journal of Neuroscience, 30(25), 8541–8550. [PubMed: 20573901]

Williams CC, Henderson JM, & Zacks RT (2005). Incidental visual memory for targets and distractors in visual search. Perception & Psychophysics, 67, 816–827. [PubMed: 16334054]

Williams CC, Zacks RT, & Henderson JM (2009). Age differences in what is viewed and remembered in complex conjunction search. Quarterly Journal of Experimental Psychology, 62, 946–966.

Wynn JS, Bone MB, Dragan MC, Hoffman KL, Buchsbaum BR, & Ryan JD (2016). Selective scanpath repetition during memory-guided visual search. Visual Cognition, 24, 15–37. [PubMed: 27570471]

Wynn JS, Ryan JD, & Moscovitch M (2019). Effects of prior knowledge on active vision and memory in young and older adults. Journal of Experimental Psychology: General, 149(3), 518–529. [PubMed: 31343184]

Wynn JS, Amer R, & Schacter DL (2020). How older adults remember the world depends on how they see it. Trends in Cognitive Sciences, 24(11), 858–861. [PubMed: 32863143]

Yang Q, Wang T, Su N, Liu Y, Xiao S, & Kapoula Z (2011). Long latency and high variability in accuracy-speed of prosaccades in Alzheimer's disease at mild to moderate stage. Dement Geriatr Cogn Dis Extra, 1, 318–329. [PubMed: 22203824]

Yang Q, Wang T, Su N, Xiao S, & Kapoula Z (2012). Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnestic mild cognitive impairment. Age, 35, 1287–1298. [PubMed: 22576337]

Public significance statement:

Older adults have more knowledge from lived experience about everyday scenes, such as what might typically be found in the rooms in a house, and past studies have shown that older adults rely more on their scene knowledge (semantic information) than young adults do when searching for objects in photographs of scenes. In our study, we tracked where young and older adults looked in photographs of everyday environments while they talked about the scenes rather than searching for objects in them. We found that semantic information predicted where young adults looked better than it predicted where older adults looked, although the parts of the scene older adults looked at were also predicted well by semantic information independently. Our results suggest that older adults may not leverage their considerable knowledge about the world equally across visual tasks (searching for an object in a photograph vs. deciding what to describe in a photograph), and should not be assumed to do so more than their young adult counterparts across the board. These findings suggest it is important to be cautious when assuming that healthy older adults' greater knowledge base is routinely consulted to compensate for other age-related changes, such as declines in perceptual skills.

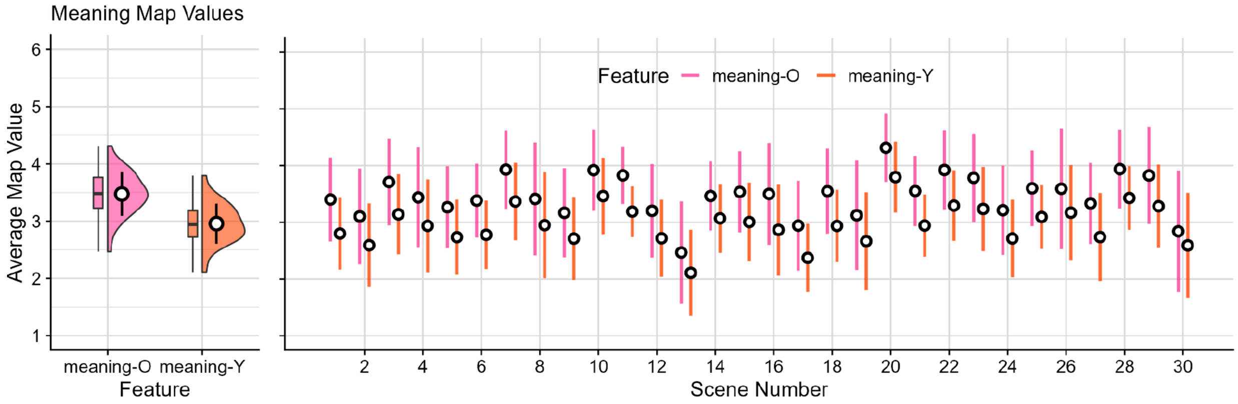


Figure 1.

Left: Hybrid violin and box plots showing the average meaning map-O (pink), meaning map-Y (orange) values computed for each scene ($N = 30$). White points superimposed over the violins indicate the grand mean, and black vertical lines indicate the standard deviation. On the box plots to the left of each violin, black horizontal lines correspond to the median, colored boxes indicate the 25% and 75% quartile boundaries, and black vertical lines show ± 1.5 IQR (the interquartile range). Right: Points showing average meaning map-O and -Y values for each scene individually (x-axis). Error bars indicate the standard deviation.

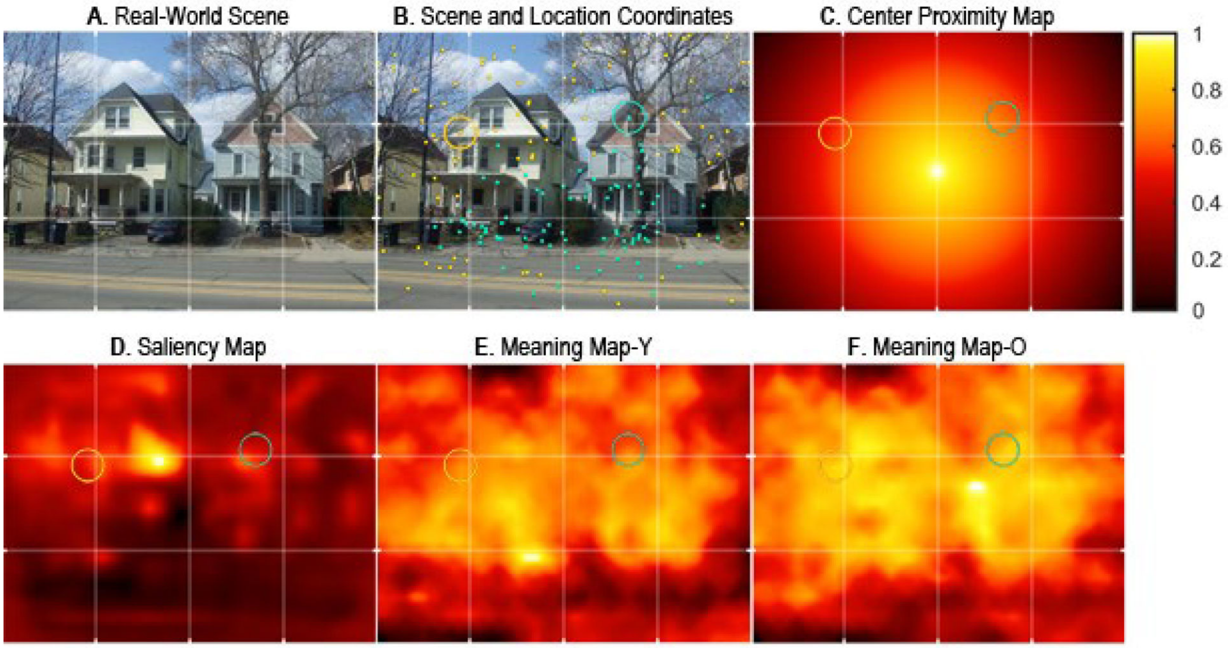


Figure 2.

Visualization of analysis approach. A) Real-world scene, B) Scene overlaid with fixated (cyan) and randomly sampled (yellow) location coordinates. Circles overlaid on B-F illustrate the mask radius used to compute average feature map values around each fixated (cyan) or sampled (yellow) coordinate. C) Center proximity map that was used to compute average center proximity values. D) Saliency map for the scene shown in A. (E) Meaning map-Y generated from young adult raters and (F) meaning map-O generated from older adult raters for the scene shown in A.

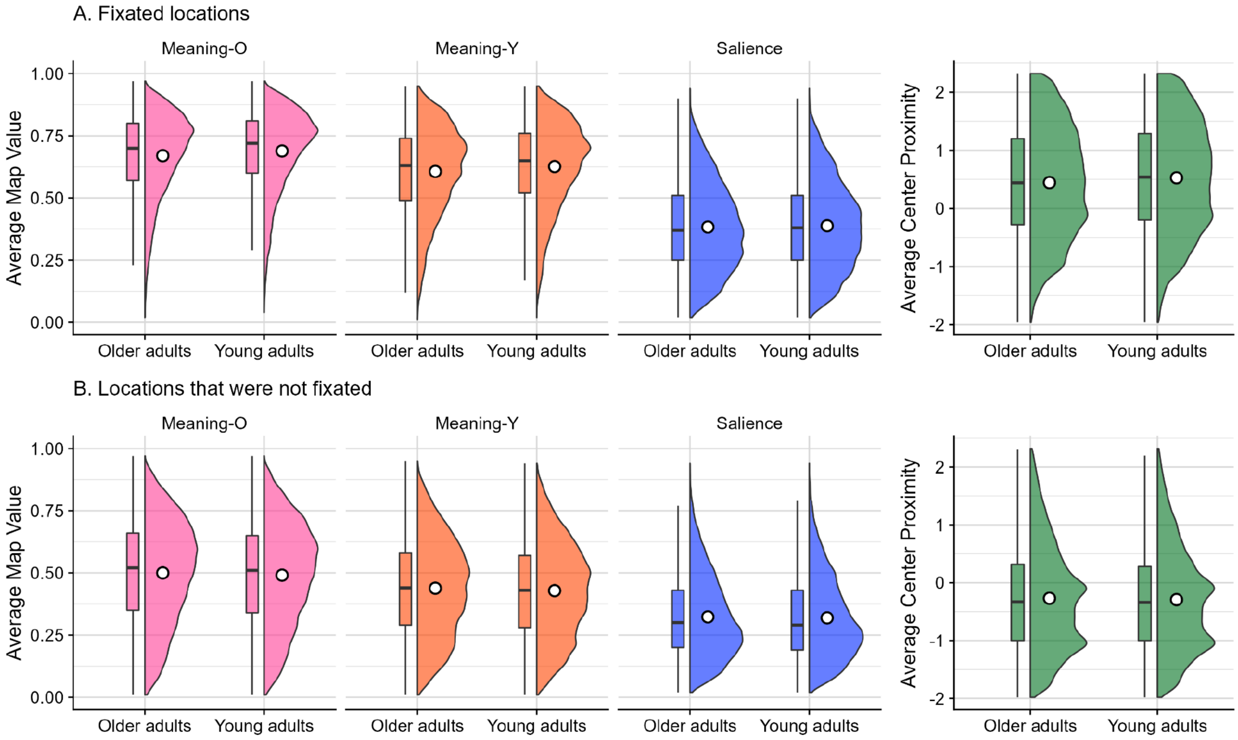


Figure 3.

Hybrid violin and box plots showing average meaning map-O (pink), meaning map-Y (orange), saliency map (blue), and center proximity (green) values for older and young adults (x-axis) corresponding to A) the region surrounding fixation coordinates and B) the region surrounding randomly sampled locations in the scene that were not fixated. White points superimposed over the violins indicate the grand mean. On the box plots to the left of each violin, black horizontal lines correspond to the median, colored boxes indicate the 25% and 75% quartile boundaries, and black vertical lines show ± 1.5 IQR (the interquartile range). Only meaning map values differed across age groups in the model (see Table 2).

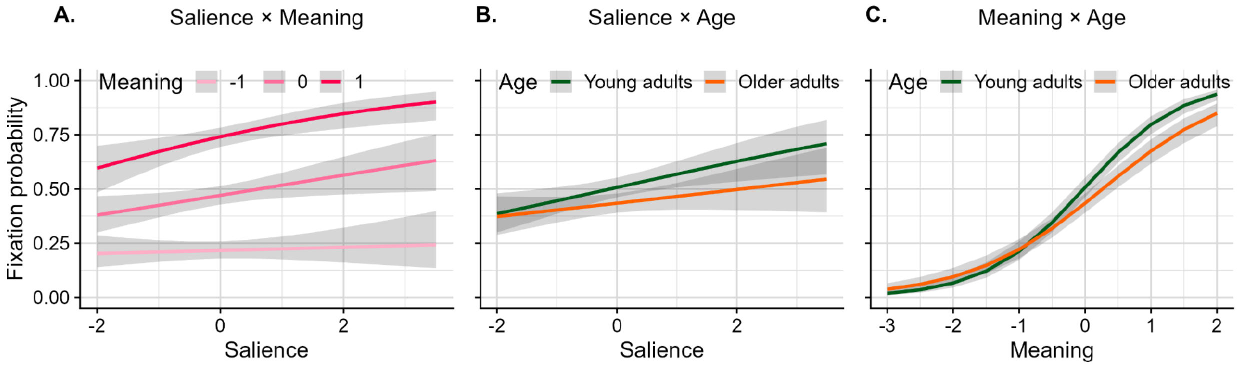


Figure 4.

Estimated fixation probability (y-axis) for each marginal or reliable interaction. Shaded gray regions indicate 95% confidence intervals. A) Interaction between meaning values (lines) and salience (x-axis), B) Marginal interaction between age (lines) and salience (x-axis), and C) Interaction between age group (older adults = orange lines, young adults = green lines) and meaning (x-axis).

Table 1

Correlations (R^2) Between Maps

Map Comparison	**Correlation (R^2)** ***M***	 ***SD***
Meaning map-Y × Meaning map-O	0.849	0.053
Meaning map-Y × Saliency map	0.126	0.101
Meaning map-O × Saliency map	0.129	0.092

Table 2

Logistic Mixed-Effects Regression Table for Locations Fixated

	Fixed effects				Random effects (SD)	
Effect	***β***	***SE***	***z***	***p***	***Subject***	***Scene***
Intercept	−0.12	0.09	−1.37	0.17	0.09	0.47
Group	−0.15	0.03	−4.63	< .0001	—	0.16
Meaning	1.17	0.08	14.79	< .0001	0.32	0.37
Salience	0.19	0.08	2.34	0.02	0.11	0.43
Center Proximity	0.50	0.08	5.94	< .0001	0.30	0.41
Group:Meaning	−0.17	0.05	−3.57	0.0004	—	0.13
Group:Salience	−0.06	0.03	−1.84	0.07	—	0.15
Meaning:Salience	0.14	0.06	2.44	0.01	0.08	0.32
Group:Center Proximity	−0.04	0.05	−0.94	0.35	—	0.12
Center Proximity:Meaning	−0.01	0.06	−0.21	0.83	0.12	0.29
Center Proximity:Salience	−0.08	0.06	−1.40	0.16	0.08	0.30
Group:Meaning:Salience	−0.009	0.03	−0.27	0.79	—	0.17
Group:Center Proximity:Meaning	−0.02	0.03	−0.72	0.47	—	0.11
Group:Center Proximity:Salience	0.03	0.02	1.20	0.23	—	0.09
Center Proximity:Meaning:Salience	0.03	0.04	0.57	0.57	0.07	0.23
Group:Center Proximity:Meaning:Salience	0.02	0.02	1.03	0.30	—	0.09

Table 3

Logistic Mixed-Effects Regression Table for Locations Fixated by Older Adults Only

	Fixed effects				**Random effects (SD)**	
Effect	β	***SE***	***z***	***p***	***Subject***	***Scene***
Intercept	−0.13	0.07	−1.75	0.08	0.08	0.40
Meaning	0.98	0.09	11.37	< .0001	0.26	0.39
Salience	0.15	0.08	1.77	0.08	0.12	0.44
Center Proximity	0.46	0.09	5.11	< .0001	0.32	0.37
Meaning:Salience	0.13	0.06	2.29	0.02	0.07	0.29
Center Proximity:Meaning	−0.03	0.06	−0.57	0.57	0.11	0.28
Center Proximity:Salience	−0.04	0.06	−0.64	0.52	0.08	0.30
Center Proximity:Meaning:Salience	0.04	0.04	1.08	0.28	0.06	0.20

Table 4

Logistic Mixed-Effects Regression Table for Locations Fixated by Young Adults Only

	Fixed effects				**Random effects (SD)**	
Effect	***β***	***SE***	***z***	***p***	***Subject***	***Scene***
Intercept	−0.15	0.11	−1.38	0.17	0.11	0.56
Meaning	1.34	0.10	13.57	< .0001	0.36	0.39
Salience	0.23	0.09	2.55	0.01	0.09	0.47
Center Proximity	0.54	0.10	5.31	< .0001	0.30	0.46
Meaning:Salience	0.16	0.08	1.97	0.049	0.08	0.42
Center Proximity:Meaning	0.01	0.07	0.08	0.93	0.14	0.32
Center Proximity:Salience	−0.11	0.06	−1.75	0.08	0.08	0.33
Center Proximity:Meaning:Salience	0.001	0.06	0.01	0.99	0.07	0.28