Age-related differences during visual search: the role of contextual expectations and cognitive control mechanisms†.

Miguel T. Borges

Eunice G. Fernandes

Moreno I. Coco

Max Planck Institute for Psycholinguistics, University of Lisbon, University of East London.

Author Note

Miguel T. Borges, Max Planck Institute for Psycholinguistics; Eunice G. Fernandes, Department of Linguistics, University of Lisbon; Moreno I. Coco, School of Psychology, University of East London.

Water Lane, E15 4LZ, School of Psychology, University of East London.

This work was supported by the Leverhulme Trust with an Early Career Research Fellowship (ECF-2014-205) awarded to MIC.

Corresponding Author: moreno.cocoi@gmail.com

†All authors have equally contributed to this work

Age-related differences during visual search: the role of contextual expectations and cognitive control mechanisms.

During the visual search, cognitive control mechanisms activate to inhibit distracting information and efficiently orient attention towards contextually relevant regions likely to contain the search target. Cognitive ageing is known to hinder cognitive control mechanisms, however little is known about their interplay with contextual expectations, and their role in visual search. In two eye-tracking experiments, we compared the performance of a younger and an older group of participants searching for a target object varying in semantic consistency with the search scene (e.g., a basket of bread vs. a clothes iron in a restaurant scene) after being primed with contextual information either congruent or incongruent with it (e.g., a restaurant vs. a bathroom). Primes were administered either as scenes (Experiment 1) or words (Experiment 2, which included scrambled words as neutral primes). Participants also completed two inhibition tasks (Stroop and Flanker) to assess their cognitive control. Older adults had greater difficulty than younger adults when searching for inconsistent objects, especially when primed with congruent information (Experiment 1), or a scrambled word (neutral condition, Experiment 2). When the target object violates the semantics of the search context, congruent expectations or perceptual distractors, have to be suppressed through cognitive control, as they are irrelevant to the search. In fact, higher cognitive control, especially in older participants, was associated with better target detection in these more challenging conditions, although it did not influence eye-movement responses. These results shed new light on the links between cognitive control, contextual expectations and visual attention in healthy ageing.

Keywords: visual search; cognitive ageing; contextual expectations; cognitive control; eye-tracking.

Introduction

Cognitive ageing has been linked to degraded performance in a variety of visual tasks such as driving, face recognition, or searching for a target among distractors (Bédard et al., 2006; Burton-Danner, Owsley, & Jackson, 2001; Foster, Behrmann, & Stuss, 1995; Hahn, Carlson, Singer, & Gronlund, 2006; Potter, Grealy, Elliott, & Andres, 2012). Older adults take longer to identify a target object (e.g., more saccades; Watson, Maylor, & Bruce, 2005), spend more time processing it (e.g., longer gaze duration; Williams, Zacks, & Henderson, 2009), and display a worse performance than younger adults when asked to remember or to search for a target embedded within other distractor objects (see Zanto & Gazzaley, 2014 for a review).

Beyond a generalised slowing of cognitive processing (e.g., Salthouse, 1994, 2012), most research on ageing points to cognitive control mechanisms to explain the difficulties often experienced by healthy older adults (e.g., Gazzaley, Cooney, Rissman, & D'Esposito, 2005). In visual search, for example, irrelevant information within the visual context has to be inhibited to orient attention towards the target object efficiently, and this operation relies on cognitive control that can be impaired in older adults (e.g., Foster et al., 1995, Schwarzkopp, Mayr, & Jost, 2016; but see Monge & Madden, 2016 for an alternative explanation based on perceptual degradation). Braver, Gray, and Burgess (2007) suggest a dual account of cognitive control mechanisms by distinguishing between proactive and reactive control. Proactive control refers to a slow and effortful mechanism that acts globally to mitigate interference before perceptual processing, which is mostly affected by ageing (e.g., Mayr, 2001). Reactive control is instead a fast and automatic mechanism that acts contingent on stimulus onset, and seems spared by ageing (e.g., Bugg, 2014).

The interplay between proactive control and access to contextual information during the visual search has been examined, mostly in young adults, by using contextual cueing paradigms (e.g., Chun, 2000). Chun and Jiang (1998) found that young observers were faster in searching a target letter T among Ls distractors when previously and repeatedly exposed to a specific spatial layout of the letters; and older adults seem to equally benefit by this pre-activation (e.g. Howard, Dennis, Howard, Yankovich, & Vaidya, 2004; Madden, Whiting, Cabeza, & Huettel, 2004; Whiting, Madden, Pierce, & Allen, 2005; but see Smyth & Shanks, 2011, for impaired learning of contextual information in older adults). Other studies on young adults have extended the findings of contextual cueing in visual search from arrays of visual objects to photo-realistic scenes, and largely corroborated the evidence that the pre-activation of a scene facilitates search (e.g., Hollingworth, 2007). However, they are facilitated only when the cue scene and the search scene vary in the local context surrounding the target but not when they instead vary in the global context (e.g., Brockmole, Castelhano, & Henderson, 2006; Chun, 2000; Hollingworth, 2009). When the scenes have different spatial layouts despite being conceptually related (e.g., two different libraries) or if the cue scene and the search scene differ in low-level features, such as colour, there is also no benefit (e.g., Goujon, Brockmole, & Ehinger, 2012).

Realistic scenes provide rich contextual information (e.g, a restaurant) that combined with the information about the search target (e.g., a basket of bread) is used by the cognitive system to guide overt attention to locations (e.g., a table) that are more likely to contain it (e.g., Bar, 2004; Eckstein, Drescher, & Shimozaki, 2006; Madden, 2007; Neider & Zelinsky, 2006). When the contextual relation between the scene and the objects is disrupted, such as a restaurant scene with a clothes iron in it, then this inconsistent object may be prioritized, e.g., attract earlier fixations compared to a consistent object (e.g., Belke, Humphreys, Watson, Derrick, Meyer, & Telling, 2008; LaPointe & Milliken, 2016; Loftus & Mackworth, 1978; Underwood, Humphreys, & Cross, 2007); and once identified, it will require a more effortful processing (e.g., longer fixation duration, Cornelissen & Võ, 2017; Henderson, Weeks, Phillip a., & Hollingworth, 1999; Võ & Henderson, 2011; but see Wu, Wick, & Pomplun, 2014 for a review of open controversies in the literature).

Very little research in cognitive ageing has looked at possible differences in visual search strategies between younger and older adults associated with contextually richer visual contexts. One exception is the study by Neider and Kramer (2011), who found that older adults relied more than younger adults on the spatial associations between the context of a pseudo-realistic scene and the target of the search (e.g., searching for a jeep on the ground) to guide their overt attention. Older adults may, in fact, over-rely on contextual information to compensate for the age-related decline in cognitive processing (e.g., Lindenberger & Mayr, 2014; Madden, 2007). Note, however, that Neider and Kramer (2011) did not explicitly manipulate contextual expectations using cueing, nor examined their interplay with explicit measures of cognitive control.

The current visual search eye-tracking study employs a contextual cueing paradigm to precisely investigate how the pre-activation of contextual information may differently modulate the identification of a target object embedded in (and consistent or inconsistent with) a naturalistic scene in older and younger adults. Our main aim is to investigate the inhibitory cognitive control strategies adopted by the two age groups to cope with irrelevant information promoted by the type of priming and the semantics of the target object. Our main goal is to seek for direct evidence for (or against) the role of cognitive control in visual search while highlighting possible age-related differences.

INSERT FIGURE 1 HERE

The present study

We report two eye-tracking visual search experiments in which participants (an older and a younger group) are initially prompted with a target cue word and then briefly primed with contextual information, either a visual scene (e.g., a picture of a restaurant, Experiment 1) or a category word (e.g., the word "restaurant", Experiment 2). Then, they are presented with a search scene where they have to find the target object, which can be either consistent or inconsistent with it (e.g., a bread basket versus an iron); see Figures 1 and 2 for the experimental design and an example of a trial. The primed information can be either congruent or incongruent with the search scene (or also neutral, but only in Experiment 2). In a trial of Experiment 1, for example, a participant may be exposed to a restaurant scene, and then have to search for a clothes iron in a different restaurant scene that is visually distinct (i.e., it is not the same scene, and they differ in low-level visual features). In this example, the experimental condition has a prime scene conveying contextually congruent information and a target object that is inconsistent with the search scene. The same structure applies to Experiment 2, in which we replace prime scenes with prime words conveying contextual information, and add a 'neutral' priming condition operationalised as a scrambled word, carrying no lexical meaning.

We compare the performance of a younger and a healthy older group on three critical measures of the search task: (a) response accuracy, which is the proportion of correctly detected target objects; and on correct trials only (b) the time to first fixation, which is the time it takes for a participant to look at the target object for the first time from scene onset, and points at the speed of target identification; and (c) the total fixation time, which is the sum of all fixations allocated to the target object, and points at the processing effort required to correctly recognize the target of the search. Each group is assessed on general cognitive abilities using the Montreal Cognitive Assessment Test (MoCA; Portuguese Version normed by Freitas, Simões, Alves, & Santana, 2011). This step is necessary to screen older participants for potential cognitive impairment resulting from neurodegenerative disease. In the analysis, we use linear-mixed effect modelling to capture the variability in baseline performance across our participants better, and z-score all time-related variable to account for general-slowing of the older group (see Method for more details). Furthermore, we test both younger and older participants on two classic tests of executive control, the Stroop and the Flanker tasks (see Rey-Mermet & Gade, 2017 for a recent review). A summary measure of the participants' performance on these two tasks, obtained using Principal Component Analysis, is then correlated with the measures of visual search discussed above. We do so to obtain direct evidence for, or against, interdependent mechanisms such as lower cognitive control leading to worsening accuracy in the visual search task. This analysis represents an essential addition to previous research on ageing, which has often argued about impairment of cognitive control mechanisms in older adults without seeking for cross-validation across tasks. In this final analysis, we use the data from both Experiment 1 and Experiment 2 to obtain a greater statistical power (i.e., more participants) as well as to compare the influence of the two types of contextual priming (scenes vs words) on visual search performance.

Our predictions

In both experiments, we expect participants to be faster at identifying the target object when it is inconsistent with the search scene. This prediction stems from previous findings which revealed that as the semantic information of objects may be processed in extra-foveal vision, inconsistent objects that violate the contextual fit of the scene attract attention earlier than consistent objects (e.g., LaPointe & Milliken, 2016; Loftus & Mackworth, 1978). Going beyond this literature, if older participants rely more on contextual expectations, then we expect them to find it more difficult to detect inconsistent objects than consistent objects, as they violate the contextual expectations of the scene by definition. Furthermore, we expect this difficulty to increase when the context is boosted through priming. Critically, if the poorer performance in visual search also stems from the inability of older adults to suppress irrelevant information (Lustig & Jantz, 2015; Zanto & Gazzaley, 2014), those older adults with better cognitive control should be better able to act upon irrelevant information, and consequently be more accurate.

In terms of eye-movement responses, one may intuitively expect that searching for an object in a restaurant scene after being primed with another restaurant scene should be facilitated relative to searching after being primed with an unrelated scene such as a pool hall. However, the two restaurant scenes crucially differ in the spatial layout of objects, which is acquired very quickly (e.g., Henderson & Ferreira, 2004), as well as in low-level visual features. Thus, the participants may find it actually harder to disengage from specific spatially-guided expectations created when primed with the same category (e.g., restaurant) than to supress the irrelevant information of a conceptually unrelated scene (e.g., pool hall). This hypothesis would be consistent with studies showing that visual search is not facilitated when the cue scene and the search scene differ in spatial layout (e.g., Brockmole et al. 2006), and that (visually) conflicting stimuli disrupt perceptual identification more than irrelevant stimuli (e.g., Vallesi, Stuss, McIntosh, & Picton, 2009). Therefore, we expect older participants to require a more extended allocation of overt attention than younger participants to correctly identify the target object when it is inconsistent with the scene, or when primed with congruent information, as they have a reduced capacity to resolve conflicting information.

The above logic may apply as well to prime words, with the critical comparison now being between neutral and incongruent (or congruent) priming. The neutral priming condition, a scrambled mask of a context word (i.e., carrying no lexical meaning), is a perceptual distractor which may likely interfere with the retention of the target identity to a greater degree than both congruent or incongruent prime words that can instead be processed and ignored if irrelevant. We thus expect participants to display a reduced detection performance when the prime is neutral, especially when the target object is inconsistent with the scene context. In line with our predictions of Experiment 1, older adults may find this condition particularly hard due to their reduced capacity to suppress irrelevant information stemming from the perceptually disruptive nature of the neutral prime, which is a scrambled word that can disrupt their retention of the target object cue, and the semantic inconsistency of the object.

Experiment 1

Methods

Participants

We recruited 53 native speakers of Portuguese in two groups: a younger group of 32 undergraduate students (28 F; Age = 21.4 ± 5.71) and an older group of 21 participants (17 F, Age = 67.4 ± 4.97) from Universities of the Third Age with no reported history of neuropsychological deficits. All participants had normal or corrected-to-normal vision. The younger participants were awarded course credits, and the older participants were paid 15 euros for taking part in the experiment. This study was approved by the Ethics Committee of the Department of Psychology, University of Lisbon, before data collection.

Design, stimuli and procedure

In a 2x2x2 experimental design, we crossed the congruency between the Prime scene and the search scene (Congruent, Incongruent) with the semantics of the target Object with respect to the search scene (Consistent, Inconsistent) within-participant, and had a Group condition (Younger, Older) between-participants (see Figure 1 for a visualisation of the design).

We used 42 experimental scenes, 42 filler scenes and 84 prime scenes, all naturalistic images, i.e., photographs. Four different versions of each experimental scene were created, with the two types of Object consistency (consistent, inconsistent) counterbalanced in two locations (left, right) within the scene to avoid the development of systematic search strategies. We, therefore, had a total of 168 unique items, divided into four lists following a Latin-Square design. Prime and filler scenes were selected from open-access databases (e.g., Flickr's "The Commons"). The prime scenes were selected according to their contextual congruency with respect to the experimental scenes while ensuring that they would never contain the target object (refer to Appendix A for the norming of the material).

Visual stimuli were presented on a 19-inch LG Flatron L194ws LCD screen with a 60Hz refresh rate, 55cm away from participants' eyes, subtending 29.3º of visual angle. Eye movement data were recorded using an SMI IVIEW X HI-SPEED eye-tracker (SensoMotoric Instruments, Teltow, Germany) at a sampling rate of 1,250Hz. Only the participants' dominant eye (assessed using a parallax test) was tracked. Chin rest and forehead support were used to stabilise the head position. A Logitech Cordless Rumblepad 2 controller was used to record the target detection responses. The experiment was developed using Experiment Center 3.2.

INSERT FIGURE 2 HERE

Each trial started with a written cue (500ms) prompting the participant with the search target. A fixation cross then appeared in the centre of the screen, displayed until it was fixated for 1000ms, after which the prime scene was presented for 250ms followed by a scrambled mask of the same scene for another 250ms. The mask was applied to reduce the perceptual activation of the prime scene before the search scene (Enns & Lollo, 2000). After the onset of the search scene, the participants had to state whether they did, or did not, see the target object by pressing a response button on a cordless joypad (see Figure 2 for an example trial).

The target object was always present in the experimental trials and always absent in the filler trials. This distribution guarantees a balanced 50-50 distribution of yes/no responses. At the beginning of each experimental session, participants completed four practice trials to familiarise with the task. A 15-point calibration procedure was performed at the beginning of the experiment and repeated every 10 trials (on 8-points) to ensure accurate tracking. The calibration was accepted when the deviation error was below 0.8º and 1.2º degrees of visual angle for the x and y-axis, respectively. On average, the younger group displayed a 0.61º ± 0.42º visual angle deviation error on the x-axis and 0.72º ± 0.58º on the y-axis while the older group displayed a 0.82º ± 0.60º error on the x-axis and 1.12º ± 0.79º on the y-axis.

At the end of the visual search task, each participant responded to the MoCA test (30 points maximum score) and performed a computerised version of the Stroop (64 randomised trials, 32 Congruent and 32 Incongruent) and Flanker (64 randomised trials, 32 Congruent – 32 Incongruent) tasks, implemented using E-Prime (Psychology Software Tools, Pittsburgh, PA). The eye-tracking session lasted for about 30 minutes, which together with the cognitive control tasks and the MoCA totalled approximately to an hour of testing.

Analyses

Raw gaze data were initially pre-processed into fixations and saccades using BeGaze 3.2 (SMI Instruments), and further analysed using R 3.4.4 (R Core Team, Vienna, Austria). Of the 2,226 total trials (i.e., 42 scenes and 53 participants), we excluded 8 items (108 trials) because participants could not correctly identify the target object above the 50% chance level, and an additional 45 trials (by participant and experimental condition) because the response times were 2.5 SDs above the mean. Of these 2,073 trials, we retained the 1,504 trials in which the target object was fixated. For these trials, we considered the response accuracy to the visual search task, and on correct trials (i.e., 1,454) the eye-movement measures for the two main stages of target detection: identification and processing. For the identification stage, we computed the time to the first fixation of the target object from scene onset. For the processing stage, we considered the total fixation time on the object, which is the sum of all fixations on the target during the trial.

We used linear-mixed effects modelling (LME) as implemented by the `lme4 (vs 1.1-18-1)` package in `R` to analyse our dependent measures (Bates, Mächler, Bolker, & Walker, 2015). We adopted the LME approach as it allows the simultaneous modelling of by-participant and by-item random effects. This approach is advantageous for two main reasons. First, it captures the intrinsic variability of the random variables when estimating the coefficients for the fixed effects, and hence it takes into account the individual differences between, for example, the participants. Second, it avoids data aggregation, a practice which leads to less precise estimates when correlations are unequal across repeated measures and may yield biased estimates, especially with small samples (Muth et al., 2016). The predictors used for the LME analysis, centred to reduce co-linearity, were: Prime (Congruent = -.5, Incongruent = .5), Object (Consistent = -.5, Inconsistent = .5) and Group (Older = -.6, Younger = .4), and our random effects were Participant (53; entered as a between-participant variable, as participants are nested into their age group), and Scene (42). In ageing research, a well-known phenomenon is that time-related responses of older participants are consistently slower than those of the younger participants (e.g., Salthouse, 1994). It becomes, therefore, crucial to normalise them so that the general slowing does not confound effects genuinely arising from the experimental conditions of interest. We followed the approach proposed by Faust, Balota, Spieler, & Ferraro, 1999, whereby we z-scored the time-dependent eye-movement measures of time to first fixation and total time of fixation independently for each participant. Effectively, this strategy makes sure that the mean response between the two age groups is approximately the same (i.e., no main effect of Group), and that any true difference between the two groups should emerge as an interaction between the Group variable and the experimental variable of interest (e.g., Prime in our study).

Here we present the response accuracy and the time-dependent attentional measures already z-scored (but refer to the Supplementary Material S2 and S3 for plots of the same measures without any transformation). We also considered the order of trials (accounting for learning strategies) as a covariate and controlled for it by residualizing it against each of the dependent variables (DVs) in a linear regression model (`DV ~ Order`, in `R` syntax), and then taking the residuals obtained as the DVs for further analyses. This approach ensures that the effects of our experimental predictors (Prime and Object) on each DV analysed are cleaned from this incidental covariate. Our initial models had a full fixed-effect structure (i.e., all main effects and interactions) with a maximal random-effect structure (i.e., random variables are included both as intercepts and uncorrelated slopes; Barr, Levy, Scheepers, & Tily, 2013). Then, in order to have models that were parsimonious in the number of parameters (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017), we reduced them utilising the `step` function of the `R` package `lmerTest` (vs 3.0-1; Kuznetsova, Christensen, Bavay, & Brockhoff, 2014). This method performs a backward selection of the model, both on the random and fixed effect structure, iteratively removing terms that do not significantly improve the model fit (thresholds of $p < 0.1$ for random effects and $p < 0.08$ for fixed effects to include marginally significant results). Table 1 reports the coefficients, standard errors, and t-values only for those predictors that were significant after model selection. We derived p-values for the fixed effects in the LME models from F-tests based on Satterthwaite approximation of the effective degrees of freedom. We also report the standardised βs, i.e., coefficients that have been normalised to be directly comparable

Results

Results are organised into two sections: (a) "General cognitive abilities and cognitive control", in which we compare the performance of the two groups on the MoCA and the Stroop and Flanker tasks, and (b) "Visual search analysis", in which we compare the two groups on the visual search task.

INSERT TABLE 1 HERE

General cognitive abilities

We assessed the general cognitive ability of the younger and the older participants with the MoCA, and their cognitive control mechanisms with the Stroop and Flanker tasks (see Table 1 for mean and standard errors of these measures). On the MoCA, the older group was significantly worse than the younger group ($t [30.3] = -4.43$, $p < 0.001$).

For the Stroop task, on a total of 3,520 trials, we excluded 41 trials for the younger group and 73 trials for the older group because they were either below the 1st or above the 99th percentile of the RT distribution for their age group. A mixed ANOVA analysis of the accuracy revealed a significant main effect of Congruency, with less accurate performances in incongruent than congruent trials [$F(1,51)= 4.6$; $p = 0.03$], and a marginal interaction between Group and Congruency, with older adults performing worse in incongruent trials [$F(1,51)= 3.57$; $p = 0.06$], and no main effect of Group. On the response time, we found significant main effects of Group, with older adults being slower than young adults [$F(1,51)= 111.38$; $p < 0.0001$], and Congruency, with longer response times for incongruent than congruent trials [$F(1,51)= 91.37$; $p < 0.0001$]. We also found a significant two-way interaction between Group and Congruency with older adults being slower than younger adults in incongruent trials [$F(1,51)= 21.96$; $p < 0.0001$].

For the Flanker task, on a total of 3,520 trials, we excluded 41 trials for the younger group and 30 trials for the older group that were either below the 1st or above the 99th percentile of the RT distribution. We only found a main effect of Congruency on the accuracy, with less accurate performance in incongruent than congruent trials [$F(1,51)= 327.85$; $p < 0.0001$]. For the response time, we found significant main effects of Group, with older adults being slower than younger adults [$F(1,51)= 30.49$; $p < 0.0001$]; and Congruency, with incongruent trials responded slower than congruent trials [$F(1,51)= 6.4$; $p = 0.01$]. We did not find a significant interaction between Congruency and Group.

These results largely confirm previous literature by showing increased difficulty in inhibiting incongruent information, especially by older participants. It must be noted, however, that this age-related interaction manifested in the Stroop but not in the Flanker task highlights a difference on the underlying inhibitory mechanisms they tap on (e.g., Spieler, Balota, & Faust, 1996).

INSERT FIGURE 3 AND TABLE 2 HERE

Visual search analysis

On response accuracy (Figure 3, upper panel), we found significant main effects of Prime, with lower accuracy when the prime scene and the search scene were semantically congruent, and Group, with younger participants being more accurate than older participants (refer to Table 2 for the model output). We also observed significant two-way interactions between Prime and Object, with lower accuracy for inconsistent objects in congruently primed scenes; and between Object and Group, with older participants committing more errors when the object was inconsistent with the scene, primarily when the prime scene was semantically congruent with the search scene (three-way interaction between Prime, Object and Group).

On time to the first fixation to the target object (Figure 3, middle panel), we only found a main effect of Object, with inconsistent objects located faster than consistent objects. When looking at the total fixation time on the object (Figure 3, lower panel), we found that inconsistent objects were looked at for longer than consistent objects (main effect of Object), especially by older participants (interaction Object, Group). Older participants also looked at the target object for longer when the prime scene and the search scene were contextually congruent (interaction Prime, Group).

Discussion

In Experiment 1, all participants had greater difficulty in correctly identifying the target object when the prime scene and the search scene conveyed the same contextual information (i.e., they were two restaurant scenes), and especially so when the target object was inconsistent with such contextual information. We argue that congruent priming may make target detection harder because: (a) the spatial layout of objects between the prime scene and the search scene must be reassessed, which is consistent with Brockmole et al. (2006) who observed no search benefit when scenes differed in global spatial layout; and (b) the activation of distractor objects consistent with the search context is further increased with respect to an inconsistent object, hence making it an even less likely target. In line with this reasoning, older adults had greater difficulties than younger adults to find an inconsistent object, especially when the prime scene and the search scene shared the same contextual information. Such difficulty may be explained by their over-reliance on contextual expectations, and by their reduced capacity to suppress it when these are irrelevant. In fact, as we shall see later in the section "*Individual differences in cognitive control and their impact on visual search*", older adults with better cognitive control were also more capable of finding inconsistent targets.

Concerning the eye-movement data, we found that the older participants looked earlier and for longer at inconsistent objects than at consistent objects. These results support the finding that the semantics of objects are, at least partly, accessed in extra-foveal vision (e.g., Belke et al., 2008; LaPointe & Milliken, 2016) and that the integration of object information violating the scene context requires a more effortful processing (e.g., Henderson et al., 1999; Williams et al., 2009). In addition to bolstering previous literature, we add that older participants display an even more effortful processing than younger adults to integrate object-scene information when the object is inconsistent, or when primed with a congruent scene. Arguably, this happens because they are required to resolve the conflict generated by the two scenes having different spatial layouts while sharing the same semantic context.

A visual scene is a cue that conveys not just conceptual, but also perceptual information about the context. It remains to be understood whether, and to what extent, a purely conceptual prime would impact on search performance. Moreover, it would be essential to examine the case where participants are primed with information that acts as a purely perceptual distractor, and it is hence "neutral" for both the scene context and the semantics of the target object. In Experiment 2, we address these questions by using words instead of scenes as the priming cue, and by including a neutral prime condition consisting of a scrambled word that does not convey any conceptual meaning.

Experiment 2

Methods

Participants

We recruited two new groups of native speakers of Portuguese: a younger group of 24 undergraduate participants (15 F; Age = 24 ± 7.2) and an older group of 24 participants (16 F, Age = 66.6 ± 4) with no recorded history of neuropsychological deficits. All participants had normal or corrected-to-normal vision. As in Experiment 1, the younger participants were awarded course credits, and the older participants were paid 15 Euro for taking part in the experiment.

Design, stimuli and procedure

The 2x2x3 design of Experiment 2 mostly followed that of Experiment 1. The only difference was that instead of using scenes as Primes, we now used words as Primes (i.e. the word "Restaurant" instead of the image of a restaurant), and included a neutral prime (i.e., a scrambled word) along with the congruent and incongruent primes. The other experimental conditions were the same as in Experiment 1 with a within-participant Object variable (Consistent, Inconsistent) and a between-participant Group variable (Older, Younger; see Figure 1 for the experimental design) crossed.

We used 62 experimental scenes and 62 filler scenes, all naturalistic images and used written prime words. Additionally, we added a neutral prime condition, consisting of a scrambled version of the incongruent or congruent word prime. Compared to Experiment 1, we added 20 experimental trials to account for the neutral prime condition, along with a further 20 filler trials to preserve an equal number of target-present and target-absent trials. The additional experimental and filler images belong to a database used by Coco, Nuthmann, & Della Sala (2017), a study also investigating object consistency using naturalistic images (refer to Appendix A for the norming of the material).

The procedure of Experiment 2 was identical to that of Experiment 1, except that the prime was now a word. When the prime was neutral, and therefore already a scrambled word, we showed a different shuffle of the same scrambled mask so that they would be visually different (see Figure 2). Participants were instructed using the same protocol as for Experiment 1, and we used the same eye-tracking calibration procedure and deviation thresholds. On average, the younger group displayed a $0.34^{o} \pm 0.17^{o}$ visual angle deviation error on the x-axis and $0.31^{o} \pm 0.16^{o}$ on the y-axis while the older group displayed a 0.48^{o}

± 0.37º error on the x-axis and 0.36º ± 0.26º on the y-axis. Participants also undertook the MoCA and the same computerised version of the Stroop and Flanker tasks.

Analyses

As in Experiment 1, raw gaze data were pre-processed into fixations and saccades using BeGaze 3.7 and further analysed using R 3.4.4. Of the 2,976 total trials (i.e., 62 scenes and 48 participants), we excluded 9 items (120 trials) because participants could not correctly identify the target object above the 50% chance level, and an additional 75 trials (by participant and experimental condition) because the response times were 2.5 SDs above the mean. Of these 2,781 trials, we retained the 2,419 trials in which the target object was fixated. We considered the same dependent measures of Experiment 1: (a) response accuracy to the visual search task and, on correct trials (i.e., 2,302), (b) the time to the first fixation of the target object from scene onset, and (c) the total fixation time on the object. We used the same modelling approach of Experiment 1 to analyse the data, and hence we refer the reader to that section for the details. The only fundamental difference in the analysis of Experiment 2 is that the Prime variable has now three levels instead of two, i.e., Congruent, Incongruent and Neutral. We have chosen Neutral as the reference level and compared Congruent and Incongruent against it.

INSERT TABLE 3 HERE

Results

General cognitive abilities

On the MoCA, the older group was significantly worse than the younger group (t [44.7] = -2.23, $p < 0.03$; refer to Table 3 for a summary of the measures).

For the Stroop task, on a total of 3,520 trials, we excluded 34 trials for the younger group and 67 trials for the older group as they were either below 1% or above 99% of the RT distribution of their age group. A mixed ANOVA analysis of the accuracy revealed only a significant main effect of Congruency, with less accurate performance in incongruent than congruent trials [$F(1,46)= 7.4$; $p = 0.01$]. On the response time, we found significant main effects of Group, with older adults being slower than younger adults [$F(1,46)= 38.76$; $p < 0.0001$], and Congruency, with longer response times for incongruent than congruent trials [$F(1,46)= 132.7$; $p < 0.0001$]. We also found a significant two-way interaction between Group and Congruency with older adults being slower than younger adults in incongruent trials [$F(1,46)= 4.12$; $p < 0.05$].

For the Flanker task, on a total of 3,520 trials, we excluded 33 trials for the younger group and 36 trials for the older group as they were either below 1% or above the 99% of the RT distribution. Regarding accuracy, we only found a significant main effect of Congruency, with less accurate performance in incongruent than congruent trials [$F(1,46)= 1628.05$; $p < 0.0001$]. On the response time, we found significant main effects of Group, with older adults being slower than younger adults [$F(1,46)= 26.6$; $p < 0.0001$], and Congruency, with incongruent trials responded to slower than congruent trials [$F(1,46) = 14.74$; $p < 0.0001$].

We did not find a significant interaction between Congruency and Group. These results are in line with Experiment 1.

INSERT FIGURE 4 AND TABLE 4 HERE

Visual search analysis

On response accuracy (Figure 4, upper panel), we found significant main effects of Group, with younger participants being more accurate than older participants, and of Prime, whereby incongruently primed trials were performed more accurately than neutral trials. We also observed a significant three-way interaction between Prime, Object and Group, whereby older adults were less accurate on detecting inconsistent objects when primed with neutral information, as opposed to incongruently primed (refer to Table 4 for the model output).

On time to the first fixation to the target object (Figure 4, middle panel), we found a main effect of Object, whereby inconsistent objects were looked at faster than consistent objects, especially when primed with incongruent, as opposed to neutral, information (a Prime and Object interaction). When looking at the total fixation time on the object (Figure 4, lower panel), we only observed a significant two-way interaction between Prime and Object, whereby inconsistent objects were looked at for longer than consistent objects when participants were primed with a neutral prime as compared to a congruent prime.

Discussion

In this follow-up experiment, we tested the impact of conceptual cueing of contextual information on the visual search. The experimental design also included primes carrying no semantic information (i.e., a scrambled word mask), defined as 'neutral' primes.

The most prominent result of this follow-up experiment is the disrupting effect of the neutral prime, which has led to a worse target detection performance compared to the incongruent priming condition, especially in older adults searching for an inconsistent target object. We speculate that a neutral cue increases the attentional load of the participants to retain the identity of the target object in memory and hence acts as a perceptual distractor (e.g., Lustig, Hasher, & Tonev, 2006). This increased load needs to be dealt with for a target to be remembered and correctly identified; and this distractor effect may be especially strong when the target is inconsistent with the scene context, and is already less readily associated with it.

This result may also be read in parallel with those of Experiment 1 with the congruent visual priming condition. A congruent scene prime created a high-level conceptual overlap with the search scene, but it also triggered a low-level perceptual (visual) interference due to differences in the spatial layout of the scenes. A neutral prime created a low-level perceptual interference as well, but in this case on the identity of the target cued. Thus, this result seems to indicate that the increased difficulty of older adults may not only relate to the inhibition of irrelevant (conceptually conveyed) contextual information, but also the inhibition of irrelevant (perceptual) non-contextual information. Regarding eye-movements, we replicated Experiment 1 by showing that inconsistent objects are looked at earlier than consistent objects, especially when the priming was incongruent (relative to neutral), and for longer when the priming is neutral (relative to congruent). Both findings seem to converge upon the idea that a neutral prime acts as a perceptual distractor and hence it disrupts eye-movement responses of target identification (time to first fixation) and processing (total fixation time).

Individual differences in cognitive control and their impact on visual search

The results reported so far show that cognitive ageing mediates the access and use of contextual information. As amply reviewed, most literature on ageing supports the idea that cognitive control is directly involved in top-down mechanisms (e.g., Gazzaley & D'Esposito, 2007), and that contextual information plays a prominent role on such mechanisms (Gilbert & Li, 2013). Often in this literature, cognitive control is discussed by looking only at group responses from a single task. Here, we want to go one step further, and explore how and to what extent participants' performance on two tasks known to tap into cognitive control (i.e., the Stroop and the Flanker) correlate with their search performance.

In order to do so, as a first step, we pool together the data from Experiment 1 and Experiment 2[1] to have greater statistical power, as well as to compare directly the influence of prime type (scene vs word) on search performance. We compute, for each participant, aggregate (mean) measures of the three DVs of the search task (i.e., response accuracy, and z-scored time to first fixation and total time of fixation) independently for each of the four

[1] We exclude the neutral priming condition from this analysis as the two datasets can only be compared on the same levels of congruency, i.e., Congruent and Incongruent prime. Additionally, in order to obtain more solid insights about the influence of prime type (scene vs word) on visual search performance, we must directly compare the two.

experimental conditions (i.e., crossing Object: Consistent vs Inconsistent; and Prime: Congruent vs Incongruent). Then, we aggregate the performance measures of response accuracy and response time (on correct trials only) of the same participants on the Stroop and Flanker executive control tasks for Congruent and Incongruent trials separately yielding a total of 8 measures for each participant. From these 8 measures, we want to obtain a single summary measure of cognitive control per each participant to correlate with his/her search performance. However, as Friedman & Miyake (2004) highlight, inhibition-related processes may be more than a single unitary construct, and different executive cognitive control tasks may tap into different aspects of response inhibition. Moreover, responses collected in any such tasks never reflect pure inhibition as they likely involve other cognitive processes (i.e., task impurity). Therefore, we use Exploratory Factor Analysis (EFA) to examine the loading of these 8 different measures into factors. Indeed, we find that these 8 measures loaded into different factors. For example, the response times to Congruent and Incongruent trials in the Stroop task load into a different factor than response times to Congruent and Incongruent trials in the Flanker task. Nevertheless, when we correlate the loading of these separate factors with the different search measures, we obtain similar results (refer to Supplementary Material A and Table S1 for the results). In a way, this result aligns well with Friedman and Miyake (2004), which indeed finds that the Stroop and Flanker load into different factors, but they are nevertheless very strongly related[2]. In the current study, we are not interested in discriminating among the different types of inhibition, and have observed very similar results when factor loadings for the Stroop and Flanker tasks are independently used as predictors of search performance. For

[2] See Figure 2 (page 112 of Friedman and Miyake, 2004), which shows a .68 relation between response inhibition and distractor inhibition.

this reason, we use Principal Component Analysis to compress the 8 measures of Stroop and Flanker data obtained for each participant into a single measure of cognitive control[3]. By doing so, we also partly address potential issues of task impurity, as our aggregated measure of cognitive control is estimated out of the participant performance in two different tasks (see Appendix B for more details on the PCA decomposition). We run the PCA independently for each age group. This approach assures that the first component of the PCA should not just reflect simple baseline differences between younger and older adults (e.g., overall slower responses in the older adults), as the groups were kept separate, and hence should genuinely index the cognitive control of each participant within his/her age group.

Statistical inference is then obtained using linear mixed-effects modelling with the same approach described above. The only difference now is the addition of prime Type (Word = -.5, Scene = .5) and the first component of the PCA[4], referred to as Cognitive Control (CC) as fixed effect predictors in the model, and that we only had Participants (101 levels) as the random effect, nested into their respective age Group.

----

[3]. Note, the results observed with Stroop and Flanker as separate EFA factors are equivalent to those obtained by using the first component of the PCA (compare Table 5 and S1).

[4] An increase on the first component of the PCA implies more difficult processing. Therefore, for interpretability, we reversed the sign of this predictor before entering it into the LME analysis, such that increase on Cognitive Control would imply easier, rather than more difficult, performance on the Stroop and Flanker tasks.

INSERT FIGURE 5 AND TABLE 5 HERE

Results and discussion

On response accuracy, we found significant main effects of Group, with younger participants being more accurate than older participants, and of Cognitive Control, whereby participants with better performances in the Stroop and Flanker tasks were also better at detecting the target object in the search task (refer to Table 5 for the model output). Older participants found it harder than younger participants to find an inconsistent target object when primed with congruent information (three-way interaction between Prime, Object and Group), mainly confirming the results presented in the *Visual search analysis* of Experiment 1 and Experiment 2. Additionally, we now find that older participants with better cognitive control were also better at detecting inconsistent targets when primed with a scene as opposed to with a word cue (four-way interaction, see Figure 5 for a visualisation of these results).

On time to the first fixation, we confirm the main effect of Object, whereby inconsistent objects were looked at faster than consistent objects. This effect was particularly strong when the participants received an incongruent prime, and it was a scene rather than a word (refer to Table 5 for the model output). As for the total fixation time on the object, we found that an inconsistent object was looked at for longer by older participants, and especially when the information primed was congruent and a visual scene (two, three and four-way interactions). We have not found any significant effect of cognitive control on eye-movement responses (the reader is referred to the Supplementary Material - Figure S4 for a visualisation of these results).

In this pooled analysis, we replicate most of the results observed when Experiment 1 and 2 were analysed separately, but also obtained some novel insights. We confirm that younger participants display a better performance than older participants on the search task, but also find that, regardless of the age group, better a cognitive control helps participants to be more accurate in the search. We also replicate that older adults had greater difficulty in suppressing congruent information when the search target was inconsistent with the context. In addition we find that older adults with better cognitive control are more accurate at detecting inconsistent objects when cued with scenes as opposed to words. This result indicates that older adults have greater difficulty in suppressing irrelevant information than younger adults but that cognitive control may be implicated in filtering it out to optimize search performance.

On the eye-movement measures, we confirm that older adults have greater difficulty in locating an inconsistent object when primed with congruent information, mainly if the primeis a scene. Once the inconsistent object is found, they also spend significantly more time than younger adults attending at it. However, we could not explain these findings on the eye-movement responses in terms of cognitive control, as measured using the Stroop and Flanker tasks. Thus, this result indicates that classic tasks indexing cognitive control may not reflect all types of inhibitory mechanisms and that additional top-down cognitive mechanisms are at work when overt attention is allocated during naturalistic scene understanding.

General Discussion

Research on healthy ageing has shown that, beyond a slowing of cognitive processing (e.g., Salthouse, 2012), older adults become over-reliant on contextual information (e.g., Gazzaley & D'Esposito, 2007; Lindenberger & Mayr, 2014) and experience difficulties in suppressing it when irrelevant to the task at hand (e.g., Lustig & Jantz, 2015; Zanto & Gazzaley, 2014). Visual search tasks require an active inhibition of distractor objects to locate the target efficiently. The inhibition of irrelevant information is a costly cognitive operation, which for older adults may be more challenging due to their reduced cognitive control (e.g., Lustig et al., 2006, Schwarzkopp et al., 2016).

Most ageing research on visual search has focused on object arrays (e.g., Potter et al., 2012) rather than more naturalistic scenes (but see Neider & Kramer, 2011). Moreover, little attention was paid to the relationship between top-down contextual information, which is key to scene understanding (e.g., Torralba, Oliva, Castelhano, & Henderson, 2006) and cognitive control, which is fundamental to selective attention (e.g., Lavie, Hirst, De Fockert, & Viding, 2004), especially in the context of cognitive ageing.

In order to better examine this relationship, we elicited the pre-activation of top-down contextual information before a search task to bias the participants towards expectations (conceptual and visual-perceptual) that may support (or hinder) the search for the target. When these expectations can hinder visual search, they have to be suppressed to maximise search efficiency. As said, the inhibition of irrelevant information is a cognitive control operation that is assumed to be impaired in older adults, and hence we predicted they would experience more difficulties in detecting a target violating contextual expectations that are irrelevant, or misleading, to the search. Moreover, we explicitly measured cognitive control using classic tasks, such as Stroop and Flanker, to assess its genuine role in the visual search.

We tested these hypotheses on two different groups of younger and older adults who performed a visual search task after being primed with contextual information (a scene or a word) congruent, incongruent or neutral (only in Experiment 2) with the context of the search scene. The semantic relationship between the target object and the search scene was also manipulated so that the former was either consistent or inconsistent with the latter. Additionally, we correlated the search performance of our participants with their performance on the Stroop and Flanker tasks, together represented using PCA, while comparing the effect of prime types (i.e., scenes vs words).

When the prime was a scene (Experiment 1), older adults were slower and less accurate than younger adults at identifying objects that were inconsistent with the search scene (e.g., a clothes iron in a restaurant). Critically, this detrimental effect on target detection was more prominent when the context of the scene was pre-activated (via, for example, exposure to a prime restaurant scene before searching in a different restaurant scene). This pre-activated congruent context reinforced expectations for the presence of consistent, rather than inconsistent, objects (conceptual overlap), and boosted expectations for a specific spatial layout of the objects building the scene. In other words, by visually priming a congruent context, we increased the expectation for consistent objects, thereby making an inconsistent object more likely to be considered a distractor and hence ignored, while the spatial layout of the objects between the prime and the search scene had also to be re-assessed. Arguably, this process of attentional re-orienting due to conflicting information is detrimental to perceptual identification (e.g., Vallesi et al., 2009). Younger adults were better able to suppress the violation of such expectations and maintain an efficient search in this condition. Older adults instead were lesser able to do so, due to their reduced cognitive control ability.

When participants were primed with simple conceptual information via words rather than picture cues, the older group still performed worse when searching for inconsistent objects, now after being neutrally, relative to incongruently primed. It is possible that both the congruent condition for the visual primes and the neutral condition for word primes tax participants' cognitive control mechanisms in similar ways: they both involve the suppression of distracting perceptual information. In Experiment 1, the distracting information was the different spatial layout for contextually congruent scenes (i.e., they mostly share the same objects but these are placed differently). In Experiment 2, the distracting information was instead the scrambled word prime that increased the attentional load to retain information about the target identity and hence acted as a perceptual distractor. More research on the effect of perceptual distractors during scene viewing in contextual cueing paradigms is needed, however. In particular, the spatial dependency of objects in the scene is a crucial factor when attention is allocated during scene understanding (e.g., Brockmole & Henderson, 2006; Wu, Wang, & Pomplun, 2014). Therefore, it would be relevant for future work to establish how a more systematic manipulation of the spatial configuration of objects between the prime and the search scene (and not just their semantic association) would modulate object recognition and identification.

In both experiments, we observed that greater cognitive control was linked to higher accuracies in trials with inconsistent objects, both for visual and word primes, and when the primes were congruent or neutral, respectively. Proactive control has been proposed as a mechanism that acts upon prior expectations. We posit that the difficulty experienced by the older participants in inhibiting the irrelevant pre-activated contextual information may be due to compromised proactive control. In future studies, it would be essential to examine whether changes in the proportion of congruent and incongruent contextual priming (e.g., 75% vs 25%, and vice-versa) may provide a different impact on target detection. We could expect that greater congruent priming may lead to a further decrease in the detection of inconsistent targets.

When looking at eye-movement responses, we found that inconsistent objects were fixated earlier, especially when using a word prime that was incongruent, rather than neutral. Older participants looked at the target objects for longer when congruently primed with a scene, and when the target of the search was inconsistent with it. The violation of object-scene information is known to be associated with more complex behavioural and neural processing (e.g., Coco, Araujo, & Petersson, 2017; Coco, Malcolm, & Keller, 2014; Mudrik, Lamy, & Deouell, 2010). We add to this literature that beside having a more effortful attentional processing to detect a target (e.g., Williams et al., 2009), older participants display additional efforts with target objects that violate the context of the scene. Future research should better examine the mechanisms of object-scene integration in cognitive ageing, perhaps by applying contextual cueing to other tasks where short-term memory is involved (e.g., change-detection). In such a task, we may expect that the recognition of a changed inconsistent object may be more difficult for older participants if the prime and search scene share the same context.

Although our data suggest that, at least for accuracy in the search task, a higher cognitive control improves visual search performance, we cannot explain the whole pattern of results only through this factor. Firstly, cognitive control capacity modulated the detection of the target also for the younger group, suggesting that this capacity plays a role regardless of age differences. Secondly, although we found group differences in the attentional patterns between older and younger adults, we did not find any effect of cognitive control on any of the eye-movement measures analysed. One possibility is that Stroop and Flanker tasks better index response inhibition rather than distractor inhibition. The need to inhibit a detection response holds throughout the visual search process, and perhaps, that's the reason why we observed effects of cognitive control only on response accuracy, rather than on the eye-movement measures, which are more sensitive to distractor inhibition. The other possibility is that the inhibitory mechanisms underlying these two tasks do not directly relate to visual search, or perhaps, they are not fine-grained enough to capture age-related differences on the allocation of overt attention. Our study, therefore, highlights the importance of collecting other measures of cognitive control to examine better its multi-faceted role in tasks involving top-down processing, such as visual search.

In sum, we have shown that the pre-activation of contextual information mediates the detection of a target object during visual search in naturalistic scenes. This pre-activation yielded different outcomes for older and younger adults, and strongly varied with the semantic fit of such a target with the context, and the cueing administered. We have also shown preliminary links between cognitive control and search efficiency, although such a finding was restricted to the detection accuracy rather than to the overt allocation of attention of our younger and older groups.

References

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629. https://doi.org/10.1038/nrn1476

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bédard, M., Leonard, E., McAuliffe, J., Weaver, B., Gibbons, C., & Dubois, S. (2006). Visual attention and older drivers: The contribution of inhibition of return to safe driving. *Experimental Ageing Research*, *32*(2), 119–135. https://doi.org/10.1080/03610730500511918

Belke, E., Humphreys, G. W., Watson, Derrick, G., Meyer, A. S., & Telling, A. L. (2008). Top-down effects of semantic knowledge in visual search are modulated by cognitive but not perceptual load. *Perception & Psychophysics*, *70*(8), 1444–1458. https://doi.org/10.3758/PP.70.8.1444

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In *Variation in working memory* (pp. 76–106).

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*(4), 699–706. https://doi.org/10.1037/0278-7393.32.4.699

Brockmole, J. R., & Henderson, J. M. (2006). Using real-world scenes as contextual cues for search. *Visual Cognition*, *13*(1), 99–108. https://doi.org/10.1080/13506280500165188

Bugg, J. M. (2014). Evidence for the sparing of reactive cognitive control with age. *Psychology and Ageing*, *29*(1), 115–127. https://doi.org/10.1037/a0035270

Burton-Danner, K., Owsley, C., & Jackson, G. R. (2001). Ageing and feature search: The effect of search area. *Experimental Ageing Research*, *27*(1), 1–18. https://doi.org/10.1080/03610730125782

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/S1364-6613(00)01476-5

Chun, M. M. & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28.

Coco, M. I., Araujo, S., & Petersson, K. M. (2017). Disentangling stimulus plausibility and contextual congruency: Electro-physiological evidence for differential cognitive dynamics. *Neuropsychologia*, *96*(November 2016), 150–163. https://doi.org/10.1016/j.neuropsychologia.2016.12.008

Coco, M. I., Malcolm, G. L., & Keller, F. (2014). The interplay of bottom-up and top-down mechanisms in visual guidance during object naming. *Quarterly Journal of Experimental Psychology*, *67*(6), 1096–1120. https://doi.org/10.1080/17470218.2013.844843

Coco, M. I., Nuthmann, A., & Della Sala, S. (2017). The role of visual attention and high-level object information on short-term visual working memory in a change detection task. *Journal of Vision*, *17*(10), 61–61.

Cornelissen, T. H. W., & Võ, M. L.-H. (2017). Stuck on semantics: Processing of irrelevant object-scene inconsistencies modulates ongoing gaze behavior. *Attention, Perception, & Psychophysics*, *79*(1), 154–168. https://doi.org/10.3758/s13414-016-1203-7

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*(11), 973–980. https://doi.org/10.1111/j.1467-9280.2006.01815.x

Enns, J. T., & Lollo, V. Di. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–352. https://doi.org/10.1016/S1364-6613(00)01520-5

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information processing rate and amount: Implications for group difference in response latency. *Psychological Bulletin*, *125*(6), 777–799.

Foster, J. K., Behrmann, M., & Stuss, D. T. (1995). Ageing and Visual Search: Generalized Cognitive Slowing or Selective Deficit in Attention? *Ageing, Neuropsychology, and Cognition*, *2*(4), 279–299. https://doi.org/10.1080/13825589508256604

Friedman, N. P., & Miyake, A. (2004). The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal ageing. *Nature Neuroscience*, *8*(10), 1298–1300. https://doi.org/10.1038/nn1543

Gazzaley, A., & D'Esposito, M. (2007). Top-down modulation and normal ageing. *Annals of the New York Academy of Sciences*, *1097*(1), 67–83. https://doi.org/10.1196/annals.1379.010

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

Gonthier, C., Braver, T. S., & Bugg, J. M. (2016). Dissociating proactive and reactive control in the Stroop task. *Memory & Cognition*, *44*(5), 778–788. https://doi.org/10.3758/s13421-016-0591-1

Goujon, A., Brockmole, J. R., & Ehinger, K. A. (2012). How visual and semantic information influence learning in familiar contexts. *Journal of Experimental Psychology. Human Perception and Performance*, *38*(5), 1315–1327. https://doi.org/10.1037/a0028126

Hahn, S., Carlson, C., Singer, S., & Gronlund, S. D. (2006). Ageing and visual search: Automatic and controlled attentional bias to threat faces. *Acta Psychologica*, *123*(3), 312–336. https://doi.org/10.1016/j.actpsy.2006.01.008

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Henderson, J. M., Weeks, Phillip a., J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 210–228. https://doi.org/10.1037//0096-1523.25.1.210

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(1), 31–47. https://doi.org/10.1037/0096-1523.33.1.31

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*, *17*(1–2), 273–291. https://doi.org/10.1080/13506280802193367

Howard, J. H. J., Dennis, N. A., Howard, D. V., Yankovich, H., & Vaidya, C. J. (2004). Implicit Spatial Contextual Learning in Healthy Ageing. *Neuropsychology*, *18*(1), 124–134. https://doi.org/10.1038/nsmb.2010.The

Kuznetsova, A., Christensen, R. H. B., Bavay, C., & Brockhoff, P. B. (2014). Automated mixed ANOVA modeling of sensory and consumer data. *Food Quality and Preference*, *40*(PA), 31–38. https://doi.org/10.1016/j.foodqual.2014.08.004

LaPointe, M. R. P., & Milliken, B. (2016). Semantically incongruent objects attract eye gaze when viewing scenes for change. *Visual Cognition*, *24*(1), 63–77. https://doi.org/10.1080/13506285.2016.1185070

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lindenberger, U., & Mayr, U. (2014). Cognitive ageing: Is there a dark side to environmental support? *Trends in Cognitive Sciences*, *18*(1), 7–15. https://doi.org/10.1016/j.tics.2013.10.006

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology. Human Perception and Performance*, *4*(4), 565–572. https://doi.org/10.1037/0096-1523.4.4.565

Lustig, C., Hasher, L., & Tonev, S. T. (2006). Distraction as a determinant of processing speed. *Psychonomic Bulletin & Review*, *13*(4), 619–625.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view." *Inhibition in Cognition*, (571), 145–162. https://doi.org/http://dx.doi.org/10.1037/11587-008

Lustig, C., & Jantz, T. (2015). Questions of age differences in interference control: When and how, not if? *Brain Research*. https://doi.org/10.1016/j.brainres.2014.10.024

Madden, D. J. (2007). Ageing and visual attention. *Current Directions in Psychological Science*, *16*(2), 70–74. https://doi.org/10.1111/j.1467-8721.2007.00478.x

Madden, D. J., Whiting, W. L., Cabeza, R., & Huettel, S. A. (2004). Age-Related Preservation of Top-Down Attentional Guidance During Visual Search. *Psychology and Ageing*, *19*(2), 304–309. https://doi.org/10.1037/0882-7974.19.2.304

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mayr, U. (2001). Age Differences in the Selection of Mental Sets: The Role of Inhibition, Stimulus Ambiguity, and Response-Set Overlap. *Psychology and Ageing*, *16*(1), 96–109.

Monge, Z. A., & Madden, D. J. (2016). Linking cognitive and visual perceptual decline in healthy ageing: The information degradation hypothesis. *Neuroscience and Biobehavioral Reviews*, *69*, 166–173. https://doi.org/10.1016/j.neubiorev.2016.07.031

Mudrik, L., Lamy, D., & Deouell, L. Y. (2010). ERP evidence for context congruity effects during simultaneous object-scene processing. *Neuropsychologia*, *48*(2), 507–517. https://doi.org/10.1016/j.neuropsychologia.2009.10.011

Muth, C., Bales, K. L., Hinde, K., Maninger, N., Mendoza, S. P., & Ferrer, E. (2016). Alternative Models for Small Samples in Psychological Research: Applying Linear Mixed Effects Models and Generalized Estimating Equations to Repeated Measures Data. *Educational and Psychological Measurement*, *76*(1), 64–87. https://doi.org/10.1177/0013164415580432

Neider, M. B., & Kramer, A. F. (2011). Older adults capitalize on contextual information to guide search. *Experimental Ageing Research*, *37*(5), 539–571. https://doi.org/10.1080/0361073X.2011.619864

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, *46*(5), 614–621. https://doi.org/10.1016/j.visres.2005.08.025

Potter, L. M., Grealy, M. A., Elliott, M. A., & Andres, P. (2012). Ageing and performance on an everyday-based visual search task. *Acta Psychologica*, *140*(3), 208–217. https://doi.org/10.1016/j.actpsy.2012.05.001

Rey-Mermet, A., & Gade, M. (2017). Inhibition in ageing: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin and Review*, 1–22. https://doi.org/10.3758/s13423-017-1384-7

Rodrigues, P. F. S., & Pandeirada, J. N. S. (2015). Attention and working memory in elderly: the influence of a distracting environment. *Cognitive Processing*, *16*(1), 97–109. https://doi.org/10.1007/s10339-014-0628-y

Salthouse, T. (1994). Ageing Associations: Influence of Speed on Adult Age Differences in Associative Learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(6), 1486–1503. https://doi.org/10.1037/0278-7393.20.6.1486

Salthouse, T. (2012). Consequences of Age-Related Cognitive Declines. *Annual Review of Psychology*, *63*(1), 201–226. https://doi.org/10.1146/annurev-psych-120710-100328

Schwarzkopp, T., Mayr, U., & Jost, K. (2016). Early Selection Versus Late Correction: Age-Related Differences in Controlling Working Memory Contents. *Psychology and Ageing*, *31*(5), 430–441. https://doi.org/http://dx.doi.org/10.1037/pag0000103

Smyth, A. C., & Shanks, D. R. (2011). Ageing and implicit learning: explorations in contextual cuing. *Psychology and Ageing*, *26*(1), 127–132. https://doi.org/10.1037/a0022014

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop Performance in Healthy Younger and Older Adults and in Individuals with Dementia of the Alzheimer's Type. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(2), 461–479. https://doi.org/10.1037/0096-1523.22.2.461

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual Guidance of Eye Movements and Attention in Real-World Scenes: The Role of Global Features in Object Search. *Psychological Review*, *113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Underwood, G., Humphreys, L., & Cross, E. (2007). Congruency, saliency and gist in the inspection of objects in natural scenes. In *Eye Movements: A window on mind and brain* (pp. 563–579). https://doi.org/10.1016/B978-008044980-7/50028-8

Vallesi, A., Stuss, D. T., McIntosh, A. R., & Picton, T. W. (2009). Age-related differences in processing irrelevant information: Evidence from event-related potentials. *Neuropsychologia*, *47*(2), 577–586. https://doi.org/10.1016/j.neuropsychologia.2008.10.018

Võ, M. L.-H., & Henderson, J. M. (2011). Object-scene inconsistencies do not capture gaze: evidence from the flash-preview moving-window paradigm. *Attention, Perception & Psychophysics*, *73*(6), 1742–1753. https://doi.org/10.3758/s13414-011-0150-6

Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2005). Search, enumeration, and ageing: Eye movement requirements cause age-equivalent performance in enumeration but not in search tasks. *Psychology and Ageing*, *20*(2), 226–240. https://doi.org/10.1037/0882-7974.20.2.226

Whiting, W. L., Madden, D. J., Pierce, T. W., & Allen, P. A. (2005). Searching from the top down: Ageing and attentional guidance during singleton detection. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(1), 72–97. https://doi.org/10.1080/02724980443000205

Williams, C. C., Zacks, R. T., & Henderson, J. M. (2009). Age differences in what is viewed and remembered in complex conjunction search. *Quarterly Journal of Experimental Psychology*, *62*(5), 946–966. https://doi.org/10.1080/17470210802321976

Wu, C. C., Wang, H. C., & Pomplun, M. (2014). The roles of scene gist and spatial dependency among objects in the semantic guidance of attention in real-world scenes. *Vision Research*, *105*, 10–20. https://doi.org/10.1016/j.visres.2014.08.019

Wu, C. C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2014.00054

Zanto, T. P., & Gazzaley, A. (2014). Attention and ageing. In *The Oxford Handbook of Attention* (pp. 927–971).

Appendix A: Norming the experimental material.

Experiment 1: The experimental material was normed in three phases. First, 10 undergraduate students (6 female, age = 21.5±1.43) were asked to name 8 objects they would expect in each of the 7 indoor contexts chosen (i.e., Bathroom, Church, Kitchen, Bedroom, Office, Restaurant, Waiting Room). 198 unique objects were produced overall and the 8 most mentioned ones, for each context, were selected as consistent targets. The 8 inconsistent objects, instead, were chosen from a pool of objects never mentioned by any of the interviewees. Second, a new group of 20 students (12 female, age = 21.2±1.23) was asked to rate the consistency of the objects within their respective contexts on a 6-point Likert scale (from *very unlikely* to *very likely*). We retained the 6 objects with the highest (consistent) and lowest (inconsistent) scores (e.g., Soap – Bathroom: mean = 5.9±0.31; Flashlight – Bathroom: mean = 1.3±0.67). Finally, a new group of 12 students (6 female, age = 22.5±1.87) was asked to rate on a 6-point Likert scale: (a) the contextual congruency between the prime scene and the search scene, (b) the likelihood that each scene would be a representative example of its corresponding context, (c) the consistency of the target object with respect to the search scene and (d) the association between each target object and its referring word cue. We show that: (a) the difference between congruous and incongruous prime scenes is statistically significant ($M = 5.4$ vs. 1.5, $t(489) = 44$, $p < 0.001$); (b) both prime scenes and the search scenes are considered as typical examples of their context ($M = 5$ for both, $t(503) = 0.167$, $p = 0.87$); (c) consistent objects are rated as more likely to be in the search scene than inconsistent objects ($M = 5.1$ vs. 1.7, $t(470) = 28$, $p < 0.001$); and (d) the target objects are rated as highly identifiable by the word cue (mean = 5.3±1.7).

Experiment 2: The additional 20 experimental scenes added in Experiment 2 to accommodate the effect size of the design for the additional Neutral priming condition have been taken from a larger dataset of a recent study by Coco, Nuthmann and Della Sala, 2017. These photographs were created following the same logic. Each scene contained a critical object that was either consistent or inconsistent with the context of the scene and the position of the object was counterbalanced left and right within the scene. Eight participants were asked to assess the consistency of the object and its linguistic denotation. Each participant saw the experimental scene and a box with a crop of the critical object along its side. He/she was asked to (a) write down the name for the critical object shown, and (b) to respond to the question “How likely is it that this object would be found in this room?” using a six-point Likert scale (1-6). On the subset of 20 scenes used in this experiment, we obtained a mean naming agreement of 96.8%. Furthermore, consistent objects were judged as significantly more likely (5.78 ± 0.57) to appear in the scene than inconsistent objects (1.88 ± 1.11), as confirmed with an independent-samples Kruskal-Wallis H test ($\chi2(1) = 129.2$, $p < .001$).

Appendix B: Correlation and Principal Component Analysis for the cognitive control tasks.

Conceptually, Principal Component Analysis (PCA) is used to reduce the dimensionality of variables that are likely correlated, and hence can be "compressed" into a smaller dimensionality, i.e., fewer variables. In particular, PCA finds the directions where most of the variance of the data spread, with each component being one of these new dimensions (i.e., the axes). Figure 1B shows the correlation of the responses for the Flanker and the Stroop tasks, for the older and younger group separately (left side), as well as, the two components of the PCA computed on those (right side), for the two different experiments. We find that the first component already takes both in Experiment 1 and in Experiment 2 a substantial amount of the variance. In particular, the first component of the PCA explains 89% of the variance in the younger group and 70% in the older group for the participants of Experiment 1, and 86% of the variance for the younger group and 87% for the older group in Experiment 2. These results suggest that the first component of the PCA can be confidently used as a predictor for the cognitive control of each participant. Moreover, by running the PCA separately for the younger and the older group, we make sure that its first component is not just accounting for the baseline difference in reaction time and performance of the two groups, but genuinely capture individual differences in the cognitive control within his/her reference age group.

INSERT FIGURE 1B HERE
