



OPEN ACCESS

EDITED BY

Stefan Pollmann,
Otto von Guericke University Magdeburg,
Germany

REVIEWED BY

Marian Sauter,
University of Ulm, Germany
Artyom Zinchenko,
Ludwig Maximilian University of Munich,
Germany

*CORRESPONDENCE

Matias J. Ison

✉ Matias.Ison@nottingham.ac.uk

[†]These authors have contributed equally to this work and share first authorship

[‡]These authors have contributed equally to this work and share senior authorship

RECEIVED 22 May 2024

ACCEPTED 06 August 2024

PUBLISHED 27 August 2024

CITATION

Barbosa A, Ruarte G, Ries AJ,
Kamienkowski JE and Ison MJ (2024)
Investigating the effects of context, visual
working memory, and inhibitory control in
hybrid visual search.
Front. Hum. Neurosci. 18:1436564.
doi: 10.3389/fnhum.2024.1436564

COPYRIGHT

© 2024 Barbosa, Ruarte, Ries, Kamienkowski
and Ison. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Investigating the effects of context, visual working memory, and inhibitory control in hybrid visual search

Alessandra Barbosa^{1†}, Gonzalo Ruarte^{2†}, Anthony J. Ries³,
Juan E. Kamienkowski^{2,4‡} and Matias J. Ison^{1‡*}

¹School of Psychology, University of Nottingham, Nottingham, United Kingdom, ²Laboratorio de Inteligencia Artificial Aplicada, Instituto de Ciencias de la Computación (Universidad de Buenos Aires – Consejo Nacional de Investigaciones Científicas y Técnicas), Buenos Aires, Argentina, ³DEVCOM Army Research Laboratory, Aberdeen Proving Ground, MD, United States, ⁴Departamento de Computación (Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires), Buenos Aires, Argentina

Introduction: In real-life scenarios, individuals frequently engage in tasks that involve searching for one of the distinct items stored in memory. This combined process of visual search and memory search is known as hybrid search. To date, most hybrid search studies have been restricted to average observers looking for previously well-memorized targets in blank backgrounds.

Methods: We investigated the effects of context and the role of memory in hybrid search by modifying the task's memorization phase to occur in all-new single trials. In addition, we aimed to assess how individual differences in visual working memory capacity and inhibitory control influence performance during hybrid search. In an online experiment, 110 participants searched for potential targets in images with and without context. A change detection and go/no-go task were also performed to measure working memory capacity and inhibitory control, respectively.

Results: We show that, in target present trials, the main hallmarks of hybrid search remain present, with a linear relationship between reaction time and visual set size and a logarithmic relationship between reaction time and memory set size. These behavioral results can be reproduced by using a simple drift-diffusion model. Finally, working memory capacity did not predict most search performance measures. Inhibitory control, when relationships were significant, could account for only a small portion of the variability in the data.

Discussion: This study provides insights into the effects of context and individual differences on search efficiency and termination.

KEYWORDS

visual search, memory search, real-world scenes, individual differences, visual working memory capacity, inhibitory control

1 Introduction

Visual search (VS) is the action of looking for a target among distractors. It is a ubiquitous task in many everyday situations, from searching for products in stores to driving. Mainly used to examine visual attention (Treisman and Gelade, 1980), visual search has been intensively researched for decades. The core manipulation in visual search is to vary the distractor set size while measuring the time until a target is detected (reaction times or RTs) and measuring

detection accuracy (Neisser, 1964). From these, a hallmark was established—the linear dependence between RT and visual set size (VSS) (Wolfe, 2020).

While much is known about elements that influence visual search when *one* sole item is searched for, search in real life is substantially more complex, often involving several objects in memory. From (Schneider and Shiffrin, 1977) classic work, encompassing both visual search and memory search (MS), hybrid search (HS) is when observers search for any of many possible targets (Wolfe, 2012a). Hybrid search tasks involve memorizing potential targets to subsequently identify their incidence in a display. In a previous study (Schneider and Shiffrin, 1977), different memory manipulations included whether the hybrid search had consistent mapping (when target sets are fixed, used throughout trials), all-new mapping (when new items compose the memory set each trial), or varied mapping (when targets become distractors and vice versa), which negatively impacted efficiency and accuracy in this order. Regardless of the specific conditions, a robust logarithmic relationship between RT and memory set size (MSS) has been consistently reported: up to MSSs of 100 for objects (Wolfe, 2012a) and words (Boettcher and Wolfe, 2015). Nonetheless, most recent hybrid search studies have not manipulated memory mappings. Solely utilizing consistent-mapping-like paradigms, high accuracies are observed (e.g., see Wolfe, 2012a). Under such conditions, speed-accuracy trade-offs that vary across set sizes cannot be evaluated as adequately as by paradigms that limit memory strengths (Nosofsky et al., 2014b). Given the prominent role of context in memory representations, such as recognizing a face or place after a single real-world encounter (Ison et al., 2015), examining paradigms with trial-by-trial changing memory sets is justified.

Most hybrid search findings are based on experiments conducted with artificial stimuli on blank backgrounds (Wolfe et al., 2011b), which may limit their ecological validity. In visual search, accumulating evidence shows that context is critical in guiding attention in real-world search (Wolfe et al., 2011a). Indeed, context guidance was found to overpower bottom-up saliency in guiding eye movements in naturalistic search (Henderson et al., 2009) and facilitate search in scenes for targets in plausible locations rather than implausible ones (Neider and Zelinsky, 2006) or blank backgrounds (Wolfe et al., 2011a). However, naturalistic scenes are complex/continuous, with set sizes impossible to define (Rosenholtz et al., 2007) but they are never random (Henderson and Ferreira, 2004). Scenes have syntax, that is, structural plausibility—humans appear on horizontal surfaces (Biederman, 1977; Torralba et al., 2006), and semantics, that is, meaningful associations—for example, a toothbrush on a sink (Vö and Henderson, 2009). In hybrid search, the limited exploration of contextual information has left mixed results. To our knowledge, only Boettcher et al. (2013, 2018) attempted to examine contextual effects in hybrid search. However, both were investigating whether context-aided memory set partition to context-relevant items at *fixed visual set size* and memory set size. This means that it is largely unknown how variable set sizes affect their relationship with context. The primary objective of this study was to explore the relationship between reaction time and set size in visual search and memory search.

A small number of hybrid search models have been proposed in the literature to explain behavioral signatures in hybrid search. Cunningham and Wolfe (2014) postulated a three-stage model for hybrid search. After the first stage of *guided* search (Wolfe, 2007), where solely feature-plausible objects affiliated to the memory set are

considered for the second stage of object recognition, a third stage of logarithmic memory search is performed through a diffusion process. Each memorized target item races as information is accumulated in parallel over time, forming N -diffusion processes, where $N = MSS$. Target selection only occurs when information from a target accumulator reaches the decision threshold. The threshold is set higher to impede exaggerated false alarms and lower to promote speed—the speed-accuracy trade-off (Chun and Wolfe, 1996). As most hybrid search paradigms follow a condition akin to consistent mapping, not much is known when targets' memory strengths are manipulated (Wolfe et al., 2015). Drawing from memory search literature, which focuses on the processes underlying memory search through memory-mapping manipulations, Nosofsky et al. (2014b) proposed an exemplar-familiarity random-walk memory search model with a logarithmic RT diffusion stage. In contrast to recent hybrid search models, it proposes that racing items are dictated by their 'memory strength', influenced by repetition effects and time presented, while they compete, and information is accumulated. Thus, while the curvilinear relationship of $RT \times MSS$ is maintained, efficiency and accuracy are affected (Nosofsky et al., 2014a,b), which can elucidate hybrid search processes for different memorization conditions. Another family of computational models that have been extensively used in decision-making tasks involving two-choice decision-making is the drift-diffusion model (DDM) (Ratcliff and McKoon, 2008). Essentially, the decision-maker accumulates evidence until a boundary is reached, and at that moment the decision is made. This process is modeled by sequential diffusion signals that drift over time. Diffusion models have been used to estimate the RT distribution in visual search tasks before (Palmer, 1975), and they may provide valuable insights into the underlying cognitive processes (Myers et al., 2022). More complex models have emerged in recent years that try to predict human scanpaths in a visual search task and one could also sample RTs from them (Kummerer et al., 2018; Bujia et al., 2022; Travi et al., 2022). The problem with this approach is that the more complex the model is, the more difficult it is to understand its underlying processes and relate them to a particular study. A secondary objective in this study was to implement a drift-diffusion model and interpret its parameter values (drift, boundary, non-decision time) in terms of the hybrid search variables, such as the visual set size (VSS), the memory set size (MSS), and the presence of the context.

Two fundamental components of executive function, namely, working memory and inhibitory control, have long been largely implicated in visual search processes. Visual working memory (VWM), which entails the maintenance and manipulation of a limited amount of visual information that serves current task demands (Luck and Vogel, 1997), has been connected to a variety of visual search processes, such as representing the search template to guide attention (Desimone and Duncan, 1995), comparing the search template to potential suitor objects (Bundesen, 1990), and influencing search facilitation in visual search under contextual cueing (Manginelli et al., 2013). Most recent hybrid search research has emulated a consistent mapping paradigm, which has been seen to produce near-error-free data (Nosofsky et al., 2014b). Studying different memory manipulations could be critical to further uncover memory's role in search, since visual working memory involvement might only be observed when targets change per trial (Woodman et al., 2007), and in memory search it is still unclear if context can restrict the memory set to scene-relevant items on a trial-by-trial basis (Boettcher et al., 2013, 2018). Inhibitory control (IC) is the critical executive function of suppressing

goal-irrelevant stimuli interference and subduing prepotent motor response (Young et al., 2018). Experimental paradigms conventionally include brusque prepotent response incitation, where one either proceeds or subdues action (e.g., go-no-go tasks) (Miyake et al., 2000). Individuals who have higher false alarms, or larger negative response bias, when subjected to signal detection theory (SDT) analysis, have more difficulty inhibiting prepotent responses (Young et al., 2018). Given the proposed importance of inhibitory control in visual search models for top-down selection and search termination (Treisman and Sato, 1990; Moran et al., 2013), an evaluation of the potential impact of individuals' inhibitory control is also merited. The last objective of this study was to examine the role of individual differences in working memory and inhibitory control in hybrid search. Based on previous literature, it is feasible to formulate specific predictions regarding the manipulations employed in this study. With mounting replications of hybrid search' RT signatures in divergent conditions, we expect all-new memorization and context to preserve *qualitatively* the main behavioral hybrid search signatures. However, we expect to see changes in search efficiency and accuracy. Predicting the role of visual working memory in hybrid search is more difficult, given the conflicting findings reported in visual search. If this study follows (Drew et al., 2016) account that a fixed amount of visual working memory is used as a conduit to transfer incoming target templates to memory, one might expect working memory capacity to correlate with RT intercepts in hybrid search, where individuals with higher working memory capacity might transfer targets through working memory faster than lower visual working memory capacity individuals. However, given our modified paradigm with targets changing per trial, analogous to Woodman et al. (2007), this proposition might not hold, and we might see higher visual working memory capacity producing smaller RT slopes as well; given high-visual working memory capacity individuals would have higher storage capacity/resource allocation flexibility as set sizes increase (Luck and Vogel, 2013). Given the potential importance of inhibitory control in search termination (Moran et al., 2013), we would expect that individuals with higher inhibitory control show larger RT intercepts than lower inhibitory control individuals, reflecting a potential to better maintain conservative thresholds with fewer false alarms and higher accuracy. In addition, by implementing a drift-diffusion model for hybrid search, we expect to link model parameters with the behavioral results and, in doing so, propose new directions for future experiments.

The rest of the study is organized as follows. Section 2 presents the results of the main experiment (Experiment 1) on the role of context in hybrid search. Section 3 shows that a simple drift-diffusion computational model can reproduce the main behavioral signatures of Experiment 1. Section 4 evaluates the extent to which individual differences in visual working memory and inhibitory control contribute to hybrid search behavior.

2 Experiment 1: context effects on hybrid search

2.1 Methods

2.1.1 Participants

An online data collection method gathered data from 110 participants, identifying as women (59), men (49), non-binary (1),

different identity (0), and non-disclosing (1). Participants were recruited via email and social media. Their ages (excluding 3 age misreports) ranged from 18 to 61 years old ($M = 27$ and $SD = 7.63$). Data were also collected for another 10 participants who were excluded due to aborting the experiment ($N = 7$) or low behavioral performance ($N = 3$). More specifically, to exclude participants with poor performance we used a threshold of 3 standard deviations below the mean accuracy (calculated from the raw data in the whole experiment), as done in previous visual search studies (Treisman, 1982). Participants were given the option to enter a draw for Amazon vouchers worth £20 each, so two people were compensated. Convenience sampling was adopted, as email and social media outreach is limited to users and not randomized. Online recruitment concentrated on Chinese-oriented social media platforms (WeChat) and Brazilian social media groups. The study was approved by the University of Nottingham School of Psychology Ethics Panel (ethics approval: S1240).

2.1.2 Materials

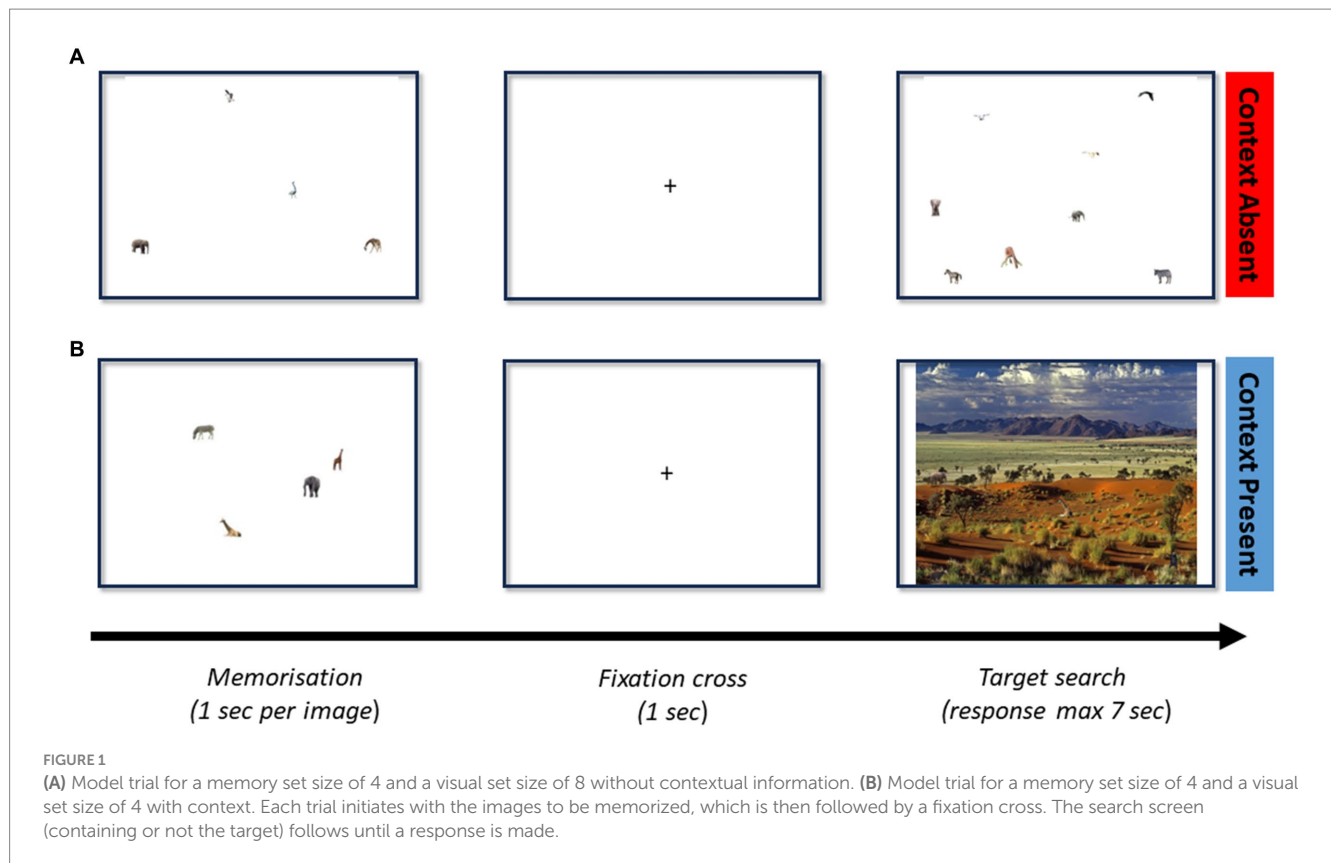
A total of 112 images with ($N = 56$) and without ($N = 56$) context were constructed. As shown in Figure 1B, the contextual images were constructed by superimposing real-world background scenes (including outdoor scenes—e.g., forest, and indoor scenes—e.g., shelf) with target and distractor stimuli. Each image contained 1, 2, 4, or 8 stimuli (visual set size, VSS), and a separate set of 1, 2, 4, or 8 stimuli were selected for memorization (memory set size, MSS). Items (targets and distractors) only appeared once during the experiment. All the targets and distractor stimuli were built from images from the COCO dataset (Lin et al., 2014) and ImageNet (Deng et al., 2009). These datasets were selected due to their extensive size and appropriate licensing. Images were then resized so that their new size would be compatible with the background image and placed according to scene syntax (i.e., no major violations of support, interposition, position, and size). All targets and distractors belonged to the same category (objects, animals, or people). The experiments were implemented in PsychoPy and executed online by participants through <https://pavlovio.org/> (Peirce et al., 2019).

2.1.3 Procedure

Participants searched 112 images with or without context and varying MSS and VSS. Following written consent, participants were instructed to press “m” for target present and “v” for target absent. The trial begins by showing the memory set (see Figure 1), on a display. A fixation cross is shown for 1 s in the center of the screen. This is then followed by the search screen (see Figure 1). The visual set contained at most 1 memory set item. Subsequent trials are promptly initiated after a response, or, after 7 s if no response is given. Target and context were present in 50% of trials. Completion of the 112 trials automatically triggered the de-briefing slides. Trials were randomized and showed one of the 4 combinations of conditions—*either target present or target absent conditions in a context absent or context present trial*. The completion time was approximately 20 min.

2.1.4 Design and data analysis

First, a descriptive analysis was conducted to assess the data's statistical metrics to provide validation of the experiments' properties and distributions delivered in an online medium. Second, a preliminary analysis was conducted to satisfy the data assumptions



necessary for the main analyses. Pre-analysis checks examined outliers, homoscedasticity, normality of residual errors, and sphericity by inspecting Cook's distance, the distribution of the residual error, and Mauchly's test. Where Mauchly's test indicated that the sphericity assumption had been violated (here, for variables with 2+ levels), Greenhouse–Geisser correction was applied. When significant and variance–covariance matrices were not homogenous, conservative Pillai's trace was reported. No outliers (maximum Cook's distance surpassing 1) were identified. For simple linear regressions, collinearity was assessed. No correlations above $r > 0.9$ were observed. By visually inspecting scatterplots of residuals against independent variables, no systemic pattern was observed (homoscedastic). In addition, examining residuals' histograms, all conformed to an acceptable departure from normality.

To identify potential large differences being associated with one of the images used in a given condition, the following process was followed: First, a one-way ANOVA over images was conducted per condition {context present, MSS, VSS, target present} with formula $rt. \sim stimulus$. For example, for {context present = 1, MSS = 1, VSS = 4, target present = 1}, there were four stimuli answered by all participants, and the ANOVA is aimed to distinguish if there were one or some of the stimuli behaving differently than the others. Then, if the F value was larger than 20 with a corresponding p -value smaller than 0.05, *post-hoc* comparisons using Tukey's HSD test ($\alpha = 0.05$) were used to determine which of the stimuli would be discarded. For the stimulus with the highest mean difference compared to the others, only one stimulus was discarded per condition at most to keep data balancing. This resulted in 6 out of 112 images being discarded. For counterbalancing, the trial order was randomized across participants

in each experimental task. Each participant experienced a unique sequence. The data were retrieved from <https://pavlovio.org/> to Excel files and matched by participant ID.

To illustrate the relationships between reaction times and set size, linear and logarithmic regression models were constructed in Python with NumPy (version 1.23.5) for visual search and memory search. These models were used to fit each participant and condition ($\{MS, VS\} \times \{CP, CA\} \times \{\text{linear}, \text{log}\}$).

Linear mixed models (LMMs) were used on the combined data from all participants to compare models for both visual search (VS) and memory search (MS). For each search type, four models were constructed: two with a linear term for set size and two with a logarithmic term. Two of these models included context and set size (VSS for visual search, MSS for memory search) as fixed effects, their interaction, and participants as a random intercept [e.g., $RT \sim VSS * Context + (1|Participant)$]. We also specified a baseline model of decreased complexity that only included context as a fixed effect [$RT \sim Context + (1|Participant)$]. To assess the sequential decomposition of the contributions of fixed effect terms, we used likelihood ratio tests, which allowed us to compare models of different complexity (Burnham and Anderson, 2004; Bates et al., 2015). For model selection, we used the Akaike information criterion (AIC). This approach allowed for a direct comparison of linear and logarithmic models while accounting for the repeated measures design of the study (Bates et al., 2015).

Python (version 3.10.5) was used for general data management and analysis. SciPy (version 1.12.0) was used to compute Pearson correlation coefficient. NumPy's polyfit was employed for linear regressions. Statsmodels (version 0.14.1) was used for the ANOVA

table and *post-hoc* analyses. Although statsmodels includes mixed models, we opted for lme4 (version 1.1.35.1) in R (version 4.3.2) for linear mixed model analysis as it is widely used and has strong community support. To compare models of increasing complexity, we used the anova() function from the lme4 package (version 1.1.35.1). Scikit-learn (version 1.3.0) was used for its ParameterGrid implementation, which facilitated dividing data between conditions.

2.2 Results

2.2.1 Online validation of experiments

Descriptive statistics of the hybrid search show properties consistent with previous lab-based experiments in literature, conferring adequate online validation. Consistent with Wolfe (2012a), this task saw a mean accuracy well above the chance level ($M=67\%$, $SEM=1\%$). The characteristic decline in accuracy when set sizes increase in divergent memory-mapping conditions was also observed: visual set size: $VSS1=75\%$ to $VSS8=55\%$; and memory set size: $MSS1=97\%$ to $MSS8=58\%$.

2.2.2 Context effects and trial-by-trial memorization in hybrid search

In this section, the following abbreviations will be used for orderliness—target present (TP); target absent (TA); context present (CP); and context absent (CA). Across all conditions, targets were correctly detected on 67% ($SEM:1\%$) of target present trials, and false alarms were produced on 26% ($SEM:0.9\%$) target absent trials. The corresponding discriminability d' -prime was 1.07.

For RT analyses, from a total of 12,320 trials across 110 participants, we excluded trials with no answer within the maximum time allowed (7 s) ($N=217$). Very short responses (less than 200 ms)

were only recorded for four trials and were not excluded from the data. Finally, we kept only correct trials which gave us a final sample of 8,042 trials (3,850 target present trials).

Linear mixed models with random intercepts per participant were used to control for repeated measurements. The inclusion of the variables of interest [VSS, MSS, $\log(VSS)$, $\log(MSS)$] and their interaction with Context was evaluated by comparing these models [e.g., $RT \sim VSS * Context + (1|Participant)$] with the baseline model, containing only the Context as a fixed effect [$RT \sim Context + (1|Participant)$]. All variables of interest yielded significant improvements in log-likelihood over the baseline model [VSS: $\Delta\chi^2$ (2 df)=215, $p < 2.2e-16$; $\log(VSS)$: $\Delta\chi^2$ (2 df)=130, $p < 2.2e-16$; MSS: $\Delta\chi^2$ (2 df)=320, $p < 2.2e-16$; $\log(MSS)$: $\Delta\chi^2$ (2 df)=459, $p < 2.2e-16$]. In this context, the χ^2 statistics represent the difference in deviance between successive models.

Linear and logarithmic dependencies were estimated for RT as a function of visual and memory set sizes (VSS and MSS, respectively) on a single-participant basis. For visual search, the classic linear dependence between RT and visual set size was replicated (Context Absent $R^2=0.54 \pm 0.03$; Context Present $R^2=0.49 \pm 0.03$). For memory search, a positive logarithmic relationship between RT and memory set size was observed (Context Absent $R^2=0.54 \pm 0.03$; Context Present $R^2=0.54 \pm 0.03$). Figure 2 exhibits the regression fits constructed for the average RT as a function of set sizes. The Akaike information criterion values were used to directly compare the linear and logarithmic models. Given the improvements in goodness of fit shown by models containing set size over the baseline model, comparisons were made for the full models. In visual search, the AIC value for the linear model (AIC_vs_linear: 10636) was lower than that for the logarithmic model (AIC_vs_log: 10717), indicating that the linear model in visual search better fits the data. Conversely, in memory search, the AIC value for the logarithmic model (AIC_ms_log: 10388)

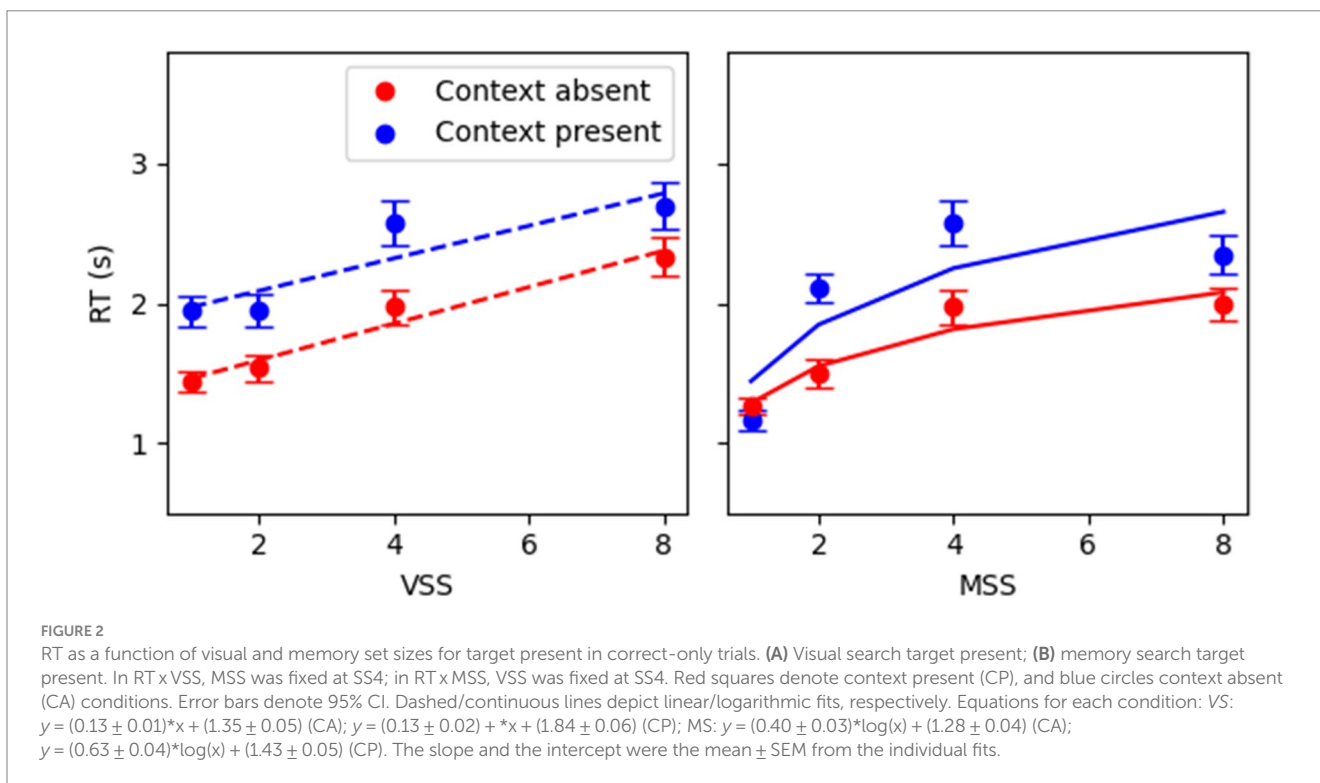


TABLE 1 Search efficiency.

Predictors		RT		
		Estimate	CI	p-value
Visual search	(Intercept)	1.23	1.11–1.34	<0.001
	Context present	0.49	0.37–0.62	<0.001
	VSS	0.12	0.10–0.14	<0.001
	Context present × VSS	−0.02	−0.05–0.01	0.133
	ICC (N = 110)	0.15		
	AIC	10,636		
Memory search	(Intercept)	1.27	1.16–1.38	<0.001
	Context present	0.10	−0.01–0.22	0.084
	log(MSS)	0.38	0.32–0.44	<0.001
	Context present × log(MSS)	0.24	0.15–0.33	<0.001
	ICC (N = 110)	0.16		
	AIC	10,388		

Both LMMs were built using the single trials from all participants with participants as random effects (N observations = 3,850). ICC, intraclass correlation coefficient. AIC, Akaike's information criterion.

was lower than that for the linear model (AIC_ms_linear: 10531). This indicates that the logarithmic model in memory search provides a more parsimonious fit to the data than the linear model.

Correct-trial RT analysis (Figure 2) presented a positive slope with visual set size ($p < 0.001$, Table 1), larger RTs when the context was present ($p < 0.001$, Table 1), and no interaction between them ($p = 0.13$, Table 1) indicating that, for target present trials, both context conditions have similar efficiencies. Moreover, the main effects were still significant in a larger model including target absent trials ($p < 0.001$, Supplementary Table S2). In this case, the interaction showed a significant effect ($p < 0.01$, Supplementary Table S2) pushed by target absent trials (Supplementary Figure S3). The target presence showed no significant main effect but significant interactions with the other co-variables.

When considering memory search, the context did not present a significant effect by itself (Table 1) but an interaction with the logarithm of the set size $\log(\text{MSS})$ ($p < 0.0001$, Table 1) indicating stronger $\log(\text{MSS})$ effect when the context was present. The $\log(\text{MSS})$ had a significant main effect by itself ($p < 0.0001$, Table 1). The significant effects were replicated when also considering the target absent trials (Supplementary Table S2), as well as the main effect of context and the main effect of target presence.

Similar models were built for accuracy as the dependent variable. There were significant main effects of visual set size and memory set size (Supplementary Table S3). No significant main effect of context, or interaction with visual/memory set sizes was observed. This suggests that participants followed a strategy in which they spent more time to locate each item but the extent to which they explored the image was the same. Both visual set size and memory set size odds ratios indicate a significant negative relation with the accuracy (odds < 1).

2.3 Discussion

Experiment 1 aimed to investigate the potential effects of context and trial-by-trial memorization on hybrid search

performance. The results replicated the characteristic RT signatures seen in the existent lab-based literature (see Wolfe, 2020), as well as the typical decrease in accuracy as set sizes increased, consistent with results observed in previous hybrid search studies (Drew et al., 2017). Generally, as RT rose, accuracy diminished. Trial-by-trial memorization and context did not qualitatively affect RT. Whereas in visual search RT increased linearly as the visual set size rose, in memory search RT increased logarithmically as memory set size escalated, with log fits exhibiting a smaller AIC score, thus providing a more parsimonious fit to the data, than linear fits. In target present conditions, the results qualitatively replicated the characteristic RT signatures of hybrid search even when context was present, and memorization occurred trial-by-trial.

In visual search, the canonical linear relationship between reaction time and set size seen in serial visual search literature and hybrid search literature was replicated (Treisman and Gelade, 1980; Wolfe, 2007, 2012a), whereas in memory search, the logarithmic relationship between reaction time and memory set size seen in traditional memory search studies and hybrid search was also observed here (Wolfe, 2012a; Nosofsky et al., 2014b; Boettcher and Wolfe, 2015). Memorization of all-new targets trial by trial does not seem to qualitatively affect the way attention is deployed and visual information is processed in non-efficient visual search (Shiffrin and Schneider, 1977).

Indeed, Schneider and Shiffrin (1977) demonstrated how manipulating memory lists to change (all-new mapping) or remain the same [consistent-mapping conditions—as seen in Wolfe (2012a) and existent hybrid search literature] on every trial provokes dramatic differences in accuracy patterns and RT efficiency in hybrid search but not RT signatures. A previously memorized target set used repeatedly throughout visual search trials can create memory reinstatement effects in memory search—that is, helps increase the distinction between memory representations and distractors (Nosofsky, 1986). Nosofsky et al. (2014b) found mean RTs are much faster, and error rates are much

lower in the consistent-mapping condition than in an all-new condition in a memory search task. Consistently, this study observed lower accuracy in comparison with previous hybrid search studies, which had pre-tested fixed memory sets throughout trials (e.g., Wolfe, 2012a; Cunningham and Wolfe, 2014).

When context was present, we observed higher RT intercepts and similar RT slopes compared to context absent conditions. The context in our stimuli did not violate scene syntax or scene semantics. Prior studies have shown that syntactic and semantic violations affect attention allocation in scene processing. Indeed, syntactic inconsistencies (e.g., floating objects) and scene inconsistencies (e.g., searching for a toaster on the floor) have been shown to impair scene guidance of attention when compared to semantically consistent objects in scenes, such as a toaster on a table (Wolfe, 2020). Differences in search slopes might emerge if distractors in our stimuli set were placed in more implausible locations than the targets, as observed by Neider and Zelinsky (2006) in a visual search task. Therefore, future studies using stimuli sets with varying degrees of semantic association between the context and memory set items are needed to uncover the influence of context on hybrid search.

While the main analyses were conducted on target present trials, we also analyzed target absent trials. In target and context absent conditions, RT shapes replicated the linear dependence between visual set size and RT and the logarithmic dependence between memory set size and RT (Wolfe, 2012b; Boettcher and Wolfe, 2015). Accuracy overall also decreased as set sizes increased. However, the quality of the model adjustments in these conditions was worse than the ones when the target was present. As explored below, this could be linked to participants' adoption of adaptive strategies for search termination.

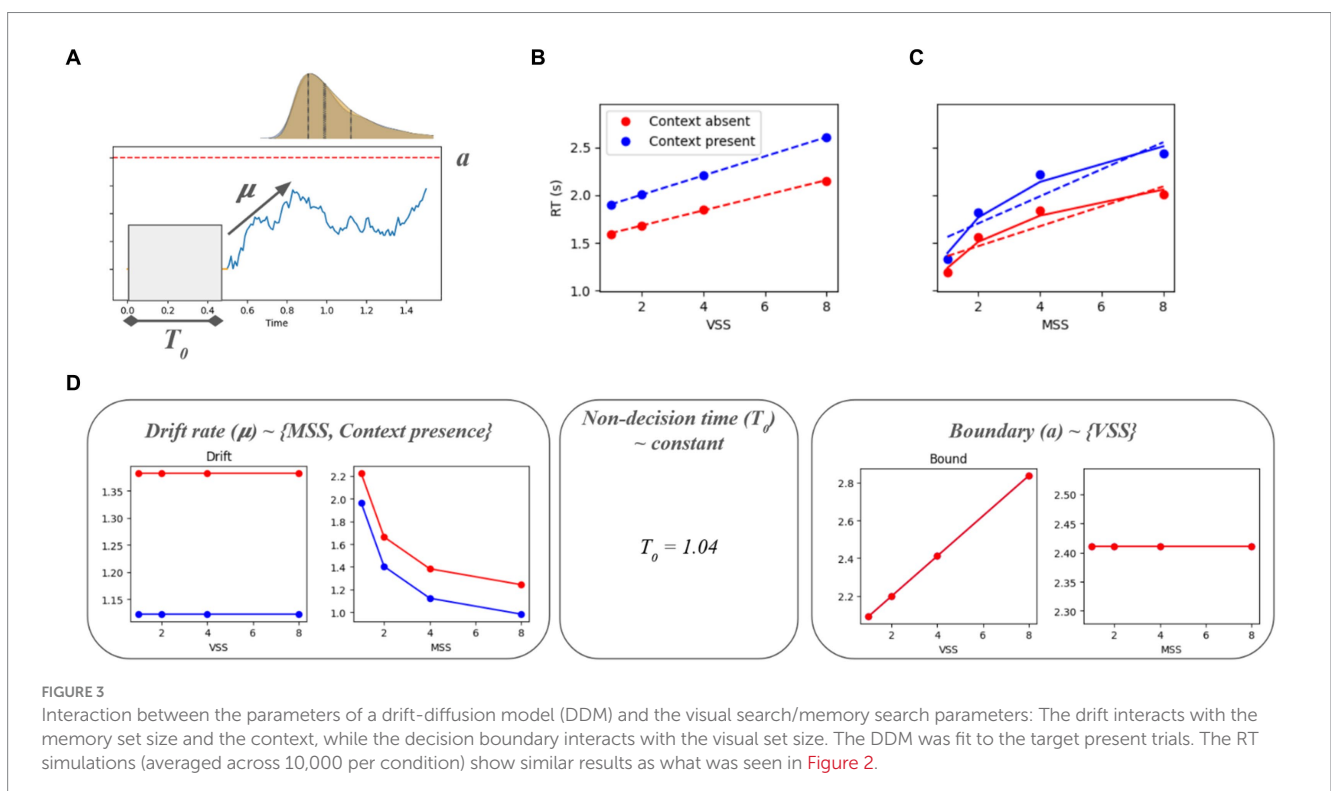
3 A drift-diffusion computational model of hybrid search

3.1 Methods

A DDM was implemented using the Python package pyddm 0.7.0 (Shinn et al., 2020). At first, the model was fitted using all the target present data using `pyddm.models.loss.LossRobustLikelihood` as the loss function and differential evolution as the fitting method (the rest of the parameters were the default). In this case, the goal was not to find the best, generalizable model but to explain its parameter values in terms of the experimental variables (MSS, VSS, and context). This would allow a better interpretation of the effects in terms of the model's parameters. The parameters considered were the drift rate (μ), the boundary (a), and the non-decision time (t_0). To keep the model simple and maintain interpretability, the drift rate was assumed to depend only on the MSS and the context, while the boundary was assumed to depend on the VSS (Figure 3). There were no constraints on the non-decision time. As a form of validation, RT simulations were performed to assess whether the $RT \times VSS$ and $RT \times MSS$ curves were like the experimental data or not. In the second step, the model was fitted including target absent data, and the boundary and non-decision time could vary with the target's presence (Supplementary Figure S4).

3.2 Results

Figure 3 shows the results of a drift-diffusion model fit for each of the conditions. Panel A describes μ (drift rate), a (boundary), T_0 (non-decision time), and how they affect a typical drift-diffusion model.



A comparison of Panels B and C with [Figure 2](#) shows that the main behavioral effects can be replicated, although an additional significant context X VSS interaction emerges. Panel D shows how those parameters relate to each experimental variable (MSS, VSS, and context). The drift rate has a multiplicative inverse relationship with the memory set size, while the context applies an intercept to the drift rate. The non-decision time remains constant across experimental conditions. The boundary is linear in terms of the visual set size. In this model, there is a constant relationship between boundary and memory set size, boundary and context, and drift rate and visual set size.

3.3 Discussion

Previous computational models of memory/visual search ([Moran et al., 2013](#); [Nosofsky et al., 2014b](#)) have provided useful information related to both search processes and termination, which was later incorporated into general models of visual search, such as Wolf's Guided Search 6.0. However, these models are particularly suitable for setups where the stimuli are artificial (i.e., white background with geometric figures) and are not defined in setups with natural scenes (where there is an inherent correlation between different regions in the image as well as added noise). This can be overcome by using data-driven approaches such as the drift-diffusion model we implemented here.

In a drift-diffusion model, the drift rate is associated with the amount of information a subject accumulates in a single time step. A higher drift rate would result in a lower RT. We observed that the drift rate goes down as the MSS goes up in a logarithmic fashion ([Figure 3](#)). One simple interpretation is that the higher the MSS, the more time is needed to process/remember the memory set, so in a fixed time interval participants will be less certain of what they are looking for when the MSS is large. The context applies a bias to the drift rate. This could be related to the fact that when looking at objects with a white background, there will typically be a few fixations in the background. However, when there is context present, there are contextual cues in the background that guide our search and the stimulus is noisier, which would add an overhead to the processing of the image.

The boundary is related to the amount of evidence needed to make a decision. The further it is from 0, the higher will be the RT. We observed that the boundary is linearly proportional to the visual set size ([Figure 3](#)). Interpreting the evidence as the proportion of the image explored until the target object is found, more objects will require increased attention when the visual set size is high, which in turn would result in a larger portion of the image being explored. This means that the higher the set size, the more evidence is needed to decide that the object has been found.

4 Experiment 2: a change detection and a go/no-go task to explore individual differences in hybrid search

4.1 Methods

4.1.1 Participants

The participants were the same as in Experiment 1. Typically, Experiment 2 was conducted after a short break following Experiment 1.

4.1.2 Materials

Visual working memory capacity was measured using the classic change detection task (CDT) adapted from previous studies ([Xu et al., 2018](#); [Balaban et al., 2019](#)). It involved 120 trials with set size 4 ($N=60$) and set size 6 ($N=60$) of colored blocks. To measure inhibitory control, the well-established go/no-go (GNG) task was employed. It was based on the parameters of previous studies ([Wessel, 2018](#); [Young et al., 2018](#)). To isolate inhibitory control from other cognitive processes, the 150 trials varied between a blue circle in go trials ($N=120$) and an orange circle in no-go trials ($N=30$).

4.1.3 Procedure

Trials began with four or six color blocks, followed by a fixation cross and, subsequently, a detection screen with one color block (for timeframes, see the [Supplementary Figure S1](#)). If the same color and location as in the initial display appeared, the "k" key was pressed. If a different color and/or location popped up, "1" was pushed. The next trial automatically started after a response without feedback. Changes occurred in 50% of trials. Once the 120 CDT trials were completed (approximately 10 min), the Go/No-Go task instructions appeared. After four practice trials with feedback, the Go/No-Go task started by displaying either blue (go stimuli; size [0.15, 0.15]cm) or orange circles (no-go stimuli; size [0.15, 0.15] cm) at a 4:1 go/no-go ratio. For go stimuli, the space bar was pressed, and for no-go stimuli, that action had to be inhibited. With an intertrial interval of 450 ms, 150 trials lasted approximately 3 min (see the [Supplementary Figure S2](#)). Part 2 lasted approximately 15 min.

4.1.4 Data analysis

Simple linear regressions were performed to assess the potential effects of individual differences in visual working memory capacity and inhibitory control on hybrid search performance. Here, the independent variables were the memory capacity K and the response bias c (see [Supplementary material](#)), and the dependent variables included the RT slopes and intercepts for each condition (context absent or present, and target absent or present). Age was also checked for its potential impact on search; however, it did not show a significant result. While some measures in this study are from direct observations, visual working memory capacity and inhibitory control are measured by K and c , respectively. K is the estimate of an individual's VWM capacity, whereas c is seen as a measure of decision/response bias. These are calculated through [Cowan's \(2001\)](#) formula and SDT analysis ([Green and Swets, 1966](#)), respectively.

Scikit-learn (version 1.3.0) and statsmodels (version 0.14.1) in Python (version 3.10.5) were used for data analysis.

4.2 Results

To ensure the validity of the visual working memory capacity K on the change detection task experiment, the normality of K was assessed using a Kolmogorov-Smirnov normality test. As in previous studies ([Balaban et al., 2019](#)), K was found to be normally distributed [$D(110)=0.065, p>0.05$]. This study's K replicates two pivotal characteristics concerning VWM capacity—(1) capacity-limited characteristic ($M=2.31$) and (2) significant individual differences denoted by $SD=0.82$

(Balaban et al., 2019). It is consistent with China-based (K:2.14) (Xu et al., 2018) and American studies (K:2.55) (Fukuda et al., 2016).

Signal detection theory analysis was applied to calculate the decision bias measure in the Go/No-Go task. A negative c mean ($M = -0.23$, $SD = 0.58$) was achieved, consistent with existing literature (Young et al., 2018).

A correlation analysis was performed to investigate the impact of individual differences in memory capacity K , decision bias c , and false alarms on hybrid search performance, measured through RT intercepts and slopes. For target present trials, none of the correlations (see Figure S5) were statistically significant, with values ranging from -0.18 to 0.18 . However, for target absent without context, memory capacity and RT slopes were positively correlated ($R = 0.18$, $p < 0.05$). Participants' decision bias c did not predict RT in any condition when the target was present, but it did predict RT intercepts when the target was absent ($R = 0.22$ for context present and $R = 0.20$ for context absent). Nevertheless, all these correlations explained little variability in hybrid search performance (maximum $R^2 = 0.05$).

4.3 Discussion

Experiment 2 assessed the possible interplay of visual working memory capacity and inhibitory control in hybrid search.

The general lack of relationship between visual working memory capacity (VWM) and hybrid search performance measures is inconsistent with accounts that propose that search/target templates reside in VWM to (1) bias attentional deployment to goal-relevant objects (Desimone and Duncan, 1995; Soto et al., 2008) and (2) be compared to suitor objects (Bundesen, 1990). If that was so, higher VWM capacity individuals, supposedly having larger storing capacities or better ability in manipulating attentional resources (see Luck and Vogel, 2013), would search more efficiently as set sizes increase, than observers with lower VWM capacity (Sobel et al., 2007). However, these results are partially consistent with the sparse hybrid search literature that previously assessed VWM capacity's role. Indeed, Drew et al. (2016) used a dual-task paradigm with CDTs in between HS tasks and found that performing a hybrid search task diminished VWM capacity by a fixed amount (e.g., one slot) regardless of set size variation. Hence, if it is assumed that a "one-item" channel/path must successively move visual items to long-term memory, VWM interaction would not be dependent on set size, that is, only a fixed amount of each individual's VWM capacity would be used for the task. This would be consistent with the largely non-significant results seen here between VWM capacity differences and hybrid search RT slopes.

An important difference between this study and other studies including (Drew et al., 2016) is that memory sets here were not kept constant, changing trial by trial. In the visual search literature, some dual-task studies indicated that loading VWM, with a change detection task embedded in a VS task, solely impacted search when target sets changed per trial, not when they were kept constant (Woodman et al., 2007; Woodman and Arita, 2011). Behaviorally, they observed a slowing of RT slopes when loading working memory. However, electrophysiological data suggested an effect on working memory representation maintenance (Fukuda and Vogel, 2009). While the results

presented are inconsistent with these findings, this points to the importance of evaluating electrophysiological data alongside behavior to pull apart the meaning of behavioral observations.

Studies investigating the effect of individual differences in inhibitory control (IC) on search, especially those with direct inhibitory control measures (Clarke et al., 2022), are markedly scarce. Given inhibition's clear importance in visual search (Beck and Kastner, 2009) and search termination (Moran et al., 2013), this study investigated the IC's role in hybrid search. The results showed that individuals with higher IC (lower negative bias) had significantly higher RT intercepts than participants with lower IC only when the target was absent (Supplementary Figure S5). These results suggest that more inhibitory control is needed to remain in search when there is no target on the visual display. This is consistent with Moran et al. (2013) competitive-guided search model that advances inhibitory links between the priority search map and the quitting unit that terminates the search when selected, and other memory search models that propose a diffusion stage with laterally inhibiting racing target items (i.e., Cunningham and Wolfe, 2014; Nosofsky et al., 2014b). In target present trials, the activation of the priority map is automatically increased, which leads to higher inhibition of the quit unit in comparison to target absent trials (Moran et al., 2013). In other words, target absent conditions inherently increase quitting probability.

5 General discussion and conclusion

This study contributes to the understanding of various elements involved in typical real-world search scenarios, where multiple targets are searched for, contextual information is available, and searching occurs after a single exposure. Unlike traditional search scenarios with items presented against a blank background, the fundamental reaction time signatures of hybrid search—linear increase with visual set size and logarithmic increase with memory set size—are preserved. Individual differences in working memory and inhibitory control only showed a modest impact on search termination when the target was absent, aligning with established visual search models such as the guided search model. In addition, we showed that a simple computational drift-diffusion model can reproduce the primary behavioral results of a hybrid search. To our knowledge, our study is the first hybrid search study to present search items against a photorealistic scene background rather than the commonly used blank backgrounds. This approach introduces a more realistic context, which is crucial for understanding search behaviors in real-world settings. Indeed, our findings have potential implications for real-world tasks, such as airport security screening, where officers must remember a set of prohibited items (memorized targets) among numerous distractors, and medical imaging analysis, where radiologists use their expertise to identify abnormal patterns in images.

While this study has made contributions toward a better understanding of search in real life, it also had some limitations. First, this study did not systematically manipulate the categories in memory sets (i.e., in each trial, targets and distractors belonged to the same category) and scene semantics/syntactics (i.e., there were no objects

placed in scenes with semantic/syntactic violations). Second, the experiment design does not allow direct comparisons between the efficiency and accuracy in all-new mapping and consistent-mapping conditions as this would require a separate comparison group. Since the mapping of the stimuli was not manipulated, the comparisons made here are related to the qualitative nature of the relationship between reaction times and set sizes. Third, the amount of time observers fixated on *individual* items was not controlled. This can also affect memory strength and working memory consolidation for each target item (Donkin and Nosofsky, 2012). Future studies could include eye tracking to better understand how memory strength impacts memory search via dwell times, and how observers encode target items in memory. This can help disentangle VWM's role in hybrid search and understand memory strength impact, which is not controlled for in the existent hybrid search corpus (Nosofsky et al., 2014b). This could also give more information to the model to work with, or even allow the use of scanpath prediction models in conjunction with, a drift-diffusion model (Travi et al., 2022). Fourth, the current drift-diffusion model does not explain the variability across subjects or stimuli. Finally, as an online study, environmental factors were uncontrolled, and attentional engagement throughout the experiment could not be assessed. This could have impacted the search in several ways including participants' appraised value of continuing the search if there were external attentional demands present. Contextual information in search should be further investigated in lab-based studies to strengthen the results reported here.

Altogether, our study exposes an intricate interplay of various behavioral mechanisms in hybrid search. While the results reported here provide some valuable insights, some key questions remain unanswered, particularly concerning the influence of scene semantics on hybrid search and the little role that visual working memory seems to have on hybrid search. To address these intriguing aspects, future research can leverage recent technical advancements, such as concurrent M-EEG and eye movement recordings (Dimigen and Ehinger, 2021; Care et al., 2023). The integration of these cutting-edge methodologies promises to unveil the underlying physiological mechanisms driving hybrid search, further deepening our understanding of this complex cognitive phenomenon.

Data availability statement

The datasets and code presented in this study are available on https://github.com/NeuroLIAA/HybridSearch_Online/. Experiments can be accessed at https://run.pavlovia.org/mison/cdt_gng_ius12/html/ and at https://run.pavlovia.org/isonlab/hybrid_search/html/.

Ethics statement

The studies involving humans were approved by University of Nottingham School of Psychology Ethics Panel (ethics approval: S1240). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AB: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Data curation. GR: Writing – review & editing, Validation, Software, Methodology, Formal analysis. AR: Conceptualization, Writing – review & editing. JK: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. MI: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by CONICET (PIP 11220150100787CO) and ARL (Cooperative Agreement Number W911NF1920240 and W911NF2120237 awarded to MI and JK).

Acknowledgments

The authors would like to thank Mingqian Luo for contributing to data collection.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2024.1436564/full#supplementary-material>

References

- Balaban, H., Fukuda, K., and Luria, R. (2019). What can half a million change detection trials tell us about visual working memory? *Cognition* 191:103984. doi: 10.1016/J.COGNITION.2019.05.021
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/JSS.V067.I01
- Beck, D. M., and Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vis. Res.* 49, 1154–1165. doi: 10.1016/J.VISRES.2008.07.012
- Biederman, I. (1977). On processing information from a glance at a scene: some implications for a syntax and semantics of visual processing. In Proceedings of the ACM/SIGGRAPH Workshop on User-Oriented Design of Interactive Graphics Systems, UODIGS 1976, 75–88
- Boettcher, S. E. P., Drew, T., and Wolfe, J. M. (2013). Hybrid search in context: how to search for vegetables in the produce section and cereal in the cereal aisle. *Vis. Cogn.* 21, 678–682. doi: 10.1080/13506285.2013.844959
- Boettcher, S. E. P., Drew, T., and Wolfe, J. M. (2018). Lost in the supermarket: quantifying the cost of partitioning memory sets in hybrid search. *Mem. Cogn.* 46, 43–57. doi: 10.3758/s13421-017-0744-x
- Boettcher, S. E. P., and Wolfe, J. M. (2015). Searching for the right word: hybrid visual and memory search for words. *Atten. Percept. Psychophys.* 77, 1132–1142. doi: 10.3758/s13414-015-0858-9
- Bujia, G., Sclar, M., Vita, S., Solovey, G., and Kamienkowski, J. E. (2022). Modeling human visual search in natural scenes: a combined Bayesian searcher and saliency map approach. *Front. Syst. Neurosci.* 16:882315. doi: 10.3389/fnsys.2022.882315
- Bundesden, C. (1990). A theory of visual attention. *Psychol. Rev.* 97, 523–547. doi: 10.1037/0033-295X.97.4.523
- Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304. doi: 10.1177/0049124104268644
- Care, D., da Fonseca, M., Ison, M. J., and Kamienkowski, J. E. (2023). Active search signatures in a free-viewing task exploiting concurrent EEG and eye movements recordings. *Eur. J. Neurosci.* 58, 2563–2578. doi: 10.1111/EJN.16057
- Chun, M. M., and Wolfe, J. M. (1996). Just say no: how are visual searches terminated when there is no target present? *Cogn. Psychol.* 30, 39–78. doi: 10.1006/COGP.1996.0002
- Clarke, A. D. F., Irons, J. L., James, W., Leber, A. B., and Hunt, A. R. (2022). Stable individual differences in strategies within, but not between, visual search tasks. *Q. J. Exp. Psychol.* 75, 289–296. doi: 10.1177/1747021820929190
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114. doi: 10.1017/S0140525X01003922
- Cunningham, C. A., and Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *J. Exp. Psychol. Gen.* 143, 1585–1599. doi: 10.1037/A0036313
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255. doi: 10.1109/CVPR.2009.5206848
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18(Volume 18, 1995, 193–222. doi: 10.1146/ANNUREV.NE.18.030195.001205/CITE/REFWORKS
- Dimigen, O., and Ehinger, B. V. (2021). Regression-based analysis of combined EEG and eye-tracking data: theory and applications. *J. Vis.* 21:3. doi: 10.1167/JVOV.21.1.3
- Donkin, C., and Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short- and long-term recognition. *The Journal of Psychological Science* 23, 625–634. doi: 10.1177/0956797611430961
- Drew, T., Boettcher, S. E. P., and Wolfe, J. M. (2016). Searching while loaded: visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. *Psychon. Bull. Rev.* 23, 201–212. doi: 10.3758/s13423-015-0874-8
- Drew, T., Boettcher, S. E. P., and Wolfe, J. M. (2017). One visual search, many memory searches: an eye-tracking investigation of hybrid search. *J. Vis.* 17:5. doi: 10.1167/17.11.5
- Fukuda, K., and Vogel, E. K. (2009). Human variation in overriding attentional capture. *J. Neurosci.* 29, 8726–8733. doi: 10.1523/JNEUROSCI.2145-09.2009
- Fukuda, K., Woodman, G. F., and Vogel, E. K. (2016). “Individual differences in visual working memory capacity: contributions of attentional control to storage” in *Mechanisms of sensory working memory attention and performance XXV*. (London: Elsevier Academic Press), 105–119.
- Green, D. M., and Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley.
- Henderson, J., and Ferreira, F. (Eds.) (2004). “Scene perception for psycholinguists” in *The interface of language, vision, and action: eye movements and the visual world* (New York: Psychology Press), 1–58.
- Henderson, J. M., Malcolm, G. L., and Schandl, C. (2009). Searching in the dark: cognitive relevance drives attention in real-world scenes. *Psychon. Bull. Rev.* 16, 850–856. doi: 10.3758/PBR.16.5.850
- Ison, M. J., Quian Quiroga, R., and Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron* 87, 220–230. doi: 10.1016/j.neuron.2015.06.016
- Kummerer, M., Wallis, T. S. A., and Bethge, M. (2018). “Saliency benchmarking made easy: separating models, maps and metrics”, in *Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science*. eds. Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y. Cham: Springer.
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al (2014). Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*. Springer International Publishing, 740–755.
- Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846
- Luck, S. J., and Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* 17, 391–400. doi: 10.1016/j.tics.2013.06.006
- Manginelli, A. A., Langer, N., Klose, D., and Pollmann, S. (2013). Contextual cueing under working memory load: Selective interference of visuospatial load with expression of learning. *Attention, Perception, & Psychophysics*, 75, 1103–1117. doi: 10.3758/s13414-013-0466-5
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The Unity and Diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/COGP.1999.0734
- Moran, R., Zehetleitner, M., Müller, H. J., and Usher, M. (2013). Competitive guided search: meeting the challenge of benchmark RT distributions. *J. Vis.* 13:24. doi: 10.1167/13.8.24
- Myers, C. E., Interian, A., and Moustafa, A. A. (2022). A practical introduction to using the drift diffusion model of decision-making in cognitive psychology, neuroscience, and health sciences. *Front. Psychol.* 13:1039172. doi: 10.3389/fpsyg.2022.1039172
- Neider, M. B., and Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vis. Res.* 46, 614–621. doi: 10.1016/J.VISRES.2005.08.025
- Neisser, U. (1964). Visual search. *Sci. Am.* 210, 94–102. doi: 10.1038/SCIENTIFICAMERICAN0664-94
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *J. Exp. Psychol. Gen.* 115, 39–57. doi: 10.1037/0096-3445.115.1.39
- Nosofsky, R. M., Cao, R., Cox, G. E., and Shiffrin, R. M. (2014a). Familiarity and categorization processes in memory search. *Cogn. Psychol.* 75, 97–129. doi: 10.1016/J.COGP.2014.08.003
- Nosofsky, R. M., Cox, G. E., Cao, R., and Shiffrin, R. M. (2014b). An exemplar-familiarity model predicts short-term and long-term probe recognition across diverse forms of memory search. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 1524–1539. doi: 10.1037/XLM0000015
- Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Mem. Cogn.* 3, 519–526. doi: 10.3758/BF03197524
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51, 195–203. doi: 10.3758/s13428-018-01193-y
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/NECO.2008.12-06-420
- Rosenholtz, R., Li, Y., and Nakano, L. (2007). Measuring visual clutter. *J. Vis.* 7:17. doi: 10.1167/7.2.17
- Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66. doi: 10.1037/0033-295X.84.1.1
- Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* 84, 127–190. doi: 10.1037/0033-295X.84.2.127
- Shinn, M., Lam, N. H., and Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife* 9:e56938. doi: 10.7554/eLife.56938
- Sobel, K. V., Gerrie, M. P., Poole, B. J., and Kane, M. J. (2007). Individual differences in working memory capacity and visual search: the roles of top-down and bottom-up processing. *Psychon. Bull. Rev.* 14, 840–845. doi: 10.3758/BF03194109
- Soto, D., Hodsoll, J., Rotshtein, P., and Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends Cogn. Sci.* 12, 342–348. doi: 10.1016/J.TICS.2008.05.007

- Torralba, A., Oliva, A., Castelano, M. S., and Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol. Rev.* 113, 766–786. doi: 10.1037/0033-295X.113.4.766
- Travi, F., Ruarte, G., Bujia, G., and Kamienskowski, J. E. (2022). ViSioNS: visual search in natural scenes benchmark. *Adv. Neural Inf. Proces. Syst.* 35, 11987–12000. doi: 10.5555/3600270.3601141
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 194–214. doi: 10.1037/0096-1523.8.2.194
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5
- Treisman, A., and Sato, S. (1990). Conjunction search revisited. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 459–478. doi: 10.1037//0096-1523.16.3.459
- Vö, M. L. H., and Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *J. Vis.* 9:24. doi: 10.1167/9.3.24
- Wessel, J. R. (2018). Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. *Psychophysiology* 55:e12871. doi: 10.1111/PSYP.12871
- Wolfe, J. M. (2007). “Guided search 4.0: current Progress with a model of visual search”, in *Integrated Models of Cognitive Systems* ed. Gray W. D. New York: Oxford Academic, 99–119.
- Wolfe, J. M. (2012a). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science* 23, 698–703. doi: 10.1177/0956797612443968
- Wolfe, J. M. (2012b). When do I quit? The search termination problem in visual search. *Neb. Symp. Motiv.* 59, 183–208. doi: 10.1007/978-1-4614-4794-8_8
- Wolfe, J. M. (2020). Visual search: how do we find what we are looking for? *Ann. Rev. Vis. Sci.* 6, 539–562. doi: 10.1146/ANNUREV-VISION-091718-015048/CITE/REFWORKS
- Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., and Sherman, A. M. (2011a). Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophys.* 73, 1650–1671. doi: 10.3758/s13414-011-0153-3
- Wolfe, J. M., Boettcher, S. E. P., Josephs, E. L., Cunningham, C. A., and Drew, T. (2015). You look familiar, but i don't care: lure rejection in hybrid visual and memory search is not based on familiarity. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 1576–1587. doi: 10.1037/XHP0000096
- Wolfe, J. M., Vö, M. L. H., Evans, K. K., and Greene, M. R. (2011b). Visual search in scenes involves selective and nonselective pathways. *Trends Cogn. Sci.* 15, 77–84. doi: 10.1016/j.tics.2010.12.001
- Woodman, G. F., and Arita, J. T. (2011). Direct electrophysiological measurement of attentional templates in visual working memory. *Psychol. Sci.* 22, 212–215. doi: 10.1177/0956797610395395
- Woodman, G. F., Luck, S. J., and Schall, J. D. (2007). The role of working memory representations in the control of attention. *Cereb. Cortex* 17, i118–i124. doi: 10.1093/CERCOR/BHM065
- Xu, Z., Adam, K. C. S., Fang, X., and Vogel, E. K. (2018). The reliability and stability of visual working memory capacity. *Behav. Res. Methods* 50, 576–588. doi: 10.3758/s13428-017-0886-6
- Young, M. E., Sutherland, S. C., and McCoy, A. W. (2018). Optimal go/no-go ratios to maximize false alarms. *Behav. Res. Methods* 50, 1020–1029. doi: 10.3758/s13428-017-0923-5