

Differential effects of visual uncertainty and contextual guidance on perceptual decisions: Evidence from eye and mouse tracking in visual search

Boris Quétard

Clermont University, Blaise Pascal University, LAPSCO,
Clermont-Ferrand, France
CNRS, UMR 6024, LAPSCO, Clermont-Ferrand, France



Jean Charles Quinton

University Grenoble Alpes, LJK, Grenoble, France
CNRS, UMR 5224, LJK, Grenoble, France



Martial Mermillod

University Grenoble Alpes, LPNC,
F-38000 Grenoble, France
CNRS, LPNC, F-38000 Grenoble, France



Laura Barca

Institute of Cognitive Sciences and Technologies,
CNR, Rome, Italy



Giovanni Pezzulo

Institute of Cognitive Sciences and Technologies,
CNR, Rome, Italy



Michèle Colomb

CEREMA,
Département Laboratoire de Clermont-Ferrand,
Clermont-Ferrand, France



Marie Izaute

Clermont University, Blaise Pascal University, LAPSCO,
Clermont-Ferrand, France
CNRS, UMR 6024, LAPSCO, Clermont-Ferrand, France



Visual search can be seen as a decision-making process that aims to assess whether a target is present or absent from a scene. In this perspective, eye movements collect evidence related to target detection and verification to guide the decision. We investigated whether, in real-world scenes, target detection and verification are differentially recruited in the decision-making process in the presence of prior information (expectations about target location) and perceptual uncertainty (noise). We used a mouse-tracking methodology with which mouse trajectories unveil components of decision-making and eye-tracking measures reflect target detection and verification. Indoor scenes were presented, including a target in usual or unusual locations or no target, and were degraded with additive noise (or no noise). Participants had to respond to the target's presence or absence. Degrading the scene delayed the decision due

to increased verification times and reduced mouse velocity. Targets in unusual locations delayed the decision and deviated mouse trajectories toward the *target-absent* response. Detection times played a major role in these effects. Thus, target detection and verification processes influence decision-making by integrating the available sources of information differently and lead to an accumulation of evidence toward both the presence of a target and its absence.

Introduction

Visual scenes from everyday life can often be very complex, including elaborate details and textures as

Citation: Quétard, B., Quinton, J. C., Mermillod, M., Barca, L., Pezzulo, G., Colomb, M., & Izaute, M. (2016). Differential effects of visual uncertainty and contextual guidance on perceptual decisions: Evidence from eye and mouse tracking in visual search. *Journal of Vision*, 16(11):28, 1–21, doi:10.1167/16.11.28.



well as a large number of objects, and they cannot be fully processed in a single glance. Thus, visual information is sampled through discrete eye fixations. As actions can be made not only to progress toward a goal, but also to reduce the uncertainty about the states of the world (Bajcsy, Aloimonos, & Tsotsos, 2016; Kirsh & Maglio, 1994), active oculomotor behavior could be directed toward minimizing uncertainty about the composition of the visual scene (Friston, 2012; Friston, Adams, Perrinet, & Breakspear, 2012; Itti & Baldi, 2009; Renninger, Verghese, & Coughlan, 2007) or object identities (Kietzmann, Geuter, & König, 2011; Kietzmann & König, 2015; Quinton, Volpi, Barca, & Pezzulo, 2014). This behavior is usually driven by particular goals, such as understanding what is depicted, gathering action-related information, or searching for a particular object (Ballard & Hayhoe, 2009; Land & Hayhoe, 2001; Rothkopf, Ballard, & Hayhoe, 2007). Most real-life tasks require fast visual processing. Therefore, efficient visual exploration has to rely on rapidly available sources of information. This includes both the visual information that can be extracted directly from the current view of the environment and prior information (e.g., knowledge and expectations about this environment) built up as a result of past exposures to similar situations, the effects of which may start even before any visual stimulation occurs.

The former was previously formalized as bottom-up saliency maps (Borji, Tavakoli, Sihite, & Itti, 2013; Itti & Koch, 2000; Itti, Koch, & Niebur, 1998), with which highly conspicuous areas of the scene (the features of which differ from surroundings) are thought to attract attention and, ultimately, the gaze. This model reproduces some experimental human data especially in free viewing (Borji et al., 2013; Itti & Koch, 2000) and predicts human gaze fixation locations better than various random models in a scene-encoding task (Foulsham & Underwood, 2008; Underwood & Foulsham, 2006). However, it predicts neither the order of the fixations in the sequence (i.e., scanpaths; Foulsham & Underwood, 2008; Underwood, Foulsham, & Humphrey, 2009) nor the oculomotor behavior in visual search tasks (Henderson, Brockmole, Castelano, & Mack, 2007; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006) or tasks in which the gaze supports navigation and goal-oriented behavior (Ballard & Hayhoe, 2009; Rothkopf et al., 2007). It has been proposed that bottom-up saliency could reflect surprise toward statistical outliers in space (Tatler, Hayhoe, Land, & Ballard, 2011) and therefore attract the gaze when relevant for the task. Considering that conspicuity-based models do not fully explain eye movements in natural behavior, other sources of information must be considered. For instance, throughout his or her life,

an observer memorizes knowledge about what objects look like and how they interact with each other and uses this knowledge to guide eye movements.

Searching for a specific object in a known scene can be made easier by memories of where this object is usually placed (Brockmole & Henderson, 2006; Brockmole & Le-Hoa Vo, 2010). Even in unknown scenes, knowledge about the laws of physics, cultural practices, or the object's function governing the object's location in the environment still apply and can be exploited. Thus, a target object located somewhere at which there is a high probability of finding it is detected and recognized more quickly than when it is placed somewhere at which there is a low probability of finding it (Malcolm & Henderson, 2010). It applies when the target is both consistent and inconsistent with the global layout of the scene (Castelano & Heaven, 2011). Hence, whether the scene is a kitchen or a bathroom, a mug would be found more easily on a sink or washbasin than on the floor. These expectations are built on a variety of sources of information, such as prior knowledge about the global layout of the scene (Castelano & Heaven, 2010; Castelano & Henderson, 2007; Pereira & Castelano, 2014), the laws of physics (Spotorno, Malcolm, & Tatler, 2014), or local information based on the relationships between the target and the other objects (Eckstein, Drescher, & Shimozaki, 2006; Pereira & Castelano, 2014; for reviews, see Võ & Wolfe, 2015; Wu, Wick, & Pomplun, 2014). Computational models that integrate contextual guidance predict eye fixations in natural scenes better than models implementing only bottom-up saliency (Ehinger, Hidalgo-Sotelo, Torralba, & Oliva, 2009; Kanan, Tong, Zhang, & Cottrell, 2009; Torralba, Oliva, Castelano, & Henderson, 2006).

How expectations stemming from the target's location influence eye movement guidance during visual search tasks has been extensively studied. However, the relationship between visual exploration and the higher-level decision-making process to which it contributes has received less attention. Although the purpose of the process is to decide whether or not the target is present in the environment, visual exploration with eye movements can be seen as a process of active evidence collection (hypothesis testing) in which each saccade reduces perceptual uncertainty and consequently enhances the quality of the choice (Friston et al., 2012). In the field of visual identification, this idea has recently been studied using ambiguous drawings representing two alternative objects or animals (Kietzmann et al., 2011) or ambiguous schematic figures interpolated between prototypical animals (Quinton et al., 2014). These two studies show how active eye behavior can disambiguate a figure with the former using an eye-tracking approach and the later a computational approach coupled with a mouse-track-

ing task. More specifically, Quinton et al. (2014) used a binary forced-choice task consisting of categorizing (un)ambiguous prototypical animals. They showed how the two alternative responses attracted mouse movement trajectories produced by participants in a way that is proportional to putative “votes” for the alternative responses at each time unit and used a computational model implementing predictive coding principles (Rao & Ballard, 1999) to reproduce analogous choice trajectories and next gaze location on the stimulus (see also Catenacci Volpi, Quinton, & Pezzulo, 2014).

This example illustrates how, by tracking mouse movements or hand movements during a choice, it is possible to measure distinct aspects of the underlying perceptual decision-making process in a binary forced-choice task (Dale, Kehoe, & Spivey, 2007; Freeman & Ambady, 2010; Song & Nakayama, 2009; Spivey, Grosjean, & Knoblich, 2005). In a typical mouse-tracking task, participants have to categorize a visual stimulus by clicking on the response buttons located in the top left and top right corners of the screen as mouse trajectories are recorded. Therefore, unlike with basic binary forced-choice tasks, the decision-making process is not only inferred through the response times (RTs) and error rates and also through various trajectory measurements whereby continuous aspects of the decision-making process can be captured (Freeman, Dale, & Farmer, 2011; Lepora & Pezzulo, 2015). Among the basic mouse-tracking measurements, the degree of deviation of each trajectory, estimated as the geometric area (area under the curve, AUC) between the actual trajectory and ideal response movement (line stretching from the start point to the accurate response), can be used as a global measure of the bias toward the alternative response in the decision-making process (Freeman & Ambady, 2010). Furthermore, trajectory deviation and mouse velocity can be analyzed as a function of time, providing useful information about how competition between alternatives unfolds over time (Farmer, Cargill, Hindy, Dale, & Spivey, 2007; Freeman & Ambady, 2011; Freeman, Ma, Han, & Ambady, 2013).

The mouse-tracking methodology has been used extensively with visual material, especially to study facial recognition of gender (Freeman, Ambady, Rule, & Johnson, 2008; Johnson, Freeman, & Pauker, 2012) and ethnicity (Freeman & Ambady, 2011; Freeman, Pauker, Apfelbaum, & Ambady, 2010) as well as the influence of stereotypes on face recognition (Freeman & Ambady, 2009; Freeman, Penner, Saperstein, Scheutz, & Ambady, 2011). It has also been used in the fields of numerical cognition (Faulkenberry, 2014; Marghetis, Núñez, & Bergen, 2014), psycholinguistics (Barca, Benedetti, & Pezzulo, 2016; Barca & Pezzulo, 2012, 2015; Farmer et al., 2007; Morett & MacWhin-

ney, 2013), and memory (Papesh & Goldinger, 2012). In visual categorization, the effectiveness of this methodology was assessed for discrimination of ambiguous categories of animals (Dale et al., 2007) and ambiguous visual stimuli representing animals (Quinton et al., 2014).

In the present study, we studied visual search by combining mouse-tracking and eye-tracking methods in order to measure aspects of, respectively, dynamical decision-making and collection of perceptual evidence. Participants had to indicate whether a target was present or absent from an indoor scene. We manipulated expectations of target location by placing it at usual or unusual locations. Noise was added to the images in order to degrade the scene and generate bottom-up visual uncertainty. To ensure that the noise would not disturb the perception of the scene context, we presented a scene preview of a smaller size (but without the target) before the visual scene itself. This allowed the participants to build a prior representation of the scene that was equivalent across noise levels (Castelhano & Henderson, 2007). We divided the visual search into two epochs defined through eye movements. The detection epoch stretches from the appearance of the scene to the first saccade landing on the target’s area of interest (AOI), and the verification epoch includes the first sequence of fixations within the target AOI. The total duration of the former (detection time or target latency) and later (verification time) can be interpreted as measurements of the time spent, respectively, finding and identifying the target (Castelhano, Pollatsek, & Cave, 2008; Malcolm & Henderson, 2009). It is to be noted that if there is no direct fixation of the target before the end of the detection epoch, the target can eventually be noticed peripherally around the end of this epoch. Despite this, detection time remains an indicator of the time spent finding the target because the next fixation after peripheral detection will be directed at the target most of the time, ending the detection epoch. Our objectives were to address the respective roles of target detection and verification in explaining the influence of both expectations about the target location and scene degradation on the decision-making process.

In a recent study, we examined how both expectation of a specific target and visual degradation of a road scene by fog contribute to the decision process, which leads to target identification by affecting detection time and verification (Quétard et al., 2015). We used mediation analysis (Preacher & Hayes, 2008) to assess whether the effects of expectations or fog on a mediating eye movement variable (detection or verification time) could account for the variations in the mouse-tracking variables. We observed that scene degradation caused by fog led to longer detection and verification times, and these effects delayed the

decision-making process and increased the deviation of the trajectory toward the incorrect response. Expecting the correct target led to shorter verification times, resulting in quicker decisions with a reduced bias toward the inaccurate response. However, we limited our analyses to two broad measurements of mouse responses (AUC and time when the largest mouse deviation was reached) whereas mouse-tracking tasks allow for a richer analysis of mouse responses. Moreover, this previous study presented a two-alternative, forced-choice task (identification of the stimulus within a visual scene). However, finding evidence that points toward one object or category rather than another does not involve the same cognitive processes as accumulating evidence of the *presence* or *absence* of a stimulus as necessary for visual search tasks. In terms of the evidence accumulation process, the former would require checking for distinctive features permitting the identification or categorization whereas the latter would require exploring the visual scene sufficiently until the target is found or enough confidence has been acquired that the target is not present.

At a methodological level, this new study specifically focuses on the effects of scene degradation caused by noise and expectations about the target location (rather than about target identity) on eye movements and mouse responses with a view to assessing active collection of evidence and perceptual decision-making in visual search. We analyzed RTs and AUC using mediation analysis and extended the analysis of the mouse responses to velocity and trajectory deviation across time (Castelhamo & Henderson, 2007). In line with previous studies, we predicted that unusual target locations would increase the target detection time and number and duration of fixations during the detection epoch (Castelhamo & Heaven, 2010, 2011; Malcolm & Henderson, 2010; Pereira & Castelhamo, 2014; Spoto et al., 2014). These studies found mixed results for the influence of target location expectations on verification times. Visually degraded scenes should also increase both detection and verification times as it has been observed that noise increases fixation durations while reducing the number of fixations on the scene and impairs the spatial distribution of fixations on the scene with shorter saccades and a higher central bias (Röhrbein, Goddard, Schneider, James, & Guo, 2015). As the active oculomotor behavior aimed at exploring the visual scene is seen as an evidence-collection process (Friston et al., 2012; Kietzmann & König, 2015; Kietzmann et al., 2011; Quinton et al., 2014), impairment of target detection and verification by both noise and targets in unusual locations should impair the decision-making process as measured through mouse movements. Hence, mouse movements should be slowed and RTs longer, and there should be a bias toward the *absent* response. The basis for this

hypothesis lies in the intrinsic asymmetry between deciding whether a target is *present* or *absent*. This choice can be fit in with the popular evidence-accumulation framework for decision-making (Ratcliff & McKoon, 2008) if it is assumed that accumulating evidence for the *present* response rests on finding a target with a known identity whereas the *absent* response aims to build up confidence that the target is not in the scene. Consequently, there should be heightened competition between *present* and *absent* responses insofar as the *absent* alternative has more weight in the decision-making process over time because the target is harder to find. As target detection and verification permit this competition to be resolved, they should influence this competition. One important assumption in mouse-tracking paradigms is that competition between the two alternative responses, here target *present* or *absent*, unfolds parallel to the mouse movement (Freeman, Dale et al., 2011; Song & Nakayama, 2009). As this competition is updated as new evidence is collected through eye movements, studying the time course of measurements, such as movement velocity and deviation toward the *absent* response with regard to when the target is detected and verified, should highlight the effects of target location and noise. In other words, we should observe these effects even before target detection and during target verification.

Method

Participants

A total of 65 psychology students (including eight males) from Blaise Pascal University took part in the experiment. They received academic credits as compensation for taking part and gave their written consent. Their average age was 20.4 ($SD = 4.1$). All of them had normal or corrected-to-normal vision (glasses or contact lenses).

Stimuli

To build the stimuli set, we selected 60 indoor scene pictures (from the categories bathroom, living room, bedroom, child's bedroom, dining room, office, and kitchen) from the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) and 60 object pictures from the POPORO database (Kovalenko, Chaumon, & Busch, 2012). The scene pictures were scaled at a resolution of 1024×768 pixels. Each target object was associated with a semantically congruent scene and placed either at a usual or an unusual location in the

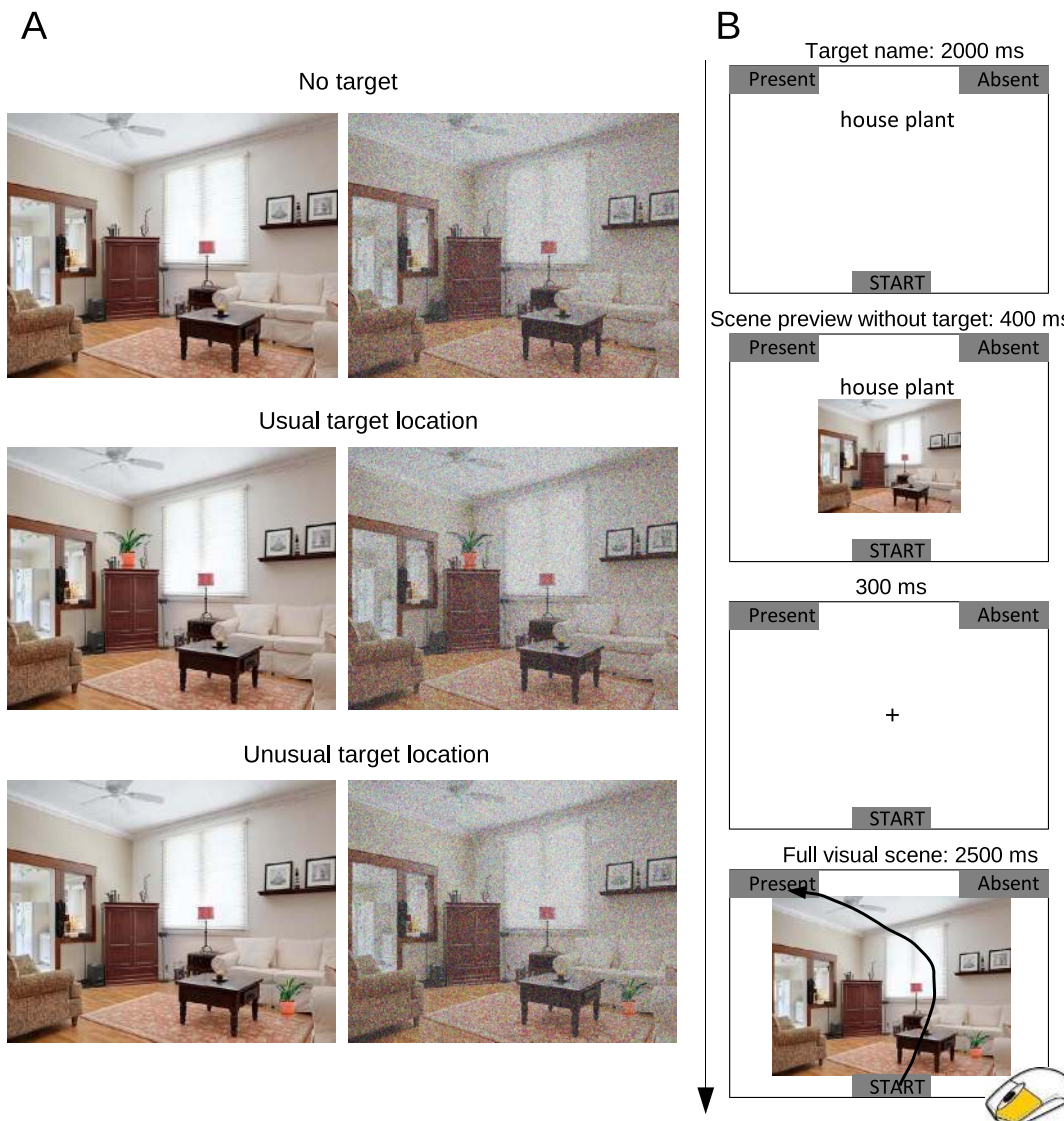


Figure 1. (A) Versions of a scene in each experimental condition. The scene either includes or does not include the target. The target is a house plant, displayed in a usual location and an unusual location. Left: without noise; right: with noise. (B) Time course of a trial.

scene or were absent from the scene (see Figure 1A). To determine the positions of each target, we controlled the laterality of the two target positions; for example, when the target's usual position was in the left part of the scene, its unusual position was in the right part and vice versa. We balanced the laterality of the targets across scenes by choosing usual and unusual locations on opposite sides of each scene. In approximately half of the scenes, the usual location was on the left side whereas the unusual location was on the right side and vice versa for the other half. Finally, we controlled for the central bias (Tatler, 2007) by choosing two approximately equidistant positions away from the center of the picture. For each location, the targets were the same size.

To assess the validity of the target positions, we pretested the scenes. We asked a total of 26 participants

to rate the likelihood of finding the target at this position in the scene (0 = not likely, 100 = very likely). We presented all of the scenes one by one in the same random order for half of the participants and in the opposite order for the other half. This method ensured that, for each scene, when half of the participants were presented with the usual location version first, followed by the unusual location, the other half saw the unusual location version first and the usual location version second. Targets positioned in a usual region ($M = 85.18$) were rated more likely to be found in this position than targets positioned in an unusual region ($M = 10.45$) as assessed using a repeated-measures ANOVA performed on the scenes as observations, $F(1, 59) = 1,314.65$, $p < 0.001$. Finally, we made sure that the rating of usual and unusual target positions differed within scenes by computing the confidence intervals for

each scene. The usual and unusual ratings were different for all the scenes with a confidence level of 99% except for one scene that had a 95% confidence level.

One scene had to be removed because of an error when building it, which we noticed after a quarter of the participants had performed the experiment. We used an algorithm to compute mean and maximal saliency within the target's AOI (Birem, Quinton, Berry, & Mezouar, 2014, adapted to the task specifics and derived from Frintrop, 2006; Itti et al., 1998) in order to check that the bottom-up saliency of the targets did not differ statistically between their two possible locations. There was no difference between usual and unusual locations for either mean, $F(1, 58) = 0.51$, $p = 0.477$, or maximal saliency, $F(1, 58) = 0.49$, $p = 0.485$.

To manipulate the visual uncertainty of the scenes, we degraded the pictures by adding a Gaussian noise to each pixel (independently on each RGB color component). The noise distributions were centered with a variance of 0.3 (for an intensity range of [0, 1]). Our final stimuli set totaled 354 pictures (59 scenes \times 3 target conditions \times 2 noise levels). We added 59 scene previews that did not include the target and were scaled at 600×450 pixels.

Apparatus

We used the MouseTracker software (Freeman & Ambady, 2010) to program and run the experiment. The stimuli were presented using a laptop with a 15.6-in. display set at a resolution of 1600×900 pixels and a refresh rate of 60 Hz. The participants were instructed to maintain their head at an approximately constant distance from the screen (61 cm). An ordinary mouse was used, and the movements were recorded by the MouseTracker software at a sampling rate of approximately 60 Hz.

Eye movements were recorded using a SMI RED-m head-free eye-tracker at a sampling rate of 60 Hz. We preprocessed the data using SMI's dispersion-based algorithm for fixation identification based on Salvucci and Goldberg (2000). We set the minimum fixation duration threshold at 80 ms and the minimum dispersion threshold at 1° (within the optimal range tested by Blignaut, 2009).

Procedure

The experiment began with a five-point eye-tracker calibration procedure and a four-point validation to establish the calibration error. Participants then completed a 16-trial training block, including natural scenes

and target animals rather than indoor scenes and target objects. Figure 1B shows the time course of a trial. To begin each trial, participants had to click with the mouse cursor on the *start* button at the bottom center of the screen with the mouse's cursor. The name of the target object appeared for a total of 2400 ms. After 2000 ms, a preview of one of the scenes appeared in the center of the screen below the target name for the remaining duration (400 ms). Then, after a blank screen with a 300-ms fixation cross, the same scene was presented for 2500 ms in one of the six conditions previously listed: degraded with noise or not, with or without a target presented in a usual or unusual location. The mouse cursor was locked in the middle of the *start* button before the scene appeared to ensure all the trajectories started from the same location. Participants were instructed to click on the *present* or *absent* response buttons in the top left and top right corners of the screen, according to whether the target was in the scene or not. If 1000 ms elapsed before they moved the mouse, a message was displayed after completion of the trial, urging them to begin their movement earlier in the upcoming trials. A response given after 2500 ms was invalid resulting in a "time out" message.

Each participant was presented with 59 trials, including 10 trials per experimental condition (3 target conditions \times 2 noise levels, nine trials for one condition due to the later removal of one scene) with only one version of each scene being presented. For instance, a participant being presented a visual scene depicting a specific kitchen with no target and no noise could not be presented the same kitchen in another experimental condition. Thus, the different versions of each scene were counterbalanced across participants using Latin squares, resulting in the formation of six sets of trials each presented to a subgroup of participants. The laterality of the response buttons was also counterbalanced with approximately half of the participants in each subgroup being presented with a leftward *present* button and a rightward *absent* button, and the other half were presented with the opposite button positions.

Results

Data processing

Eight participants were excluded from the analysis for the following reasons: invalid eye-tracking data due to glasses or lenses making the calibration impossible (five) and a high number of "time out" trials combined with weak calibration (three). We analyzed only the trials in which the target was present because our eye-

movement variables could not be calculated in target-absent trials. We excluded the trials with invalid eye-tracking data: In 4.1% of the trials, gaze was directed toward the target when the scene appeared, in 15.5% there was no fixation on the target, and in 1.4% the detection or verification time was less than 80 ms or more than 2500 ms. We then discarded trials in which the participant's response exceeded the time limit (5.6% of the trials with valid eye-tracking data) or in which he or she clicked on the wrong response (8.4%). Overall, the analyses were performed on 68% of the trials, including a target (1,491 trials). The low percentage of valid trials is not surprising because it was not possible to instruct the participants to fixate the target directly when responding because they were involved in a dynamic motor task in which they had to control their mouse movement.

We remapped the correct response buttons on the right side and the alternative on the left side of the screen, and we rescaled the x -axis between -1 and 1 and the y -axis between 0 and 1.5 in accordance with traditional mouse-tracking literature (Freeman & Ambady, 2010). Given the mouse-tracker time resolution, the time interval between each point on a trajectory was 16 ms. As eye-tracking measurements, we calculated the detection and verification times as well as the mean duration of fixations and number of fixations during the detection epoch. As temporal measurements of the dynamics of mouse responses, we calculated the RTs and velocity profiles of the trajectories. As measurements of the spatial deviation of the mouse toward the *absent* response, we calculated the AUC and mouse deviation profiles along the x -axis.

To assess the effects of the target location (usual vs. unusual location) and noise (no noise vs. noise) on the eye-tracking and mouse-tracking measurements, we used linear mixed modeling (LMM), which can take account of the variability of the parameters across participants and visual scenes simultaneously without having to perform separate *by-subject* and *by-item* analyses (Baayen, Davidson, & Bates, 2008). In our study, *by-item* analysis corresponds to *by-scene* analysis. This is interesting in our experiment because we used Latin squares to distribute the trials of each condition across both scenes and participants. Furthermore, these models are particularly well suited to our data because they enable efficient handling of repeated-measures data sets with imbalanced data and missing values. When required, we included the target location and noise as well as their interaction as fixed factors. We recoded the location variable as -0.5 for usual and $+0.5$ for unusual probability locations and the noise variable as -0.5 without noise and $+0.5$ with noise. As random parameters, we included the variance of the intercepts and slopes as well as the correlations between random parameters for both the subject and scene variables.

However, due to nonconvergence of the model testing the RTs, we took out the slope variances for the subject variable after ensuring they were nonsignificant. In that way, we used the maximum random structure allowing the models to converge (Barr, Levy, Scheepers, & Tily, 2013). To estimate the p values of the fixed effect parameters, we used the Kenward-Rogers approximation for the degrees of freedom. The “lme4” R package was used to compute the LMM and the “pbkrtest” R package to estimate the p values.

Eye movement measurements

Eye movements during the detection and verification epoch were analyzed. The detection epoch included all the fixations on the scene (including the latency of the first saccade) before the first saccade on the target AOI. Three measurements were analyzed for this epoch, namely total detection time, the number of locations fixated, and the mean duration of the fixations. The verification epoch spanned from the first fixation on the target's AOI to the first saccade ending outside the target AOI. Following the example of Malcolm and Henderson (2009, 2010), if the fixation following the first fixation on the target was not on the target but was followed by a fixation on the target, it was considered part of the verification epoch because this fixation could be part of a preprogrammed sequence of saccades. Verification time was computed as the sum of the durations of the fixations within the target AOI and included in the verification epoch.

Detection epoch

See Table 1 for mean detection times, number of fixations, and fixation durations by experimental condition. Targets positioned in an unusual location resulted in significantly longer detection times, $b = 204.59$, $t(62.8) = 8.44$, $p < 0.001$, and more fixations during this epoch, $b = 0.76$, $t(51.9) = 8.86$, $p < 0.001$, than when the target was in a usual location. No main effect of noise was found on detection times, $b = 13.03$, $t(27.5) = 1.12$, $p = 0.27$, or on the number of fixations, $b = -0.06$, $t(31.9) = -0.52$. There was no interaction between the two variables on detection times, $b = -20.77$, $t(1308) = -0.93$, but a marginal interaction was recorded for the number of fixations, $b = -0.22$, $t(1342.2) = -1.92$, $p = 0.056$. Therefore, positioning a target in an unusual location increased the number of fixations slightly less in a scene degraded with noise than in a scene with no noise, but the difference did not reach the significance threshold. Finally, mean fixation durations contrasted with these results: The only effect found with this measure was that fixation durations increased for a noisy scene, $b = 10.60$, $t(36.3) = 2.70$, $p =$

	Usual location		Unusual location	
	Without noise	With noise	Without noise	With noise
Detection time, ms	388.9 (9.9)	406.9 (10.8)	570.4 (15.2)	557.8 (15.4)
Fixations number	1.74 (0.04)	1.78 (0.05)	2.51 (0.07)	2.31 (0.06)
Fixations duration, ms	199.6 (3.1)	210.0 (3.7)	200.8 (3.4)	211.6 (3.7)
Verification time, ms	336.9 (6.7)	387.2 (9.4)	317.3 (6.8)	374.5 (9.6)

Table 1. Means (standard errors) for each eye movement measurement by target location and noise conditions.

0.011, whereas there was no main effect of target location, $b = 1.45$, $t(42.8) = 0.58$, and no interaction, $b = -0.82$, $t(1321.4) = 0.06$.

Verification epoch

Verification time was the only measurement used to characterize the verification epoch (see Table 1 for means). The only effect found was an effect of noise, $b = 63.84$, $t(34.2) = 7.87$, $p < 0.001$, with longer verification times when the scene was degraded by noise. No effect of target location, $b = -12.6$, $t(49.1) = -1.10$, $p = 0.277$, and no interaction between target location and noise, $b = 7.73$, $t(1320.3) = 0.56$, were found.

RTs and temporal dynamics of the mouse response

In order to analyze the dynamics of the responses, we first used mediation and moderation analyses to investigate how target location and noise affected the RTs and how the detection and verification times influenced location and noise effects on the RTs. Second, we further analyzed the dynamics of the responses, focusing on the time course of the velocity profiles of the mouse trajectories in relation to target detection and target verification.

Response times

The RT spanned from the appearance of the scene to the response click. By making the target harder to find, both scene degradation and unusual target locations should result in longer RTs. We used mediation and moderation analyses to assess whether the detection and verification times affected the effects of target location and noise on RT. The detection or verification times moderated the effects of the independent variable (IV), i.e., target location or noise, on the RT when there was an interaction between the IV and the moderator. Thus, depending on the moderator values, the effect of the IV on the RT will vary. On the other hand, with mediation analyses, it is possible to estimate whether

the effect of the IV on the mediator explained part (or all) of the effect of the IV on the RT. As there was an effect of target location on detection time and an effect of noise on verification time, the effect of target location on RT should be mediated or moderated by detection time whereas the effect of noise on RT should be mediated or moderated by verification time.

We assessed the mediative effects of detection and verification times using a bootstrapping procedure (Preacher & Hayes, 2008). It consists of estimating the indirect effect of the IV on the RT through the mediator. This indirect effect AB is calculated as the product of the IV's effect on the mediator (path A) multiplied by the mediator's effect on the RT while controlling for the effect of the IV (path B). The 95% confidence interval of this effect was assessed with a nonparametric bootstrapping procedure comprising 10,000 samples. To measure the size effect, we estimated the mediated proportion of the IV's total effect (path C). This can be done by calculating the proportion by which the IV's effect is reduced when controlling for the effect of the mediator (path C'). Although this measurement is not perfect, it is a simple way of obtaining information about the effect size (for a short discussion, see Hayes, 2009). We estimated these parameters using LMMs with both IVs and the mediator (when needed) as well as all the possible interactions as fixed effects. We specified the same random structure for all models, i.e., the variance of the intercept and slopes of the IVs and the correlations between random parameters nested within the subject and scene random factors.

Figure 2A displays the mean RT for each condition, and Figure 2B displays in schematic form how target detection and verification times mediate the effects of target location and noise on RT. Our LMMs revealed a main effect of both noise, ($b = 71.27$, $t(50.9) = 4.7$, $p < 0.001$, and target location, $b = 148.69$, $t(55.5) = 6.07$, $p < 0.001$, on RT. No interaction was found, $b = -28.58$, $t(1361.1) = -0.99$. Responses were faster when the scene was not degraded by noise as opposed to degraded and when the target was placed at a usual location as opposed to an unusual one. Detection times mediated the effect of target location on RT. When we controlled for the effect of detection time, the sign of the parameter of target location expectation (direct effect

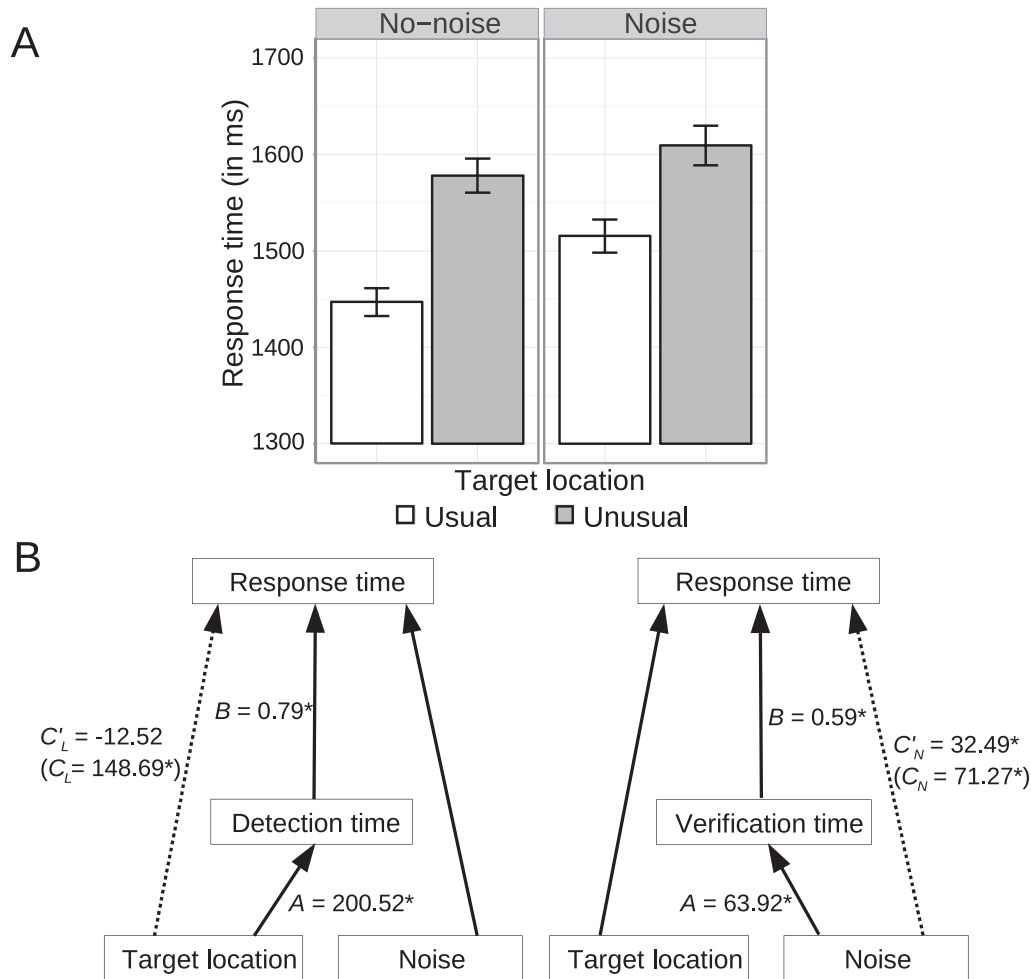


Figure 2. (A) Mean RTs as a function of target location and scene degradation. White bars: target in usual location; gray bars: target in unusual location. Error bars denote the standard errors. (B) Mediation of the effects of target location and noise by the eye movement measurements, i.e., target detection and verification times. Dashed arrows denote an effect mediated by detection or verification times. The model parameters are estimated with LMMs (see Results). When an effect is mediated, parameters A and B are displayed (indirect effect = $A \times B$) as well as the direct effects C' (effect of the independent variable when we controlled for the effects of the mediator) and total effect C (effect of the independent variable without including the mediator in the model); L indices: effect of target location; N indices: effect of the noise. *The parameter is significant ($p < 0.05$).

C') was reversed and lost significance, $b = -12.52$, $t(61.5) = -0.75$, and the indirect effect of target location expectation through detection time was significantly different from 0 ($AB = 159.40$, 95% bootstrap CI [135.15, 185.62]). When detection times are included in the model as a predictor, the target location effect disappears. Conversely, when we controlled for the effect of verification time, the direct effect of noise on RT was reduced by 54.4% and remained significant, $b = 32.49$, $t(55.6) = 2.19$, $p = 0.032$. The indirect effects of noise through verification time was significant ($AB = 38.14$, 95% CI [27.76, 52.17]). Finally, the indirect effect of target location through verification times and of noise through detection times were not significant (respectively, $AB = -7.60$, 95% CI [-17.52, 1.21] and

$AB = 10.66$, 95% CI [-9.77, 28.72]). Thus detection time mediated the effect of target location on RT whereas verification time mediated a large part of the effect of noise.

Velocity profiles

We further analyzed the dynamics of the competition between responses by means of mouse velocity as a function of time. Velocity profiles can estimate the beginning of the response movements as well as how the commitment toward the correct response builds up as time unfolds (Hehman, Stolier, & Freeman, 2015). Our underlying hypotheses are that greater difficulty finding and identifying the target due to unusual

location and/or noise should reduce the commitment toward the *present* response, hence reducing overall velocity. We analyzed the time course of velocity in relation to the moments when the target was detected (end of the detection epoch, corresponding to the detection time) and the end of target verification (end of the verification epoch), thus examining how it evolves during the detection and verification epochs. Finally, because accumulating evidence during the detection and verification epochs should affect the ongoing decision, mouse movements should begin during the detection epoch (and, thus, before the target is properly fixated). To perform this analysis, for each trajectory, we centered the time coordinates on the end of the detection epoch by subtracting the detection time from each time coordinate. Thus, the 0-ms time indicates approximately the moment along the response trajectory when the target was detected, referenced later on as target detection. The interval between each time stamp was 16 ms (~ 60 Hz time resolution). We measured the velocity for each time step of each trajectory. Then, we smoothed the velocity profiles of each trajectory using a convolution with a triangular kernel over a sliding window of 21 time steps. This smoothing step is required on account of the stochasticity of mouse events on most computer devices. This method allowed us to analyze the velocity prior and posterior to the moment when the target was detected, and it also allows for raw time analysis on a data set in which mouse movement onsets vary within a large time interval (2500 ms). More commonly used alternatives include time-normalizing the trajectories (Hehman et al., 2015) at the cost of losing part of the time information whereas our key objective was to study the raw timing of the responses compared to the target detection and target verification.

Figure 3 displays the mouse velocity as a function of time centered on detection of the target for each experimental condition. For each time step, the velocity was estimated by computing one intercept-only LMM (testing velocity against 0) *per* experimental condition. These LMMs included the intercept variations across both participants and scenes as random parameters. As inferential statistics, we computed the 95% Wald confidence intervals from the LMMs (shown in Figure 3). For targets in a usual location, velocity peaks occurred, respectively, for no noise and noise conditions, 528 ms and 560 ms after target detection and 191 ms and 173 ms after the end of target verification (estimated velocity peak minus mean verification time, shown as solid and dashed vertical lines in Figure 3A, B). For targets in unusual locations, they occurred 480 ms after target detection (for both no noise and noise) and, respectively, 163 ms and 106 ms after the end of target verification. The velocity peak of the profiles was higher when the scenes were not degraded by noise than

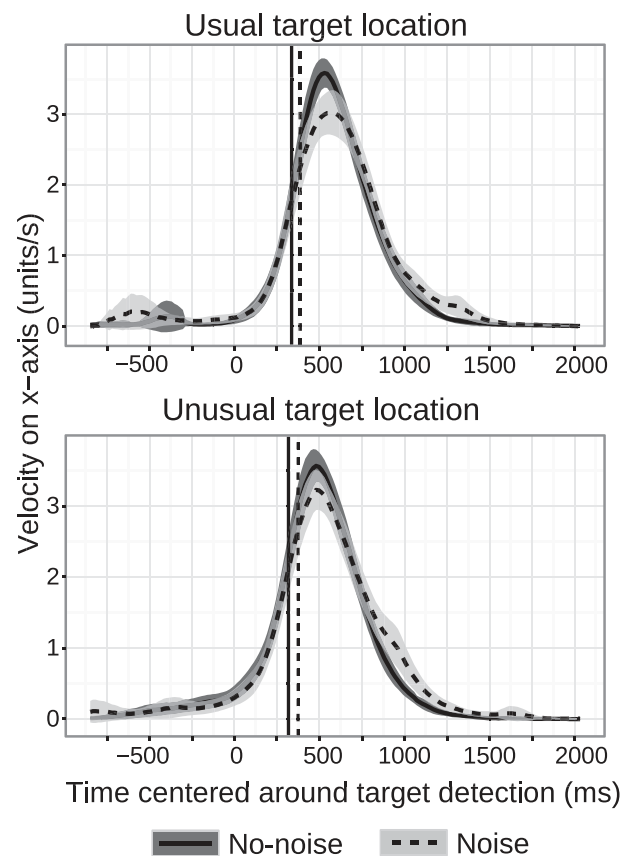


Figure 3. Mouse velocity as a function of time for scenes nondegraded by noise (solid lines, dark gray) and scenes degraded by noise (dashed lines, light gray). Top: target in a usual location; bottom: target in an unusual location. The x-axis represents time centered on detection of the target; 0 represents the moment the target was first fixated. Velocity and the error intervals (95% Wald confidence intervals) were estimated at each time frame using LMMs. On each plot, the black vertical lines denote the average moments when target verification ended (verification time). No noise: solid lines, noise: dashed lines.

when they were degraded irrespective of whether the target was in a usual (no noise: $b = 3.58$, 95% CI [3.38, 3.79]; noise: $b = 3.03$, 95% CI [2.71, 3.34]) or an unusual location (no noise: $b = 3.56$, 95% CI [3.34, 3.79]; noise: $b = 3.22$, 95% CI [2.94, 3.51]). The velocity was higher for nondegraded scenes from 416 ms to 592 ms (12 steps) after target detection for targets in a usual location and from 448 ms to 560 ms (eight steps) for targets in an unusual location. However, there was no effect of target location on the velocity around the velocity peak. A result of the reduced velocity for degraded scenes is that the end of the mouse movement exhibited a significantly higher velocity than for nondegraded scenes from 1104 ms to 1440 ms (22 steps) for targets in usual locations and from 816 ms to 1552 ms (47 steps) for targets in unusual locations. Finally, we

found evidence that mouse velocity began to increase even before target detection (0 on the x -axis of Figure 3). Overall, for targets in usual locations, velocity was significantly greater than 0 up to 272 ms before target detection and increasing from 0.056 units/s to 0.106 units/s once the target was detected. For targets in unusual locations, the velocity was significantly above 0 up to 640 ms before target detection and increased from 0.069 units/s to 0.338 units/s when the target was detected. Furthermore, when the scene was not degraded, the velocity for targets in unusual locations was significantly higher than for targets in usual locations from 288 ms before target detection to 416 ms after target detection (45 steps). When the scene was degraded, the velocity for targets in unusual locations was significantly higher than for targets in usual locations from 128 ms before target detection to 496 ms after target detection (40 steps). This indicates that, in conditions in which target detection is harder due to unusual target locations, thus leading to longer detection times, the response movement was initiated even before the first fixation on the target.

Mouse trajectory spatial deviation

Measuring the deviation of the mouse movements from the ideal response trajectory is a way of estimating the attraction toward the incorrect *absent* response in the decision-making process. As an overall deviation measurement, we calculated the AUC. In a similar way as we did for RT, we used mediation and moderation analysis to estimate the contribution of target detection and verification to the effect of target location expectations and noise on the AUC. Finally, we studied spatial deviation over time centered on target detection by analyzing the deviation on each non-normalized time step, using a similar method to the one used for the velocity profiles.

Area under the curve

The AUCs were calculated after time-normalizing the trajectories into 101 time steps (as usually done in mouse-tracking experiments; Freeman & Ambady, 2010). Making the target harder to find by means of both scene degradation and unusual target locations should increase the AUC. This implies that the difficulty finding a target should increase the competition between *present* and *absent* responses, and because target detection and verification should play parts in resolving this competition, they should mediate or moderate these effects. As with the RT analyses, the effects of target location should be mediated/moderated by detection times as only target location had an effect on detection times, and respectively, the effects of

noise should be mediated/moderated by verification times. Figure 4A displays the mean AUC for each condition and the averaged mouse trajectories for each target location condition. Only the main effect of target location on AUC was significant, $b = 0.22$, $t(49.8) = 3.35$, $p = 0.002$. Deviation from the ideal response trajectory was greater when the target was in an unusual location. No noise effect nor interaction was found, respectively, $b = 0.07$, $t(29.6) = 1.59$, $p = 0.12$ and $b = -0.04$, $t(1339.3) = -0.50$. Mediation and moderation analyses revealed that verification time had an effect on AUC, $b = 0.0005$, $t(1151.5) = 3.05$, $p = 0.002$, but did not moderate or mediate the effect of target location. On the other hand, when detection times were included in the model, the effect of target location was no longer significant, $b = -0.046$, $t(39.3) = -0.98$, $p = 0.333$. Detection time interacted significantly with target location, $b = 0.0007$, $t(907.7) = 4.13$, $p < 0.001$. Figure 4B displays this moderation model. Detection times increased the deviation toward the inaccurate response to a greater extent when the target was in an unusual location as opposed to a usual location.

Deviation on x -axis over time

To investigate the relationship between the trajectory deviation and the detection times further, we studied the deviation on the x -axis (x -deviation) prior and posterior to the moment the target was detected. Given that only target location affected the AUC, we focused on the x -deviation only on this variable. As the alternative responses are spatially distinguished on the x -axis, an analysis of x -deviations can index the ongoing competition across time. For example, greater deviations can indicate that the incorrect response is considered a valid alternative. As with velocity, we analyzed the time course of x -deviation in relation to when the target was detected (detection time) and the end of target verification (end of the verification epoch). We hypothesized that an unusual location should increase the deviation. As accumulating evidence toward an *absent* response is aimed at accumulating confidence that the target is not in the scene, the x -deviation should accordingly begin to increase during the detection epoch as it is during this epoch that the *target-absent* response competes with the *target-present* response. As we did for velocity profiles, for each trajectory, we centered the time coordinates on the detection time so that a 0-ms time coordinate indicates the moment on the response trajectory when the target was detected. The x -deviation was the difference between the coordinate on the x -axis of each point of the response trajectory and the x -coordinate of its perpendicular projection point on the ideal response trajectory (straight line from *start* to *present* button). A positive x -deviation value denotes a deviation toward

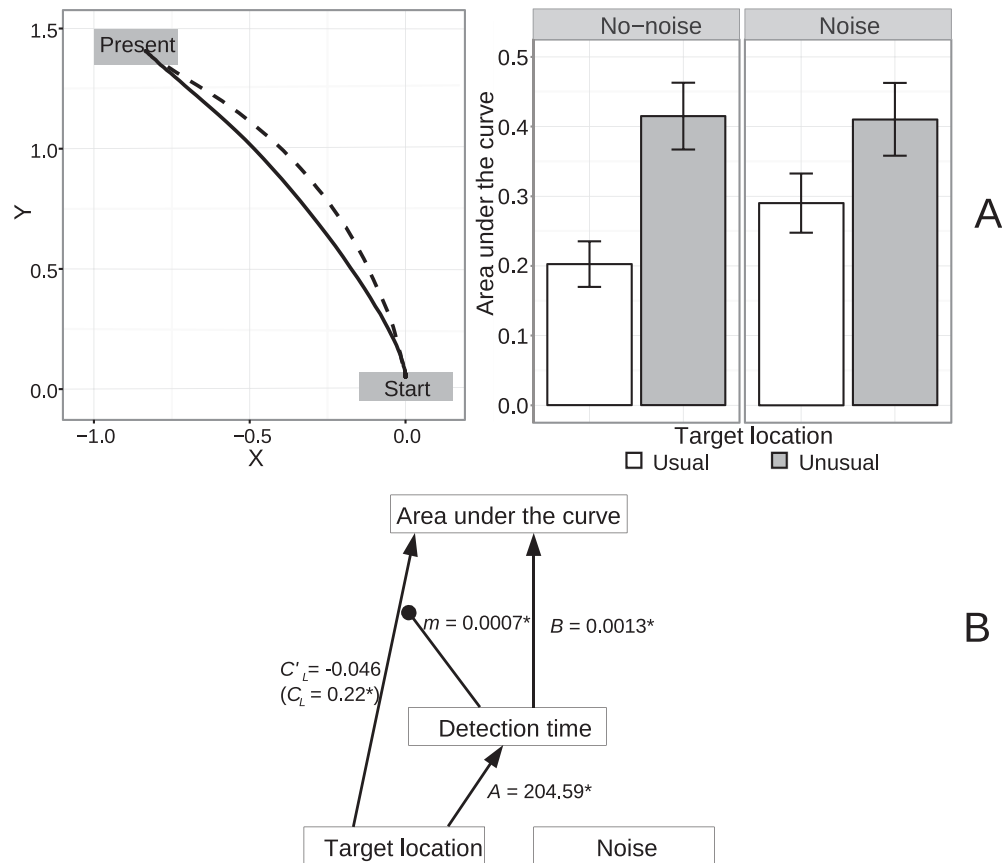


Figure 4. (A) Left: mean mouse trajectories as a function of usual (solid curve) and unusual (dashed curve) target location. Right: AUCs as a function of target location and scene degradation. White bars: target in usual location; gray bars: target in unusual location; error bars denote the standard errors. (B) Moderation of the effects of the target location and noise by the target detection time. The line ending with a black dot indicates a moderation effect with m being its parameter (moderation corresponds to an interaction between target location and detection time). Parameters A and B are displayed together with the direct effect C' (effect of the independent variable when the moderator is included in the model) and total effect C (effect of the independent variable without including the moderator in the model). L indices: effect of target location. *The parameter is significant ($p < 0.05$).

the *absent* response compared to the ideal trajectory. We smoothed the x -deviation profiles of each trajectory by using a convolution with a triangular kernel over a sliding window of 21 time steps. Figure 5A displays the x -deviation as a function of time centered on detection of the target for usual (solid line, dark gray error) and unusual target locations (dashed line, light gray error). The 95% Wald confidence intervals included in Figure 5A were estimated with intercept-only LMMs for each time step. As stated previously, these LMMs included the intercept's variation across both participants and scenes as random parameters. Two observations are relevant here. First, as with the results with the AUC, the overall deviation was greater for targets in an unusual location than in a usual location. This was the case from 304 ms before target detection to 640 ms after target detection (60 steps). For targets in an unusual location, the x -deviation attained its maximum value ($b = 0.115$, 95% CI [0.087, 0.143]) around 384 ms after

target detection, corresponding to the average end of target verification (343 ms, dashed vertical line). For targets in a usual location, the maximum x -deviation ($b = 0.063$, 95% CI [0.045, 0.08]) occurred 560 ms after the target detection, which corresponds to an approximate 200-ms delay after the end of target verification (360 ms, solid vertical line). This result may seem counter-intuitive, but given that the target was detected sooner when in a usual location and thus verified sooner, participants had much more time to finish their movement and validate their response before the time granted to respond elapsed. Second, we observed a significant deviation toward the *absent* response before target detection. The trajectory deviated up to 128 ms before target detection for targets in a usual location and 512 ms for targets in an unusual location. As reported previously, the deviation was significantly greater for targets in unusual locations starting 304 ms before target detection (Figure 5A).

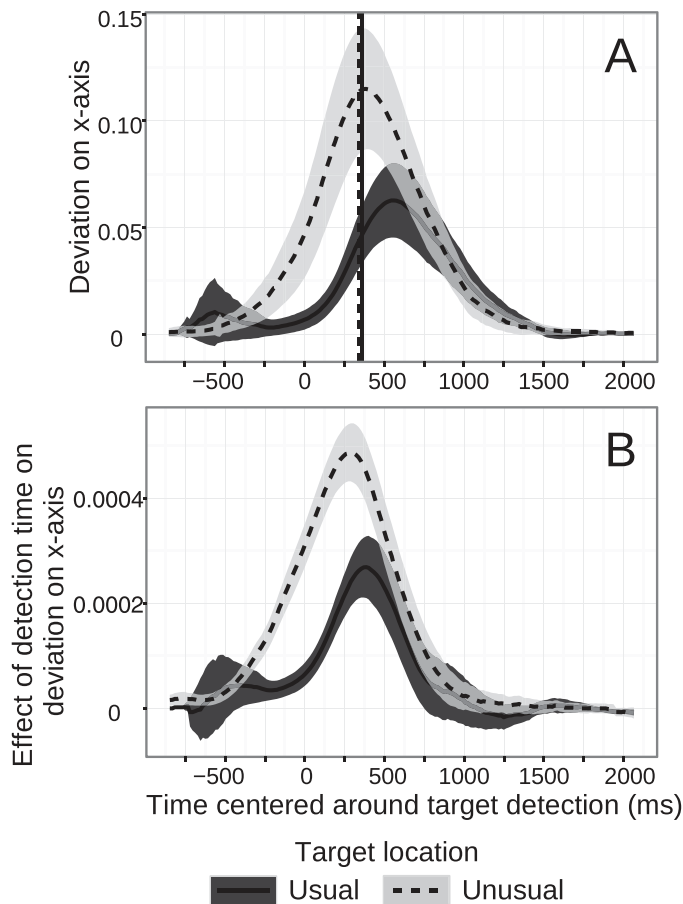


Figure 5. (A) Estimated deviation on the x -axis as a function of time and target location. (B) Estimated effects of target detection times on the x -deviation as a function of time and target location. Solid lines, dark gray: target in a usual location. Dashed lines, light gray: target in an unusual location. The x -axis represents the time centered on the detection of the target; 0 represents the moment the target was first fixated. X -deviations and the effect of detection at each step were estimated using LMMs, as well as the error intervals (95% Wald confidence intervals). The black vertical lines denote the average moments when target verification ended (verification time). Usual target location: solid, unusual target location: dashed.

With mediation and moderation analysis, we previously found an interaction between detection time and target location expectations on the AUC in which the effect of detection time was greater for targets in unusual locations than in usual locations. We examined the time course of the effect of the detection times on the x -deviation across these conditions more closely by plotting the parameters of this effect across time steps for usual and unusual target location (estimated using LMM with the same random parameters as in the last analysis). This type of analysis is similar to the estimation of correlations across time suggested by

Hehman et al. (2015), but we aimed to obtain less biased estimates of the relationship between detection times and deviation by taking account of the nested structure of our data (estimating the variance of x -deviation across participants and scenes). This is displayed in Figure 5B. Overall, detection times had a greater influence on the x -deviations when the target was in an unusual location (max effect: $b = 0.00049$, 95% CI [0.00043, 0.00054]) as opposed to a usual location (max effect: $b = 0.00027$, 95% CI [0.00021, 0.00033]). This is consistent with the interaction previously reported between detection times and target location expectations on the AUC. We also observed an effect of detection times on the x -deviation even before the end of the detection epoch, especially for targets in an unusual location. Hence, in the trials in which the target was detected later, the deviation toward the *absent* response was greater even in the time steps preceding the effective detection.

Discussion

We devised a novel study of visual search and perceptual decision-making, in which participants were required to report on whether a target object was present or absent from a visual scene and in which both the predictability of target location and noise levels were manipulated. To sum up, the main research questions underlying this study addressed (a) how expectations about target location and visual uncertainty due to noise drive the active visual evidence collection process through oculomotor behavior and (b) how these two factors differentially affected the decision about whether the target is present in the scene or absent. Our results (summarized in Box 1) show that targets in usual locations and nondegraded visual information facilitated, respectively, target detection and verification, and consequently, they highlight the differential roles of these two processes in the encompassing decision-making process.

Hence, both eye-movement and mouse-tracking measurements indicate that predictable target locations and absence of noise facilitate visual search. Increasing noise level in the scene impaired target verification and led to longer fixations during the detection epoch. By contrast, placing the target in an unusual location impaired its detection (compared to a usual location) and increased the number of fixations required to detect it but had no effect on target verification. The decision-making process was also impaired by noise and unusual target location. Both manipulations led to lower performance on the overall RTs. Scene degradation by noise reduced overall mouse velocity and increased it in the latest parts of the response

movement whereas targets in unusual locations increased the bias toward the incorrect *absent* response. Finally, our results indicate that detection and verification processes played a role in the decision-making process, respectively, by mediating the effects of target location and scene degradation on RT. Additionally, detection times moderated the effect of target location on response selection such that unusual target locations impaired the AUC to a greater extent for high target detection times. Both velocity and deviation were significantly different from zero for several time steps before the first fixation on the target, especially when it was in an unusual location.

Contribution of contextual knowledge and bottom-up visual information with respect to eye movement

The results for eye movement measurements are consistent with previous studies that investigated contextual guidance. Visual noise did not impair target detection but impaired the verification process as well as the duration of the fixations during the detection epoch in keeping with previous studies using degraded or low-contrast visual scenes (Harley, Dillon, & Loftus, 2004; Röhrbein et al., 2015). In visual search, variations in the verification time may reflect the identification of the target whereas variations in the duration of the fixation in the detection epoch may reflect that more time is required to reject the fixated location as a potential target location. Thus, in our experiment, degraded visual information made it harder to both reject the fixated location and identify the target.

Considering the effect of usual target locations, previous studies (Castelhano & Heaven, 2010, 2011; Malcolm & Henderson, 2010; Pereira & Castelhano, 2014; Spotorno et al., 2014) consistently showed a reduction in overall detection time and the number of fixations during the detection epoch, reflecting the number of locations visited before finding the target. Along with these studies, our results confirm that expectations about target location improve target detection (as is normally the case for targets in usual locations). As extensively discussed in the literature (Brockmole & Le-Hoa Vo, 2010; Ehinger et al., 2009; Torralba et al., 2006), this may be due to contextual knowledge guiding oculomotor behavior in visual search. However, contrasting results have been reported for the verification epoch. Some showed that contextual information facilitated the verification process (Castelhano & Heaven, 2011; Malcolm & Henderson, 2010) whereas others did not (Castelhano & Heaven, 2010; Pereira & Castelhano, 2014; Spotorno et al., 2014) as in the present study. It is to be noted that there were two main differences with these previous

studies due to our mouse-tracking paradigm. First, participants were subjected to considerable time pressure, and second, subsequent eye movements were needed on the response buttons for mouse movement control after the verification epoch. Thus, our study goes against the idea that identifying a target in an unusual location requires longer gaze allocation than a target in a usual location, especially in situations in which time is limited and subsequent eye movements are needed to perform a motor task efficiently. However, spending more time fixating the target in order to identify it or spending more time fixating a location in a scene before rejecting it as a potential target location could still be a valid strategy for reducing uncertainty about the decision in contexts in which the speed–accuracy trade-off is not biased toward quick responses (unlike in our task in which participants had to respond within a specified time limit and begin their response movement early).

Distinct contributions of detection and verification processes upon decision-making

Our results clearly distinguish between the contributions of target detection and identification to the decision-making process, respectively, influenced by target location and noise. Variations of RT due to target location were explained by the effect of target location on target detection time. Note however that we cannot conclude that detection times mediate the entirety of the effect because this claim relies on a null hypothesis (Rucker, Preacher, Tormala, & Petty, 2011), the effect of target location being nonsignificant when controlling for detection times. Conversely, a large proportion of RT variation (54.4%) due to degradation of the visual scene was explained by the effect of noise on target verification time. Thus, these results show that wrong expectations about the target's location and visual uncertainty due to noise both impair decision-making by requiring prolonged collection of evidence at different levels (target detection and verification), thus slowing the decision process. The results of the present study and the fact that we found no interaction between target location and scene degradation with our mouse-movement measurements could indicate that the influence of expectations and visual uncertainty on perceptual decision-making are processed additively and separately by influencing two distinct subprocesses: target detection and verification. In this respect, an analysis of the dynamics of the response movements indicated that the movements began toward the end of the detection epoch when the target had yet to be fixated. As the aim of the task is to find the target, this epoch is critical. Indeed, contrary to classic mouse-tracking experiments using two-alternative forced

choice, the target on which the decision has to be made is not directly and centrally available right from the beginning of the trial but has to be found first. On the other hand, the moment velocity was highest was situated around the end of the verification epoch and shortly after, thus coincidental with fixations made toward the response buttons after visual exploration of the scene. At this stage, degradation of the visual scene by noise was critical. Although we found no strong evidence of noise delaying the velocity peak, the maximum amplitude of the velocity profile was lower whereas velocity was higher in the latest parts of the response movement. This could indicate less commitment to the *present* response with the degraded target being harder to recognize and distinguish from the background, thus generating uncertainty. This would induce a more careful response with lower maximum velocity but faster toward the end to enable the response to be validated within the time constraints.

Relationship between expectations for target location and target detection upon response selection

Contrary to noise, unusual target locations did not reduce mouse velocity compared to usual target locations. Instead, they induced a spatial bias toward the *absent* response throughout a large portion of the response movement, starting before detection of the target (from approximately 304 ms before target detection to approximately 640 ms after target detection, and mean response times lasted around 1450 ms to 1600 ms). In addition, detection times moderated the effect of target location on the AUC. Increasing detection times inflated the mouse trajectory deviation toward the *absent* response to a larger extent for unusual target locations than for usual target locations. Consistent results were found for x -deviations across time as the effect of detection times was greater for unusual target locations during most of the response.

This indicates, first, that expectations about the target location influence the final decision about the presence of the target by affecting eye movements made to detect it and, second, that the inflation of the target location effect by detection time could be explained by a combination of two effects: the delay of detection as a result of unusual target location (wrong expectations) and a mechanism whereby certainty about the *target-absent* response builds up as time is spent unsuccessfully searching for the target through eye movements. In our results, we found some evidence of such a mechanism. We found that both mouse velocity and x -deviations began to increase before the end of the detection epoch and that the effect was larger and occurred sooner when targets were in unusual locations

(Figure 5A). This indicates growing competition between the *absent* and *present* alternatives beginning before the moment the target was detected. Furthermore, trials with longer detection times showed bigger x -deviations even before target detection, especially for targets in unusual locations (Figure 5B). Another way to express it would be to say that detection times correlated with increased deviation before the actual detection of the target. This may indicate that the competition increases with the difficulty finding the target but also while the target is actively searched with eye movements. The longer the visual search is unsuccessful, the more the *target-absent* alternative competes with the accurate *target-present* response and the attraction toward the *target-absent* alternative may reflect an indirect accumulation of evidence stacked in favor of this response owing to the lack of evidence in favor of the *present* response. The effect of detection time on x -deviation reached its maximum around the end of the verification epoch, indicating that the time spent fixating the target in order to verify its identity is needed to counteract the evidence that has accumulated toward the *target-absent* alternative and finally converge toward the accurate *target-present* response. The gradual increase in mouse velocity up to target verification consistently supports this interpretation. By extension, with this model, if the target is not found, the gradual commitment toward the *absent* response would finally lead to the decision to stop searching for the target. This could be plausible as a mechanism allowing for search termination (for review, see Wolfe, 2012) as a possible way an agent searching the environment for an object in an area decides to stop and to search in another area.

However, we need to remain careful when considering these conclusions given that, in our experiment, the examination of how target detection and verification relate to mouse movements is based on the incidental measurement of detection and verification times as invoked variables, and thus, only limited conclusions can be drawn about causation. Furthermore, the uncertainty about the visual input was directly manipulated with noise and top-down expectations using usual and unusual target locations, and we found that their effects on eye movements and mouse response movements could be linked. Our study tested how noise and target location influenced eye movements (active collection of evidence), in turn influencing decisions through mouse movements in a feed-forward fashion. What our study does not cover is how the accumulation of evidence toward each alternative itself influences how eye movements are directed to collect further information (see Kietzmann et al., 2011; Kietzmann & König, 2015). Nonetheless, it provides an interesting direction for further studies addressing both the causal relationship between active

vision and decision-making and the search termination problem in visual search tasks. For instance, an important step would be to manipulate detection times and verification times directly in similar mouse-tracking tasks, for instance, with designs inspired by the scene onset delay paradigm (Henderson & Smith, 2009; Nuthmann, Smith, Engbert, & Henderson, 2010). Comparing human data with simulated data using computational models implementing target detection and verification processes could also provide interesting insights. Additionally, because expectations about target location are linked to detection times, it would be important to control for their effects, for instance, by using nonmeaningful targets in real-world scenes, such as white dots as used by Brockmole and Henderson (2006) and Brockmole and Le-Hoa Vo (2010), or by using artificial grids instead of natural scenes.

Conclusion

On a theoretical level, the results of our study indicate that the two distinct visual subprocesses, i.e., target detection and target verification, may play different roles in decision-making, notably contributing in the influence of noise and expected target location on the decision process. Noise degrading bottom-up visual information slowed down decision-making by increasing target verification, consequently moderating the commitment to the decision (as reflected in the speed of the mouse movement). Target location influenced decision-making by affecting eye movements relating to target detection, consequently moderating the bias toward the *absent* response. Additionally, we suggested that selecting the *absent* response builds up from the lack of evidence in favor of the *present* response as a result of the time spent unsuccessfully searching for the target. Importantly, our results describe a close interaction between bottom-up and top-down processes with which both visual uncertainty and expectations drive the accumulation of evidence during visual search. On a behavioral level, this is an extension of existing models of top-down recognition that suggest fast and automatic interactions between bottom-up and top-down visual streams (Bar, 2004; Beffara et al., 2015; Kauffmann, Chauvin, Guyader, & Peyrin, 2015). This raises important questions about bottom-up and top-down integration within and across fixations during scene exploration. For example, one might consider nuanced integrative schemes in which the relative contributions of top-down and bottom-up information are weighted by their relative precision (or inverse variance) and address the issue of how the time course of integration relates to the precision of information and confidence levels (Ambrosini, Pezzulo, & Con-

stantini, 2015; Friston et al., 2012; Pezzulo, Rigoli, & Friston, 2015). Further psychophysical, neuroimaging, and modeling studies will need to establish how the current behavioral data could be understood within the theoretical frameworks provided in cognitive neurosciences.

At the methodological level, we think it is particularly interesting to study eye guidance in real-world scenes by considering the decision-making process underlying the visual search task to which the eye movements contribute. In that respect, one effective solution is to combine mouse tracking with eye tracking. This study provides an early experimental account of the potential of such a method. It is a potentially fruitful perspective for goal-oriented vision in that it focuses on a process common to a broad range of perceptual functions, including target identification (Trapp & Bar, 2015), face perception (Freeman & Ambady, 2011; Freeman, Penner et al., 2011), word reading (Barca et al., 2016; Barca & Pezzulo, 2012, 2015), and choosing between available actions for humans (Lemonnier, Brémond, & Baccino, 2014). If both experimental approaches were taken into account by means, for instance, of mouse-tracking tasks and/or signal detection theory and computational approaches implementing decision-making with eye movements (e.g., Friston et al., 2012; Quinton et al., 2014) as well as fixation durations (e.g., Nuthmann et al., 2010), it would be particularly beneficial for further understanding of how information is continuously and actively collected in order to reduce the uncertainty surrounding competing perceptual decisions in visual search.

Box 1: Summary of the main results

Eye movement measures: Targets placed in unusual locations led to detection time and more fixations during the detection epoch compared to targets in usual locations. In contrast to these results, noise increased verification time and fixation duration during the detection epoch.

Response time: Both targets in an unusual location and degradation with noise increased the RT. The effect of target location was mediated by detection time whereas the effect of noise was mediated by verification time, thus clearly distinguishing the influence of both on the RT.

Mouse velocity on x-axis: The *x*-velocity began increasing before the moment of target detection and was higher for unusual target locations than for usual from 288 ms to 128 ms before target detection to 416 ms to 496 ms after target detection. The peak of the *x*-velocity profiles occurred 106 ms to 191 ms after the end of target verification, and their amplitude was decreased by noise. Additionally,

noise increased x -velocity in the latest parts of the profile. Hence, target location influenced the early part of the response movement whereas noise influenced the late part.

Area under the curve (global attraction by absent response): Although noise had no significant influence, the AUC was higher for targets in unusual locations than for targets in usual locations. This effect was modulated by detection times: Higher detection times further increased the AUC for targets in unusual locations compared to usual locations.

Deviation on x -axis: The deviation from the ideal trajectory (straight movement) began increasing before the moment of target detection. It increased earlier and overall was larger for targets in unusual locations, indicating a competition between responses early on. Furthermore, detection time increased x -deviation before the moment of target detection and through most of the mouse trajectory, and the effect was greater for targets in unusual location.

Keywords: visual search, perceptual decision-making, eye movements, mouse tracking, context information, scene perception

Acknowledgments

This study was approved by the French Committee for the Protection of Persons “South-East VI” (Comité de Protection des Personnes “Sud-Est VI”; IRB00008526, CE33, 2013). The authors would like to thank Mélanie Marcotte and Jessica Guilhe for their contribution to the visual scene selection and data collection. This work was funded by grants from the French program “investissement d’avenir” managed by the National Research Agency (ANR) from the European Union (Auvergne European Regional Development Funds—ERDF—of Auvergne region) and from the “Région Auvergne” in the framework of the IMobS3 LabEx (ANR-10-LABX-16-01).

Commercial relationships: none.

Corresponding author: Marie Izaute.

Email: marie.izaute@univ-bpclermont.fr.

Address: Clermont University, Blaise Pascal University, LAPSCO, Clermont-Ferrand, France.

References

- Ambrosini, E., Pezzulo, G., & Constantini, M. (2015). The eye in hand: predicting others’ behavior by integrating multiple sources of information. *Journal of Neurophysiology*, *113*(7), 2271–2279, doi:10.1152/jn.00464.2014.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412, doi:10.1016/j.jml.2007.12.005.
- Bajcsy, R., Aloimonos, Y., & Tsotsos, J. K. (2016). Revisiting active perception. *Perception*, 1–22.
- Ballard, D. H., & Hayhoe, M. M. (2009). Modelling the role of task in the control of gaze. *Visual Cognition*, *17*(6–7), 1185–1204, doi:10.1080/13506280902978477.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629, doi:10.1038/nrn1476.
- Barca, L., Benedetti, F., & Pezzulo, G. (2016). The effects of phonological similarity on the semantic categorisation of pictorial and lexical stimuli: Evidence from continuous behavioural measures. *Journal of Cognitive Psychology*, *28*(2), 159–170, doi:10.1080/20445911.2015.1101117.
- Barca, L., & Pezzulo, G. (2012). Unfolding visual lexical decision in time. *PLoS One*, *7*(4), e35932, doi:10.1371/journal.pone.0035932. [Article]
- Barca, L., & Pezzulo, G. (2015). Tracking second thoughts: Continuous and discrete revision processes during visual lexical decision. *PLoS One*, *10*(2), e0116193, doi:10.1371/journal.pone.0116193. [Article]
- 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, doi:10.1016/j.jml.2012.11.001.
- Beffara, B., Wicker, B., Vermeulen, N., Ouellet, M., Bret, A., Funes Molina, M. J., & Mermillod, M. (2015). Reduction of interference effect by low spatial frequency information priming in an emotional Stroop task. *Journal of Vision*, *15*(6):16, 1–9, doi:10.1167/15.6.16. [PubMed] [Article]
- Birem, M., Quinton, J.-C., Berry, F., & Mezouar, Y. (2014). SAIL-MAP: Loop-closure detection using saliency-based features. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4543–4548). IEEE, doi:10.1109/IROS.2014.6943206.
- Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, *71*(4), 881–895, doi:10.3758/APP.71.4.881.
- Borji, A., Tavakoli, H. R., Sihite, D. N., & Itti, L.

- (2013). Analysis of scores, datasets, and models in visual saliency prediction. In *2013 IEEE International Conference on Computer Vision* (pp. 921–928). IEEE, doi:10.1109/ICCV.2013.118.
- Brockmole, J. R., & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology*, *59*(7), 1177–1187, doi:10.1080/17470210600665996.
- Brockmole, J. R., & Le-Hoa Vo, M. (2010). Semantic memory for contextual regularities within and across scene categories: Evidence from eye movements. *Attention, Perception, & Psychophysics*, *72*(7), 1803–1813, doi:10.3758/APP.72.7.1803.
- Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics*, *72*(5), 1283–1297, doi:10.3758/APP.72.5.1283.
- Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review*, *18*(5), 890–896, doi:10.3758/s13423-011-0107-8.
- Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 753–763, doi:10.1037/0096-1523.33.4.753.
- Castelhano, M. S., Pollatsek, A., & Cave, K. R. (2008). Typicality aids search for an unspecified target, but only in identification and not in attentional guidance. *Psychonomic Bulletin & Review*, *15*(4), 795–801, doi:10.3758/PBR.15.4.795.
- Catenacci Volpi, N., Quinton, J. C., & Pezzulo, G. (2014). How active perception and attractor dynamics shape perceptual categorization: A computational model. *Neural Networks*, *60*, 1–16, doi:10.1016/j.neunet.2014.06.008.
- Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, *35*(1), 15–28, doi:10.3758/BF03195938.
- 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, doi:10.1111/j.1467-9280.2006.01815.x.
- Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*, *17*(6–7), 945–978, doi:10.1080/13506280902834720.
- Farmer, T., Cargill, S., Hindy, N., Dale, R., & Spivey, M. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science: A Multidisciplinary Journal*, *31*(5), 889–909, doi:10.1080/03640210701530797.
- Faulkenberry, T. J. (2014). Hand movements reflect competitive processing in numerical cognition. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *68*(3), 147–151, doi:10.1037/cep0000021.
- Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8*(2):6, 1–17, doi:10.1167/8.2.6. [PubMed] [Article]
- Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, *20*(10), 1183–1188, doi:10.1111/j.1467-9280.2009.02422.x.
- Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241, doi:10.3758/BRM.42.1.226.
- Freeman, J. B., & Ambady, N. (2011). Hand movements reveal the time-course of shape and pigmentation processing in face categorization. *Psychonomic Bulletin & Review*, *18*(4), 705–712, doi:10.3758/s13423-011-0097-6.
- Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General*, *137*(4), 673–690, doi:10.1037/a0013875.
- Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*, 1–6, doi:10.3389/fpsyg.2011.00059. [Article]
- Freeman, J. B., Ma, Y., Han, S., & Ambady, N. (2013). Influences of culture and visual context on real-time social categorization. *Journal of Experimental Social Psychology*, *49*(2), 206–210, doi:10.1016/j.jesp.2012.10.015.
- Freeman, J. B., Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology*, *46*(1), 179–185, doi:10.1016/j.jesp.2009.10.002.
- Freeman, J. B., Penner, A. M., Saperstein, A., Scheutz,

- M., & Ambady, N. (2011). Looking the part: Social status cues shape race perception. *PloS One*, *6*(9), e25107, doi:10.1371/journal.pone.0025107. [Article]
- Frintrop, S. (2006). *VOCUS: A visual attention system for object detection and goal-directed search*. In J. G. Carbonell & J. Siekmann (Eds.). Berlin Heidelberg, New York: Springer.
- Friston, K. (2012). Embodied inference and spatial cognition. *Cognitive Processing*, *13*(Suppl. 1), S171–S177, doi:10.1007/s10339-012-0519-z.
- Friston, K., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, *3*, 1–20, doi:10.3389/fpsyg.2012.00151. [Article]
- Harley, E. M., Dillon, A. M., & Loftus, G. R. (2004). Why is it difficult to see in the fog? How stimulus contrast affects visual perception and visual memory. *Psychonomic Bulletin & Review*, *11*(2), 197–231, doi:10.3758/BF03196564.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420, doi:10.1080/03637750903310360.
- Helman, E., Stolier, R., & Freeman, J. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Process & Intergroup Relations*, *8*(3), 384–401, doi:10.1177/1368430214538325.
- Henderson, J. M., Brockmole, J. R., Castelano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. P. G. Van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Amsterdam: Elsevier.
- Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, *17*(6–7), 1055–1082, doi:10.1080/13506280802685552.
- Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306, doi:10.1016/j.visres.2008.09.007.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259, doi:10.1109/34.730558.
- Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*(1), 116–131, doi:10.1037/a0025335.
- Kanan, C., Tong, M. H., Zhang, L., & Cottrell, G. W. (2009). SUN: Top-down saliency using natural statistics. *Visual Cognition*, *17*(6–7), 979–1003, doi:10.1080/13506280902771138.
- Kauffmann, L., Chauvin, A., Guyader, N., & Peyrin, C. (2015). Rapid scene categorization: Role of spatial frequency order, accumulation mode and luminance contrast. *Vision Research*, *107*, 49–57, doi:10.1016/j.visres.2014.11.013.
- Kietzmann, T., Geuter, S., & König, P. (2011). Overt visual attention as a causal factor of perceptual awareness. *PloS One*, *6*(7), e22614, doi:10.1371/journal.pone.0022614. [Article]
- Kietzmann, T., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. *Vision Research*, *110*, 76–86, doi:10.1016/j.visres.2015.02.023.
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549, doi:10.1207/s15516709cog1804_1.
- Kovalenko, L. Y., Chaumon, M., & Busch, N. A. (2012). A pool of pairs of related objects (PO-PORO) for investigating visual semantic integration: Behavioral and electrophysiological validation. *Brain Topography*, *25*(3), 272–284, doi:10.1007/s10548-011-0216-8.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565, doi:10.1016/S0042-6989(01)00102-X.
- Lemonnier, S., Brémond, R., & Baccino, T. (2014). Discriminating cognitive processes with eye movements in a decision-making driving task. *Journal of Eye Movement Research*, *7*(4), 1–14.
- Lepora, N. F., & Pezzulo, G. (2015). Embodied choice: How action influences perceptual decision making. *PLOS Computational Biology*, *11*(4), e1004110, doi:10.1371/journal.pcbi.1004110.
- Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11):8, 1–13, doi:10.1167/9.11.8. [PubMed] [Article]
- Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2):4, 1–11, doi:10.1167/10.2.4. [PubMed] [Article]

- Marghetis, T., Núñez, R., & Bergen, B. K. (2014). Doing arithmetic by hand: Hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *The Quarterly Journal of Experimental Psychology*, 67(8), 1579–1596, doi:10.1080/17470218.2014.897359.
- Morett, L. M., & MacWhinney, B. (2013). Syntactic transfer in English-speaking Spanish learners. *Bilingualism: Language and Cognition*, 16(1), 132–151, doi:10.1017/S1366728912000107.
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382–405, doi:10.1037/a0018924.
- Papesh, M. H., & Goldinger, S. D. (2012). Memory in motion: Movement dynamics reveal memory strength. *Psychonomic Bulletin & Review*, 19(5), 906–913, doi:10.3758/s13423-012-0281-3.
- Pereira, E. J., & Castelano, M. S. (2014). Peripheral guidance in scenes: The interaction of scene context and object content. *Journal of Experimental Psychology: Human Perception and Performance*, 40(5), 2056–2072, doi:10.1037/a0037524.
- Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, 134, 17–35, doi:10.1016/j.pneurobio.2015.09.001.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891, doi:10.3758/BRM.40.3.879.
- Quétyard, B., Quinton, J.-C., Colomb, M., Pezzulo, G., Barca, L., Izaute, M., ... Mermillod, M. (2015). Combined effects of expectations and visual uncertainty upon detection and identification of a target in the fog. *Cognitive Processing*, 16(Suppl. 1), S343–S348, doi:10.1007/s10339-015-0673-1.
- Quinton, J.-C., Volpi, N. C., Barca, L., & Pezzulo, G. (2014). The cat is on the mat. Or is it a dog? Dynamic competition in perceptual decision making. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(5), 539–551, doi:10.1109/TSMC.2013.2279664.
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87, doi:10.1038/4580.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922, doi:10.1162/neco.2008.12-06-420.
- Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision*, 7(3):6, 1–17, doi:10.1167/7.3.6. [PubMed] [Article]
- Röhrbein, F., Goddard, P., Schneider, M., James, G., & Guo, K. (2015). How does image noise affect actual and predicted human gaze allocation in assessing image quality? *Vision Research*, 112, 11–25, doi:10.1016/j.visres.2015.03.029.
- Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision*, 7(14):16, 1–20, doi:10.1167/7.14.16. [PubMed] [Article]
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 6, 359–371.
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the symposium on eye tracking research & applications - ETRA '00* (pp. 71–78). New York: ACM Press, doi:10.1145/355017.355028.
- Song, J.-H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, 13(8), 360–366, doi:10.1016/j.tics.2009.04.009.
- Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences, USA*, 102(29), 10393–10398, doi:www.pnas.org/cgi/doi/10.1073/pnas.0503903102.
- Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2014). How context information and target information guide the eyes from the first epoch of search in real-world scenes. *Journal of Vision*, 14(2):7, 1–21, doi:10.1167/14.2.7. [PubMed] [Article]
- Tatler, B. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14):4, 1–17, doi:10.1167/7.14.4. [PubMed] [Article]
- Tatler, B., Hayhoe, M., Land, M., & Ballard, D. (2011). Eye guidance in natural vision: Reinterpreting saliency. *Journal of Vision*, 11(5):5, 1–23, doi:10.1167/11.5.5. [PubMed] [Article]
- Torrallba, A., Oliva, A., Castelano, 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, doi:10.1037/0033-295X.113.4.766.
- Trapp, S., & Bar, M. (2015). Prediction, context, and

- competition in visual recognition. *Annals of the New York Academy of Sciences*, 1339, 190–198, doi:10.1111/nyas.12680.
- Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruity influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, 59(11), 1931–1949, doi:10.1080/17470210500416342.
- Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition*, 17(6–7), 812–834, doi:10.1080/13506280902771278.
- Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, 18(3), 321–342, doi:10.1080/09541440500236661.
- Võ, M. L.-H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, 1339, 72–81, doi:10.1111/nyas.12667.
- Wolfe, J. M. (2012). When do I quit? The search termination problem in visual search. In *Nebraska Symposium on Motivation* (Vol. 59) (pp. 183–208). New York: Springer, doi:10.1007/978-1-4614-4794-8_8.
- Wu, C. C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology*, 5, 1–13, doi:10.3389/fpsyg.2014.00054. [Article]
- Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 3485–3492). IEEE, doi:10.1109/CVPR.2010.5539970.