Mental Model Updating and Eye Movements

by

Hanbin Go

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2020

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

Two studies investigated what eye movements can reveal about how we process surprising information and use it to update mental models. Mental models guide our actions to make decisions in a dynamic environment. Participants made saccades to visual targets presented one at a time, radially, around an invisible perimeter, while their eyes were tracked. Target locations were normally distributed and changed at an unannounced point during the task. In Experiment 1, the distribution changed to one with non-overlapping regions of target locations. Saccadic latencies were slower when targets appeared in areas of low as opposed to high spatial probability. The length of time participants looked at targets, dwell time, increased when unexpected, low probability events occurred. In a second study, the mean of the distribution was held constant, but variance changed in three ways; a narrow-to-wide variance shift; a wide-to-narrow shift, and a no-shift condition. Hence distribution shifts were not as apparent as study 1, especially for the wide-to-narrow variance shift. Participants reported trials on which they perceived a shift in the target distribution, via a mouse click. Participants were poor at determining the distributional shifts. On trials with reported distribution change, participants dwelled on the target longer and were slower to generate saccades. When presented with a narrow-to-wide distribution shift, saccadic latencies were slower for targets from the new wide-distribution (unexpected, low probability), however, no changes were observed in dwell time, suggesting that participants deemed the highly surprising events as random noise with no predictive value for future events, and hence felt no need to update their predictive model. Results suggest that slower saccadic latency reflects surprise, whereas longer dwell time reflects updating of a mental model.

Acknowledgements

I would like to express my sincerest gratitude to my supervisors, Dr. Britt Anderson and Dr. James Danckert, for their constant support, and guidance. I have been extremely lucky to have supervisors who have assisted me in developing key research skills. I would also like to extend a sincere thank you to my thesis reader, Dr. Mike Dixon for his feedback and guidance during the final stages of this thesis. Additionally, my gratitude goes to all members in the lab for their advice and insightful suggestions. Lastly, a heartfelt thank you goes to my family and partner who have encouraged and supported me throughout this process.

Table of Contents

List of Figures

Chapter 1

General Introduction

1.1 Purpose

Humans are not capable of representing everything in their environment. The world is both noisy and ever-changing, demanding that we create and update condensed representations that encapsulate the rules and regularities governing the environment. The utility of any given representation – or mental model – depends on the accuracy with which it represents regularities in the world and how flexibly it can be updated when contingencies change (Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Danckert, Stöttinger, Quehl, & Anderson, 2012). For example, a baseball player at bat may have a mental representation of the pitcher he is facing. Perhaps, through prior experiences with this pitcher he expects to see mainly fastballs. This highlights an important component of mental models – they derive utility from their predictive success. If the pitcher now throws a fastball, the model is deemed accurate. If the pitcher throws a slider the batter now faces a dilemma – is this new pitch just random noise (a fluke), or does the batter need to update his mental model to include new expectations (Filipowicz, Valadao, Anderson, & Danckert, 2018; Johnson-Laird, 2010).

In the everyday example given above, the unexpected information may prompt an updating event. In some sense, the batter must determine whether unexpected or surprising events are: random fluctuations where the existing model need not be changed; events that necessitate a fine tuning of an existing model; events that warrant the adoption of a completely new model. There is an important distinction to be made here between low probability but unexpected events and expected events (unlikely in a particular observation) – both of which may be surprising to us, but for different reasons (O'Reilly et al., 2013). In statistical learning terms, one can create a mental model capturing the regularities of events that can be defined stochastically (Collins & Koechlin, 2012). Low probability events, although surprising, are still predicted by the model. An unexpected event on the other hand, is one that is surprising because it is not predicted by the model in any way (Nassar, Wilson, Heasly, & Gold, 2010; O'Reilly et al., 2013; Teigen & Keren, 2003).

Regardless of what makes an event surprising, research has suggested a linear relationship between surprise and updating (McGuire, Nassar, Gold, & Kable, 2014) – with more surprising events leading to an increased probabilistic updating of the associated motor response (Mars et al., 2008). However, work by Filipowicz and colleagues (2018) suggests that this may not be the case. Participants in this study estimated the distribution of ball drops in a virtual Plinko game. At unannounced time points the distribution presented to subjects shifted in various ways that could be quantified in terms of surprise level. Poor updating performance was observed for shift in distributions that were characterised by the highest level of surprise, as though participants discounted the most surprising events relative to a learned distribution (Filipowicz, et al., 2018).

Research by O'Reilly and colleagues (2013) examined different responses to surprising events using eye tracking as a metric of updating performance. Participants were asked to make saccades to targets presented in an annular array around a fixation. Targets were chromatically coloured, each colour was associated with a distinct distribution of target locations (e.g., all the targets from the first distribution would be red, and targets from the subsequent distribution would be blue). A change in target colour signaled to the participant that there had been a shift in the distribution. In addition, grey targets, not belonging to either distribution were presented on 20% of the trials. These events were surprising but uninformative – they did not signal any change to the underlying distribution of target locations. In contrast, a change in target colour was surprising but was also informative to the subject. In both instances, the time to initiate a saccade (saccadic latency) to the surprising event was slower. In other words, regardless of whether or not a surprising event signalled a change in distribution (e.g., target colour changed from red to blue) or was uninformative (i.e., a grey target), saccadic latencies were slowed. In contrast, the amount of time participants spent looking (i.e., dwell time) at these surprising, unexpected events did differ. Participants looked longer at targets signalling a change in distribution (e.g., a red target where previous targets had been blue), when compared to targets that did not signal a change (i.e., grey targets). In other words, saccadic latency seemed to track whether or not an event was unexpected/surprising, whereas dwell time was associated with the need to update a mental model of the expected distribution of target locations.

Although participants in the O’Reilly study were unaware of the *timing* of a change in distributions, the fact that a change had occurred was explicitly signalled by the change in target colour. In many instances changes in environmental contingencies are not so explicitly signalled (Albrecht & O’Brien, 1993). In this thesis, the aim was to examine the capacity of participants to update mental models when the change in distributions was not explicitly represented by a concomitant change in perceptual characteristics of the events to be represented. To do this, O’Reilly’s design was modified so that changes in the distribution of target locations occurred at unannounced time points and was not associated with any change in colour of the targets. In addition, as in the work described above by Filipowicz and colleagues (2018), surprise was quantified so that the relation between updating and surprise could be examined.

In the first experiment presented here there were no surprising but uninformative events (i.e., the grey targets in the O'Reilly et al., study). Rather, we contrasted saccadic latency and dwell time for low and high probability events within learned distributions. Essentially, we sought to ascertain whether these saccade metrics could distinguish between low and high probability events within an established mental model. In addition, these same metrics were then contrasted for target events across a shift in distributions. In other words, the pre-shift analyses examined whether saccade metrics would distinguish between low and high probability events that would both be predicted by a current mental model. Such low probability events may be considered surprising but *expected*, given they are predicted by the learned mental model. In contrast, examining how saccade metrics change in the post-shift distribution was cast in terms of responses to surprising *unexpected* events.

The second experiment presented here replicated and extended the results of the first by presenting target distributions that overlapped (there was no overlap in distributions presented in Experiment 1), as well as presenting targets that were uninformative of the underlying distributions (10% of trials). In both experiments, an attempt was made to capture explicit reports of the participant's mental representation of target distributions. In the O'Reilly work discussed above, saccade metrics were taken as an implicit measure of the underlying mental model participants are using to predict impending target locations. In the experiments presented here we sought to associate saccade metrics with more explicit reflections of the underlying representation (i.e., participants were told to make eye movements to locations where they expected targets to appear in Experiment 1, and to click on a mouse when they detected a distribution shift in Experiment 2) – explicit measures which would reveal the mental model held by the participants. It was hypothesized that the participants would have slower saccadic latencies towards unexpected/surprising events, and would dwell on targets longer during fine tuning, and updating of their mental models.

Chapter 2

Experiment 1

2.1 Introduction

In Experiment 1, saccades and dwell time metrics were observed while participants completed a saccadic eye movement response task. These objective metrics of performance were observed to examine their association with the contingencies manipulated in the experiment – whether individuals could build and update a mental representation of where the targets are likely to appear (i.e., spatial distribution of targets). In other words, we sought to address whether participants could utilize internal model to facilitate efficient saccadic responses to the targets, and make appropriate changes to their model when contingencies change (i.e., distribution shift). It is possible that future target locations could be anticipated due to the fact that targets generally appeared in similar location for the same distribution run. The task was similar to the O'Reilly et al., (2013) task, in which participants made speeded eye movements towards targets, and at an unannounced point, the distribution shifted. However, unlike the O'Reilly et al., task, the presentation of a new distribution of targets was not signalled explicitly by a colour change. Rather, the new distribution was indicated by targets in completely new locations that were never used in the preceding distribution. The goal was to examine whether participants could detect the distribution shift with the absence of explicit signals (i.e., colour change in the O'Reilly et al., task) and whether eye movements can implicitly reflect surprise and updating of mental models. It was hypothesized that participants would exhibit longer saccadic latencies for targets which appeared in low probability/unexpected locations (surprising events). Whereas, longer dwell times would be expected when a shift in the distribution was noticed and prompted an updating of the mental model.

2.2 Methods

2.2.1 Participants

Forty undergraduate and graduate students from the University of Waterloo participated in this in-lab experiment in exchange for a course credit or $10. Seven participants could not complete the study due to calibration difficulties. Therefore, a total of thirty-three participants were included in the analysis (median age = 20 years; 24 right eye-dominant; 28 female). This study was given ethics clearance by the University of Waterloo's Office of Research Ethics, and all participants provided informed written consent prior to their participation.

2.2.2 Apparatus and Procedure

Participants viewed the stimuli from a gamma corrected 19-inch CRT monitor (horizontal refresh rate of 91.1 kHz, vertical refresh rate of 85.0 Hz). To reduce head movements, a chin and forehead rest was used with a fixed distance of 60 cm from the monitor. Eye movements were recorded using an EyeLink 1000 Plus eye tracker, with a sampling rate of 1000 Hz. The experiment began with a practice block with targets that appeared from a uniform distribution around a circular perimeter (i.e., the targets did not appear in one specific area of the circle, as they were equally distributed). In this block, the participants had the opportunity to be familiarized with making saccadic responses to targets that appeared in different areas of the circle.

In a single trial, participants fixated their gaze on a central fixation point which prompted a visual target to appear from an underlying spatial distribution. Participants were to make a saccade to the target as quickly as possible and were instructed to fixate back on the central fixation point to prompt a new target. Every five trials, they were instructed to look at the areas where they would predict the next few dots to appear (Figure 2.1).

All targets were equidistant from the central fixation, appearing on a circular perimeter not visible to the participants. This was to ensure that the position of the targets was only controlled by a single parameter, the degrees of angle from horizontal; put differently, latency of saccades would not be influenced by saccadic amplitude, given that all targets were equidistant from the central fixation point.

This experiment was designed to be gaze contingent. That is, targets were not presented until a central fixation was established. This was done by fixating within 1.5 degrees of visual angle from the fixation circle in the centre of the screen. Additionally, to ensure that participants did indeed process the target, it turned white only when the gaze was within 1.0 degrees of visual angle of the target location and remained white as long as the gaze was within 1.5 degrees of visual angle of the target.

A total of nine fixation points were used to calibrate participants to the eye tracker prior to beginning the blocks. Once calibrated, participants were instructed to maintain their head rested on the chin and forehead rest to reduce head movement. If any head movement occurred, participants needed to be re-calibrated to the eye tracker. Re-calibration also occurred when the trial timed out, as prompted by the experiment (i.e., when participants were not able to locate and fixate on a target within 5 seconds).

Every five trials, the participants completed an “estimate” phase, and had the option to explicitly update their current model. To do so, participants gazed at a range of possible locations inside a darker grey ring (which visually guided their gaze to be within 5.0 to 11.0 degrees of visual angle away from the centre of the screen). To assist the participants with the awareness of their gaze positions, a trace of their last five gaze positions was visually represented on the screen (i.e., a trail of black dots).

Participants completed a total of three blocks, with 150 experimental trials in a given block. Before the start of each block, participants were told to disregard their previously held information regarding the distribution of target locations. To avoid confusion, targets in each of the three blocks had a distinct target colour (i.e., red, green, or blue). Changes in the target colour were not related in any way to the changes in the target distributions; the target colour solely represented the beginning of a new block.

Within each block there was a shift in the distribution. The shift occurred on average after 75 trials (with shifts occurring between the 60th and 90th trial). Thus, two separate distributions were presented in a given block. The target dots were generated from a circular gaussian distribution, with a fixed mean (randomly selected to be between 0 and 2π) and variance for each distribution run. The distribution mean and variance abruptly shifted to new values to signify the start of a new distribution run. The new distribution was selected from with a non-overlapping region from the circle (Figure 2.2).

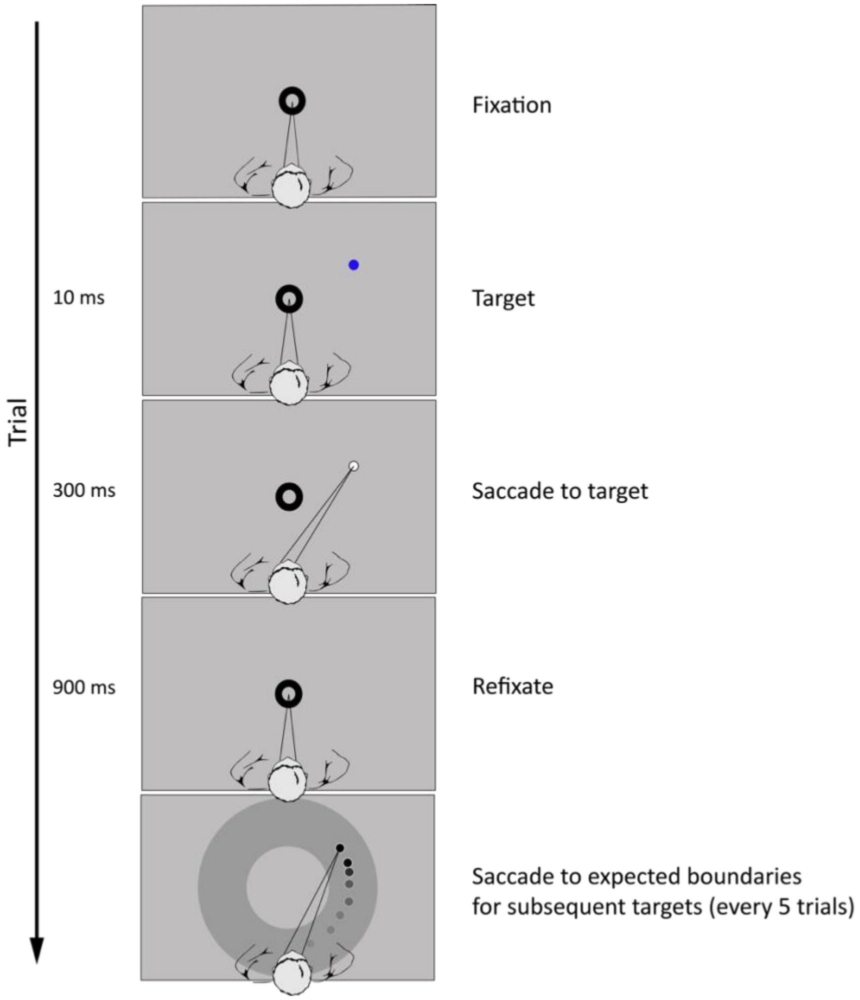

Figure 2. 1. Time course of each trial for Experiment 1. Once the gaze was on the central fixation, a target appeared with no delay, from a gaussian distribution on an invisible circular perimeter, and participants were to saccade to the target as quickly as possible. Once participant's gaze was fixated on the target, the target turned white, as a form of feedback. Every five trials, participants looked at the areas where they expect the next few target dots to appear. The black dots in the bottom panel indicate the fact that participants were instructed to look not at a single location, but to scan the range of possible locations they thought were most likely for the subsequent targets.

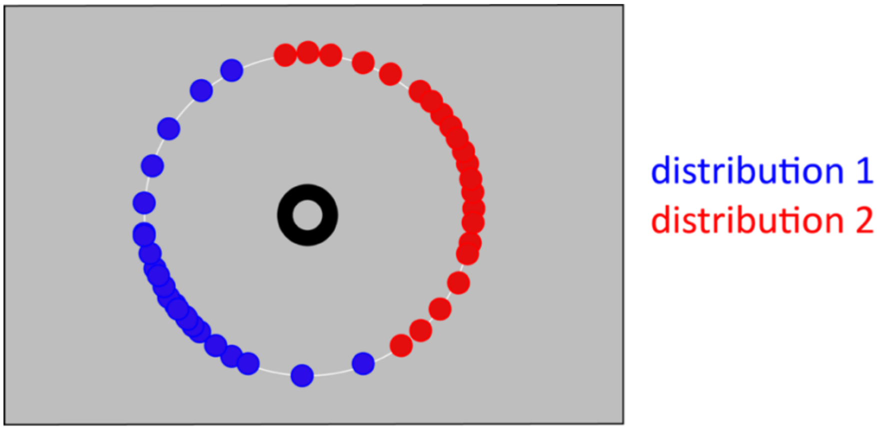

Figure 2. 2. The shift in the distribution of targets for Experiment 1 occurred at an unannounced time point and was only indicated by the non-overlapping region of target locations. The distribution change was not associated with any change in target colour, as the colour remained the same for the entire block (e.g., target colours were blue for both distribution 1 and distribution 2 in the same block). The colour differentiation in this figure is for visual illustration purposes, where blue displays targets drawn from the first distribution run, and the red target dots signify targets drawn from the subsequent distribution. Furthermore, targets were normally distributed, hence target appearance near the mean was more frequent than targets further away from the mean.

2.2.3 Data Analytic Plan

Participants learned two discrete distribution of targets in a given block. In a given distribution run, the targets appeared in similar region over multiple trials as it was normally distributed. Thus, most of the targets appeared around the mean of the distribution. Within a distribution, the targets appearing around the mean can be categorized as high probability, and targets appearing around the tails of the distribution as low probability. Specific details on how this was computed is discussed in the preceding section. Differences between high and low probability target locations were examined for both saccadic latency and dwell time. Saccadic latency was a reaction time (units of time in seconds) from the onset of the target, to the time taken to leave the central fixation threshold (i.e., 1.5 degrees of visual angle). In other words, the time taken to initiate a saccade. Dwell time was the time lapsed from the arrival of gaze on the target threshold (i.e., 1.0 degrees of visual angle), to leaving the fixation (i.e., a threshold of 1.5 degrees of visual angle). Saccadic latency and dwell time were measured using a Python script with elements from the PsychoPy library.

Learning a Distribution Run

Participants completed a total of three blocks, and within each block, participants completed two distribution runs (an original distribution whose properties participants had to learn about and a shifted distribution). A distribution run was categorized into high and low probability target locations. Low-probability targets were determined as the 10 target locations that deviated the most from the median. We then calculated the spread (θ_1 and θ_2) of the low probability target points from both tails of the distribution. Lastly, high-probability targets were determined as targets that appeared between the lower threshold ($\text{median}_{\text{distribution}} - \theta_1$), and the upper threshold ($\text{median}_{\text{distribution}} + \theta_2$).

Update to a Shifted Distribution

To reiterate, in a given block, participants were presented with the initial distribution and were abruptly presented with a new shifted distribution of targets in which there was no spatial overlap with the initial distribution. To examine whether dwell time and saccadic latency differed for the targets from the shifted distribution, we computed the mean of both dwell time and saccadic latency for the 5 trials preceding the distribution shift, and 5 trials following the distribution shift.

Estimate Phase

The goal of the estimate phase was to provide a measure for when the participants' updated their mental model. Given the exploratory nature of this study, it was observed that looked at locations in the estimate phase were poor at capturing the explicit representation of updating. Participants were instructed to look at the areas where they expect the next few dots to appear, however based on their looking behaviour it was rather ambiguous as to whether they had updated their model in response to the shift in distribution. This is because the instructions were open to interpretation (e.g., “next few dots” may be interpreted as looking at all possible future target locations or looking solely at where the next five targets are going to appear). For instance, if subject 1 were to scan the range of all possible locations (e.g., estimate of the spatial distribution of target locations), versus subject 2 glancing at a specific area (e.g., high probability region), when computing for the differences in the mean gaze locations (angle in radians) between each of the estimate trials, the angular distance for subject 1 would be smaller compared to subject 2 (e.g., greater variation in the mean gaze locations across each estimate phase for subject 2), however no conclusion can be drawn about model updating. Thus, estimate phase was not analyzed for this study.

2.2.4 Hypothesis

As individuals saccade toward target stimuli drawn from a single distribution with which they were familiar, it is predicted that the planning of saccades would be shorter for high-probability (as opposed to low-probability) locations. Furthermore, since dwell time is linked to unexpected, rather than merely infrequent events, a non-significant difference in dwell time is to be expected between high and low probability-locations, since participants’ mental models include both types of locations.

A distribution shift at an unexpected time point, which present targets from a non-overlapping region from the learned distribution, would be a novel, unexpected, location.

Following such an unexpected event, it was hypothesized that participants will have longer saccadic latencies when compared to latencies from the preceding known distribution. Crucially, in accordance with the surprising event, they should spend a longer time looking at the target drawn from the new distribution (noted by an increase in dwell time).

2.3 Experiment 1: Results

2.3.1 Data Screening

Outliers for saccadic latencies were determined by the Interquartile range rule, calculated by adding 1.5 * IQR to the third quartile. Therefore, saccadic latencies that were slower than the specified rule, were considered outliers (< 4.0%) and were removed from the analyses. Factors that may have contributed to such outliers were calibration inconsistencies and blinks.

2.3.2 Learned Distribution

Saccadic Latency

Saccadic latencies were contrasted for high and low probability events in both distribution 1 and 2. A linear mixed-effects model was used to predict saccadic latency, using R Studio (RStudio Team, 2017) to run the nlme package (Pinheiro & Bates, 2019) in the R statistical analysis environment (R Core Team, 2017). With a fixed effect of probability conditions (i.e., high and low) and a random effect of participants to predict saccadic latency. The analysis yielded a marginal main effect of probability condition on saccadic latency, ($\text{estimate}_{\text{high-probability}}$ = -0.005, SE = 0.003, $t(372) = -1.91$, $p = 0.057$). Depicting shorter saccadic latencies for high probability events when compared to low probability events (Figure 2.3).

Dwell Time

A linear mixed-effects model with a fixed effect of probability conditions (i.e., high and low) and a random effect of participants, predicting dwell time revealed a negligible effect, ($\text{estimate}_{\text{high-probability}} = 0.001$, $SE = 0.007$. $t(372) = 0.07$, $p = 0.94$). No significant difference in dwell times were observed between high and low probability events. (Figure 2.4).

2.3.3 Update to a New Distribution

Saccadic Latency

Saccadic latencies were contrasted between 5 trials preceding and 5 trials following the distribution switch. A repeated measures ANOVA with factors of distribution (i.e., distribution 1 versus distribution 2) and block (i.e., block 1, 2, and 3) yielded a significant main effect of distribution shift, ($F(1,32) = 7.70$, $MSE = 0.00$, $p < 0.05$). Additionally, performance was homogenous across all blocks, ($F(2,32) = 1.32$, $MSE = 0.00$, $p = 0.28$). Lastly, there was no distribution by block interaction, ($F(2,64) = 0.84$, $MSE = 0.00$, $p = 0.44$). The analysis was followed up with a paired-samples t-test to make post hoc comparisons between distribution conditions, which revealed slower saccadic latencies to targets after the distribution switch, ($MD = 0.02$, $t(34) = 2.97$, $p < 0.01$; Figure 2.5)

Dwell Time

A repeated measures ANOVA with factors of distribution (i.e., distribution 1 and 2) and block (i.e., block 1, 2, and 3) revealed longer fixation times on targets from the new distribution, ($F(1,32) = 12.11$, $MSE = 0.01$, $p < 0.005$), and a main effect of block, ($F(2,32) = 14.10$, $MSE = 0.01$, $p < 0.001$). In addition, there was no significant main effect of distribution by block interaction, ($F(2,64) = 1.66$, $MSE = 0.01$, $p = 0.20$). A paired-samples t-test post hoc analysis revealed that participants dwell on the targets longer after the distribution switch, (MD = 0.03, $t(34) = 2.81$, $p < 0.01$; Figure 2.6)

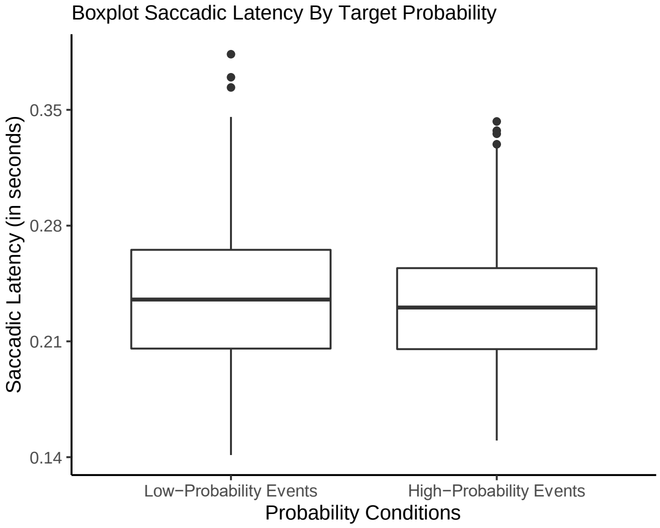

Figure 2. 3. Saccadic latency by target probability. Spatial distribution of target locations were categorized as low and high probability. Participants had shorter saccadic latencies (in seconds) for target locations that were more probable, as opposed to less probable locations

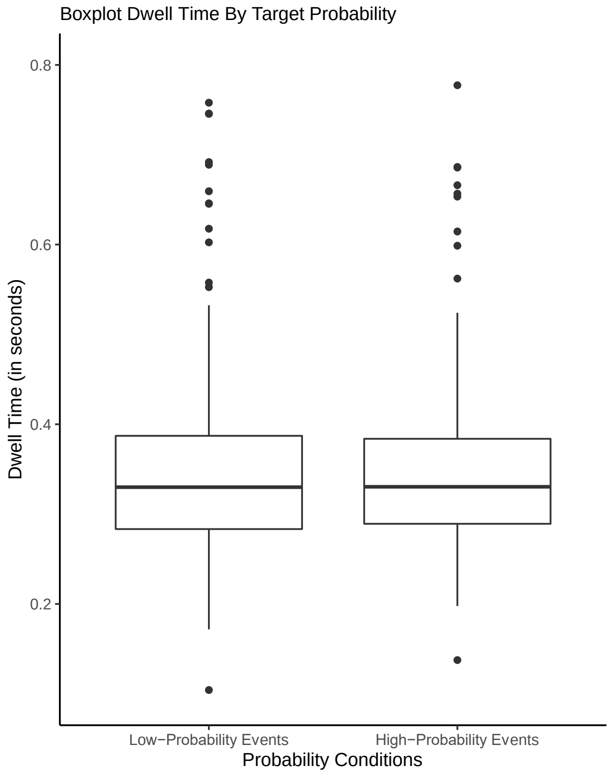

Figure 2. 4. Dwell time by target probability. No differences in dwell times were observed between low and high-probability target locations within a distribution.

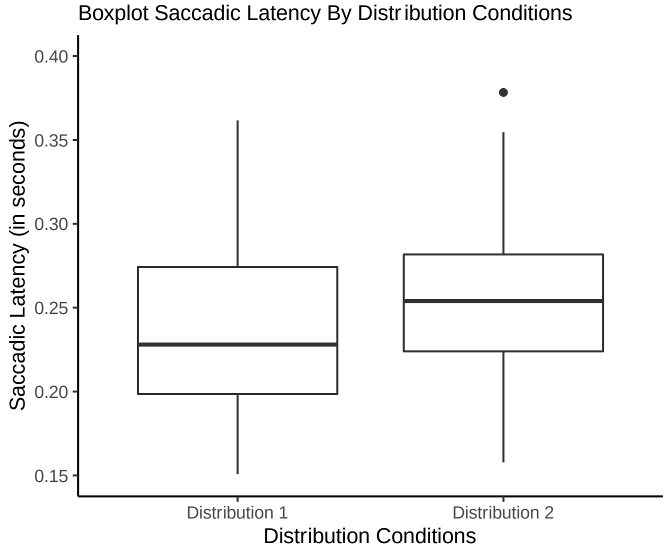

Figure 2. 5. Saccadic latency by distribution conditions. Mean saccadic latencies (in seconds), comparing 5 trials preceding the distribution shift, and 5 trials following the distribution shift. Saccades planning were slower for the unexpected low probability events.

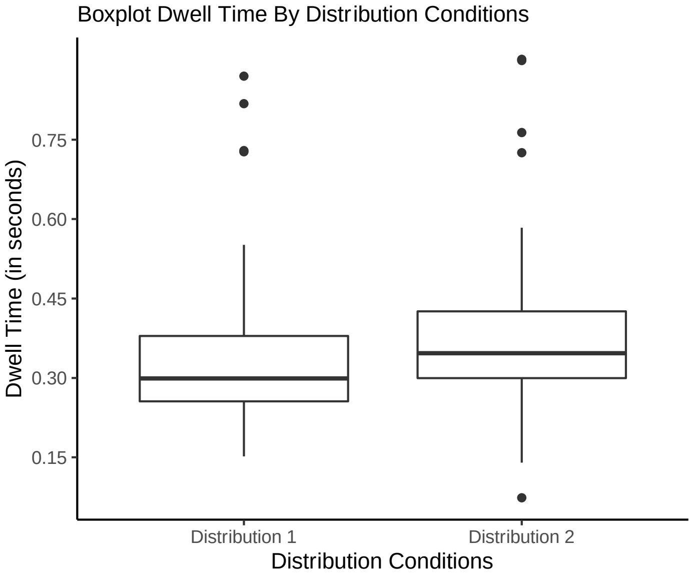

Figure 2. 6**.** Dwell time by distribution conditions. Comparing the mean dwell times (in seconds) of 5 trials preceding the distribution shift and 5 trials following the distribution shift. There was an increase in dwell time following the distribution shift.

2.4 Discussion

2.4.1 Learned Distribution

The main goal of the first experiment was to explore what we can infer about mental model updating through eye movements in an oculomotor task. Participants demonstrated their abilities to detect statistical regularities dispersed around the perimeter of an invisible circle (i.e., used prior knowledge of the spatial distribution to make speeded eye movements, where time to initiate a saccade was longer for low probability events). To perform optimally, participants would need to accurately represent where target dots were most likely to appear – in other words, create a mental model of the target distribution. This would aid the participants in accurately making speeded eye movements towards target dots. When learning a distribution, participants would have to account for both high and low probable regions of where the targets may appear to elicit goal-directed saccades; accounting for all likely polar coordinates of future targets. The analysis revealed shorter saccadic latencies for high probability events compared to low probability events. Meaning, individuals were marginally better at planning their speeded eye movements to regions where targets appeared more frequently. However, it is important to note the speed and accuracy trade off. If a saccadic landing error occurs due to preference in speed over accuracy, a correction would need to be made by a secondary saccade towards the target, which could reflect the slower saccadic latencies for the low probability events (Wu, Kwon, & Kowler, 2010). Overall, slower generation of saccadic motor plan was observed for target locations that were infrequent (i.e., low probability), yet expected to be included in the distribution - suggesting that probability influences reaction time of saccade planning.

No difference in dwell times were observed for low and high probability events, meaning participants spent equal amount of time fixating on targets that appeared in both low and high probability events. Unlike saccadic latency, low probability, yet expected events did not influence dwell time.

2.4.2 Surprise and Updating

Both saccadic latencies and dwell times were longer for targets, when a distribution shift had occurred. Targets from the new distribution were novel, unexpected, and hence surprising events, as they fell in locations that never contained targets in the previous distribution. Furthermore, the distribution shift itself would be surprising to the participants, as it was abrupt and unannounced. Despite the absence of an explicit representation by a change in a perceptual characteristic (i.e., new distribution was explicitly signaled by a change in target color in the task by O’Reilly and colleagues, (2013)), participants were able to reorient their saccadic responses to targets presented from the new distribution, however slower saccadic latencies were observed. The slower saccadic latency could reflect the time needed to extract information from noisy sensory signal (Carpenter & Williams, 1995), thus when the target is unexpected (e.g., surprising), the saccadic latency is longer. Longer saccadic latencies were followed by longer dwell times, possibly reflecting the participants updating their mental model.

2.4.3 Limitations

The present study failed to investigate the relationship between eye movement metrics with a more explicit report of the underlying representation (i.e., via explicit looking behaviour reflecting where participants predicted future target locations to lie based on their mental model of the distribution). Thus, Experiment 2 was designed to address this issue. Furthermore, changes in environmental contingencies may not be as transparent (i.e., distribution shift to non-overlapping target locations). Hence, the question to address is whether eye tracking metrics could provide insight on surprise and updating when changes in the environment are more ambiguous. Lastly, there were no surprising but uninformative events (i.e., an event with no predictive value for future events), therefore, this was also examined in the second Experiment.

Chapter 3

Experiment 2

3.1 Introduction

The first Experiment demonstrated that individuals can detect statistical regularities of an environment, and saccadic latencies were responsive to target probabilities. Furthermore, longer saccadic latencies reflect infrequent target locations, where longer dwell time reflects updating following surprising events. Experiment 2 examined whether saccadic latencies and dwell times could provide an implicit representation of surprise and updating when distribution shifts are not as evident. Thus, less salient distribution shift conditions were implemented: a wide to narrow variance shift (wide-to-narrow), a narrow to wide variance shift (narrow-to-wide), and a no shift in mean or variance (no-shift) (Figure 3.1). The different shift conditions allowed to vary the surprise factor, as wide-to-narrow would be considered a medium-surprise condition, as opposed to narrow-to-wide (high-surprise condition) (Filipowicz et al., 2018). Furthermore, this Experiment incorporated trials that had no predictive value relating to the distribution of target locations (uninformative trials). This was done to dissociate the effects of surprise from the process of updating a mental model. In theory, participants would be surprised by low probability events from both the trials with predictive value and uninformative trials. However, would solely update on the trials with predictive value, as participants were informed in advance that uninformative trials are not related in any way to the underlying distribution of targets locations. The uninformative trials were not perceived by a colour change (i.e., grey targets in O’Reilly et al., task (2013)), but rather by their spatial location. This was done by varying the distance of the target locations from the centre (Figure 3.2). Lastly, a refinement of explicit report was made, as individuals were to make a single mouse click as soon as they detected a shift in the distribution, this report was measured to examine the relationship between implicit and explicit representation of updating.

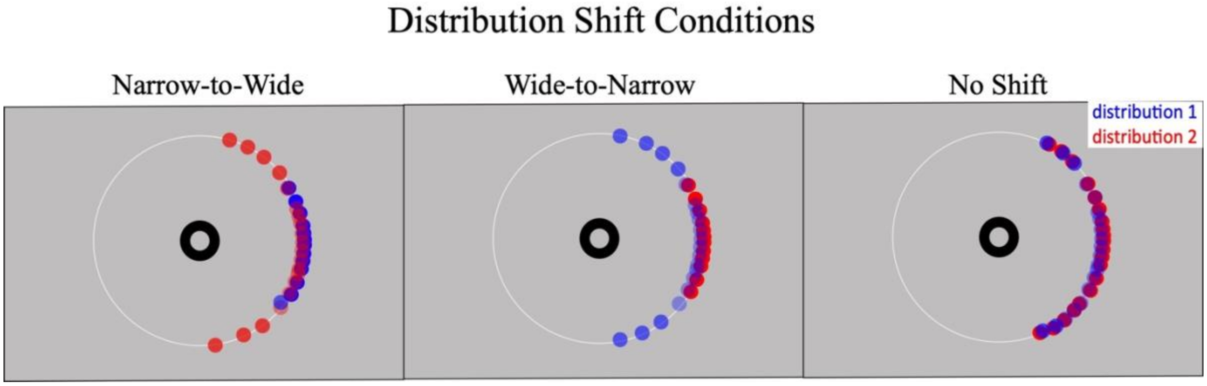

Figure 3. 1. A diagram of distribution shift conditions for Experiment 2. At an announced point in the experiment, the distribution shifted: narrow-to-wide (both distributions had the same mean, but different variance) or wide-to-narrow (same mean, but different variance as well), or no shift occurred (same mean and variance). The blue dots in this diagram represent potential targets from distribution 1, and red dots represent possible targets from distribution 2. The colour difference in this diagram is for visual illustration purposes, the participants viewed a series of target dots and the targets remained the same colour for the entire experimental block (e.g., same target colour for narrow-to-wide shift). Furthermore, the circular perimeter was not visible to the participants.

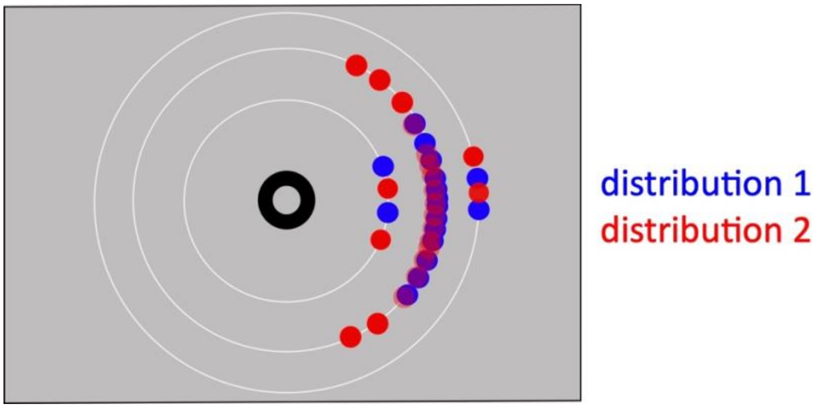

Figure 3. 2. A diagram of potential target locations in a given block, which consisted of two distribution runs. Uninformative trials were the targets that appeared closer or further away from the central fixation by 2.0 degrees of visual angle. This diagram is an example of a narrow-to-wide distribution shift.

3.2 Methods

3.2.1 Participants

Ninety-eight undergraduates from the University of Waterloo participated in the study in exchange for a course credit. Fourteen participants were not able to continue with the study, due to calibration issues. Thus, a total of eighty-four participants were included in the data analyses. (Median age = 19; 49 right eye-dominant; 66 female). The experiment was approved by the University of Waterloo's Office of Research Ethics and all participants provided informed written consent prior to their participation.

3.2.2 Apparatus

Experiment 2 used the same monitor, display settings, viewing distance, eye-tracker, chin and head rest, and sampling rate as the first experiment (see 2.2.2). In addition to target locations that were previously used, (8.0 degrees of visual angle away from the central fixation), this current study included uninformative trials, with targets that appeared 6.0 and 10.0 degrees of visual angle away from the central fixation. There was a total of 15 uninformative trials (10%) in a block of 150 trials, put differently, ~7 uninformative trials (10%) in a distribution run.

3.2.3 Procedure

Experiment 2 followed a similar procedure to the first Experiment, with a few changes. Instead of measuring explicit representation of updating with explicit looking behaviour (i.e., by having participants look at regions where they predict the next few target locations to appear); participants were instructed to click the mouse when they detected a distribution shift (Figure 3.3). Furthermore, uninformative trials were introduced, and the participants were informed that the targets that appeared closer or further away had no predictive value relating to the distribution of target locations they had to learn. Additionally, the last block of the experiment was always the no-shift in distribution condition, and the first two blocks were counter balanced to be either wide-to-narrow or narrow-to-wide shift. The no-shift in distribution condition was forced to be the last block of the experiment to provide a contrast between the subtle shifts in distributions (of which potentially some participants may not have been aware), and a no-shift condition. Similar to Experiment 1, the distribution shift occurred abruptly, however it occurred exactly half-way through the block (75 trials per distribution), which ensured that all distribution runs received even exposure – equal number of trials. Participants completed a total of 3 blocks, and the duration of the experiment took roughly twenty minutes.

3.2.4 Data Analytic Plan

Data Screening

A total of eighty-four participants were included in the data analyses. In this experiment, we extracted the saccadic latencies and dwell times from the EDF files of the EyeLink 1000 plus eye-tracker. Trials with multiple saccades to detect a target were excluded from the analyses (~30% of trials), factors such as blinks, calibration inaccuracies causing overshoot of saccadic responses, other user errors, played a role in saccadic landing errors. Of course, it is important to address, the limitation of this exclusion criteria, as saccadic landing errors could have been due to participants preference in speed over accuracy (Wu et al., 2010). Furthermore, outliers were excluded from the data analyses, determined by the Interquartile range rule, calculated by adding 1.5 * IQR to the third quartile. Thus, saccadic latencies that were slower than a specified rule were considered outliers (< 1.1%).

Distribution Shift Conditions

The mean latency in saccades and mean dwell times were analyzed for the three distribution shift conditions. Similar to Experiment 1, 5 trials preceding and following the distribution shift were analyzed. Recall that the distribution shifts were designed to be less transparent compared to the first Experiment (i.e., distribution shift had non-overlapping target locations). It was hypothesized that a narrow-to-wide distribution shift would be more evident, hence it was expected to be followed by longer saccadic latencies and dwell times. Furthermore, a wide-to-narrow distribution shift was predicted to be a very subtle shift that most participants would be unable to detect. The distribution shift would result in targets appearing from what was a high probability target location from the preceding distribution, thus it was hypothesized to observe shorter saccadic latencies for wide-to-narrow shift, and no changes in dwell-time – failing to update their model. Lastly, no difference in saccadic latencies and dwell times was predicted for the no-shift condition.

Explicit Reports of Distribution Shift

Mean saccadic latencies and dwell times on trials with reported distribution shift (i.e., reported via mouse clicks) were compared to trials in which participants did not report a change in the distribution (i.e., no mouse clicks). The analysis of the explicit report was independent of the distribution shift conditions, meaning the mouse clicks were collapsed across narrow-to-wide, wide-to-narrow, and no-shift distribution shift conditions. It was hypothesized that longer saccadic latencies and dwell times would be observed on trials where participants reported a distribution shift – participant's belief that a change in distribution has occurred.

Uninformative Trials

Uninformative trials (i.e., nears and fars) were analyzed to dissociate the effects of surprise from updating. To reiterate, these trials did not have any predictive value to future target locations, and since it only occurred 10% of the trials, uninformative trials could be deemed surprising, but uninformative. To examine whether uninformative trials were in fact surprising, difference in mean saccadic latencies were examined against uninformative trials and trials where individuals explicitly reported a change in the distribution – updating trials. It was hypothesized that no difference in saccadic latencies would be observed, as both events are expected to be surprising. To determine, whether individuals implicitly updated their mental model for uninformative trials, mean dwell times for uninformative trials were compared to mean dwell times for updating trials. It is then hypothesized that participants would not update their model for uninformative trials, thus longer dwell times for updating trials, as compared to uninformative trials, would be expected.

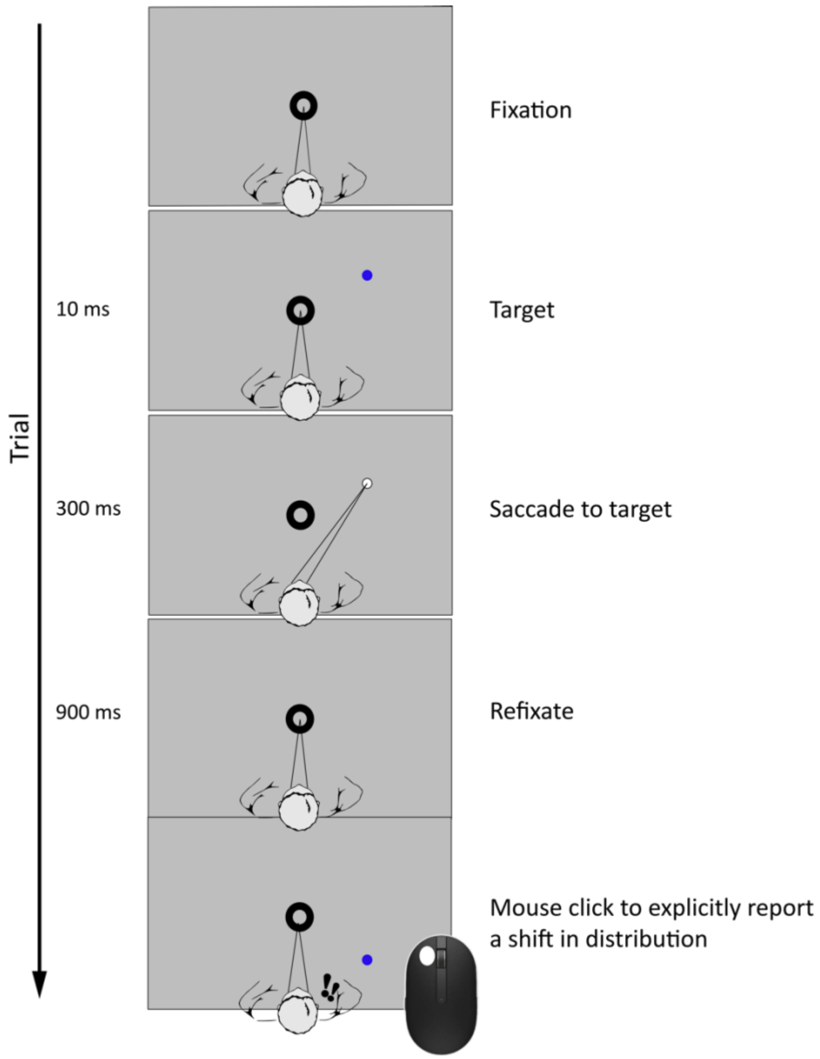

Figure 3. 3**.** Time course of each trial for Experiment 2. Once the gaze was on the central fixation, a target appeared with no delay, from a gaussian distribution around the invisible circle, and participants were to locate the target as quickly as possible. Once participant's gaze was fixated on the target, the target turned white, as a form of feedback. Participants clicked the mouse, when they believed that the distribution has shifted.

3.3 Experiment 2: Results

3.3.1 Distribution Switch Conditions

Narrow-to-Wide

A linear mixed-effects model predicting saccadic latency, with distribution shift condition as a fixed effect, and participants as a random effect, yielded a significantly longer saccadic latencies when the distribution shifted from narrow-to-wide, ($\text{estimate}_{\text{wide}}= 0.023$, SE = 0.008, $t(80) = 2.81$, $p < 0.01$) (Figure 3.4).

A linear mixed-effects model predicting dwell time, with a fixed effect of distribution shift, and a random effect of participants, revealed a non-significant difference in dwell time for a narrow-to-wide shift, ($\text{estimate}_{\text{wide}}= 0.003$, SE = 0.016, $t(80) = 0.16$, $p = 0.87$) (Figure 3.5).

Wide-to-Narrow

A linear mixed-effects model predicting saccadic latency, with a fixed effect of distribution shift condition and participants as a random effect, yielded a significant main effect of distribution shift, ($\text{estimate}_{\text{narrow}} = -0.019$, SE = 0.008, $t(81) = -2.48$, $p < 0.05$), with shorter saccadic latencies for the wide-to-narrow shift (Figure 3.4).

Similarly, a linear mixed-effects model predicting dwell time, with the fixed effect of distribution shift and the participants as a random effect yielded a significantly shorter dwell times when the distribution shifted from wide-to-narrow, ($\text{estimate}_{\text{narrow}} = -0.055$, SE = 0.020, $t(81) = -2.78$, $p < 0.01$) (Figure 3.5).

No-Shift

Lastly, linear mixed-effects models, with a fixed effect of distribution shift, and participants as a random effect, revealed no significant differences in saccadic latencies

($\text{estimate}_{\text{no-shift}}$ = 0.002, SE = 0.009, *t*(79) = 0.25, *p* = 0.81) (Figure 3.4) as well as, dwell time ($\text{estimate}_{\text{no-shift}}$= 0.000, SE = 0.014, *t*(79) = -0.02, *p* = 0.98) for the no-shift (i.e., control) condition (Figure 3.5).

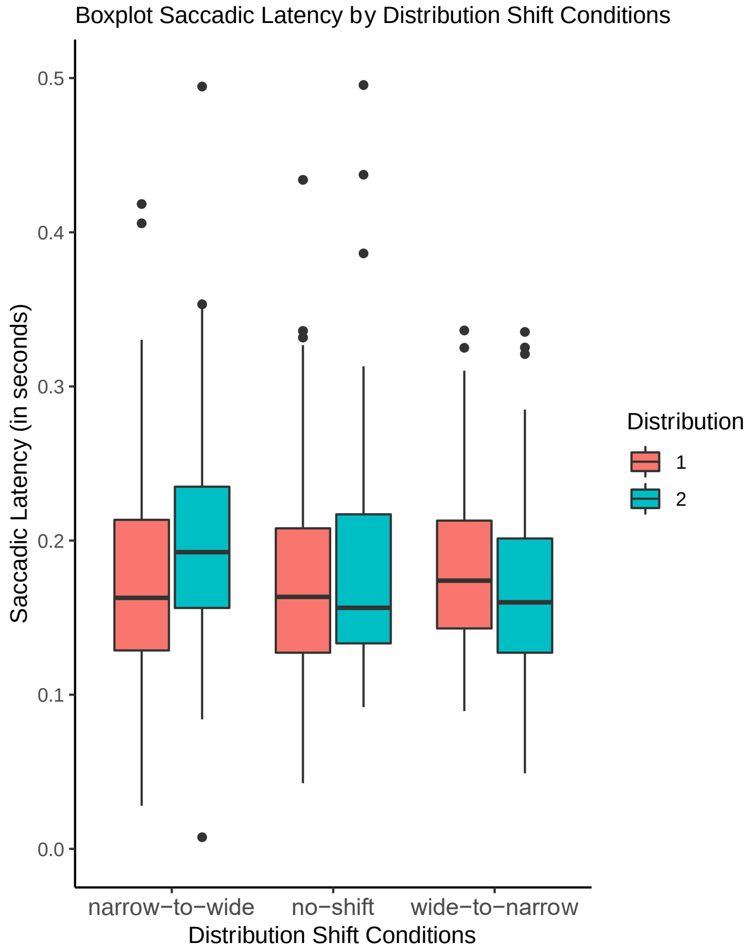

Figure 3. 4. Saccadic latency by distribution shift conditions. Participants' saccadic latencies (in seconds) were longer when the distribution shifted from narrow-to-wide, whereas shorter saccadic latencies were observed when the distribution shifted from wide-to-narrow. No differences were observed in saccadic latencies for no-shift condition.

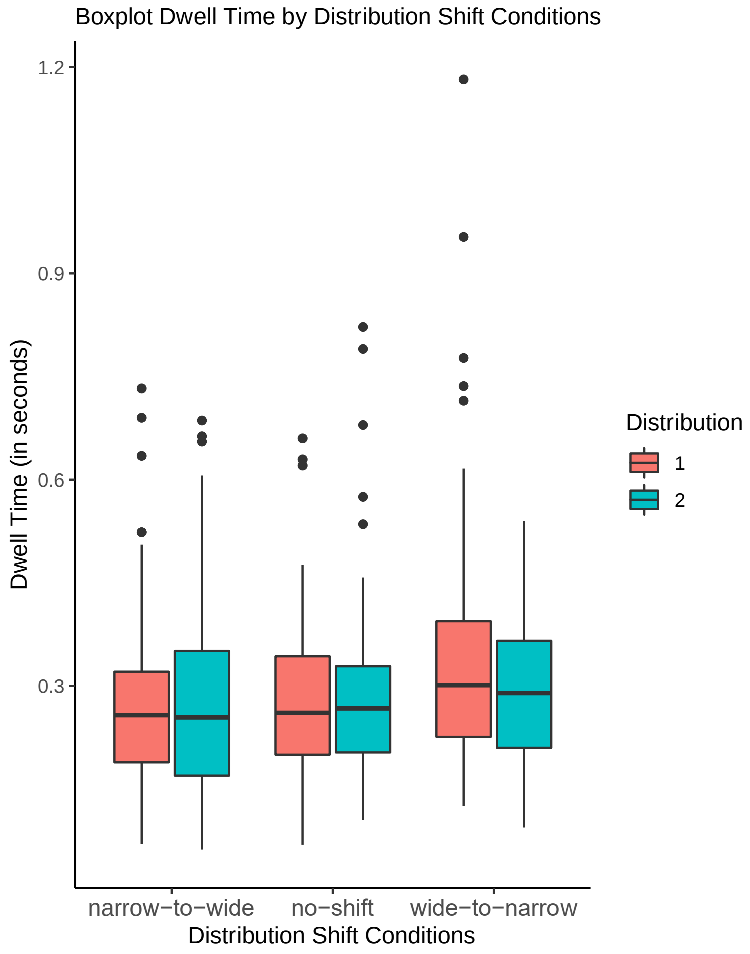

Figure 3. 5. Dwell time by distribution shift conditions. No differences in dwell times (in seconds) were observed for the narrow-to-wide and no-shift conditions. However, in the wide-to-narrow shift condition, shorter dwell times were observed for the "narrow" distribution, when it shifted from the "wide" distribution.

3.3.2 Explicit Reports of Distribution Shift

The median explicit reports of distribution shift (i.e., mouse clicks) that participants reported in this task was *Mdn* = 8, *SD* = 13.61. (Figure 3.6).

Linear mixed-effects models, with a fixed effect of explicit reports, and participants as a random effect of yielded a significantly longer saccadic latencies on trials with reported distribution shift, ($\text{estimate}_{\text{reported}}$= 0.012, SE = 0.005, $t(347) = 2.49$, $p < 0.05$) (Figure 3.7), and yielded a significantly longer dwell times on trials with reports of distribution shift, ($\text{estimate}_{\text{reported}}$= 0.080, SE = 0.014, $t(347) = 5.88$, $p < 0.001$) (Figure 3.8).

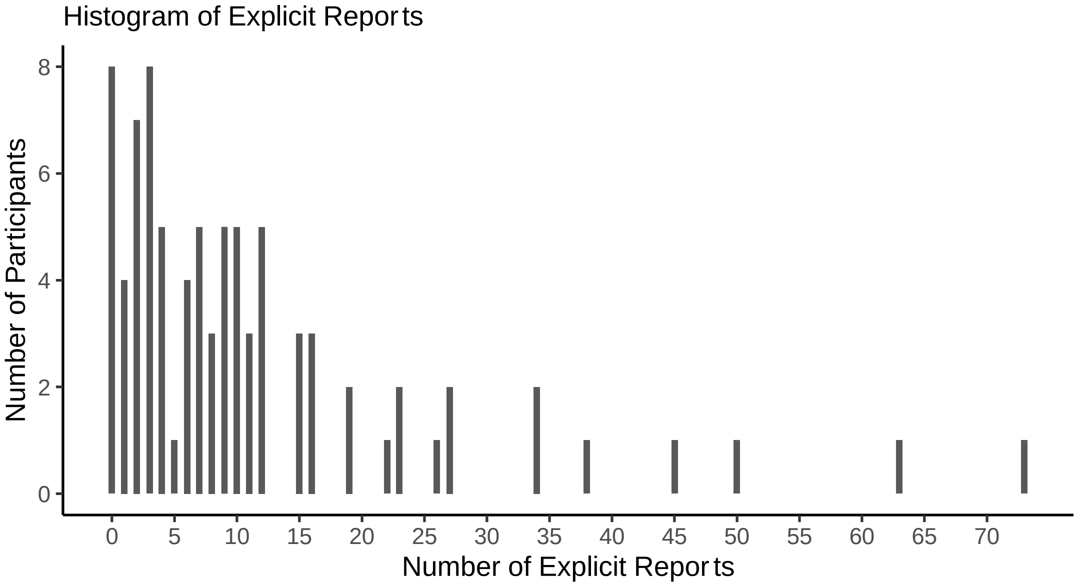

Figure 3. 6. Histogram of number of explicit reports of distribution shift (i.e., via mouse click) that participants made throughout this experiment.

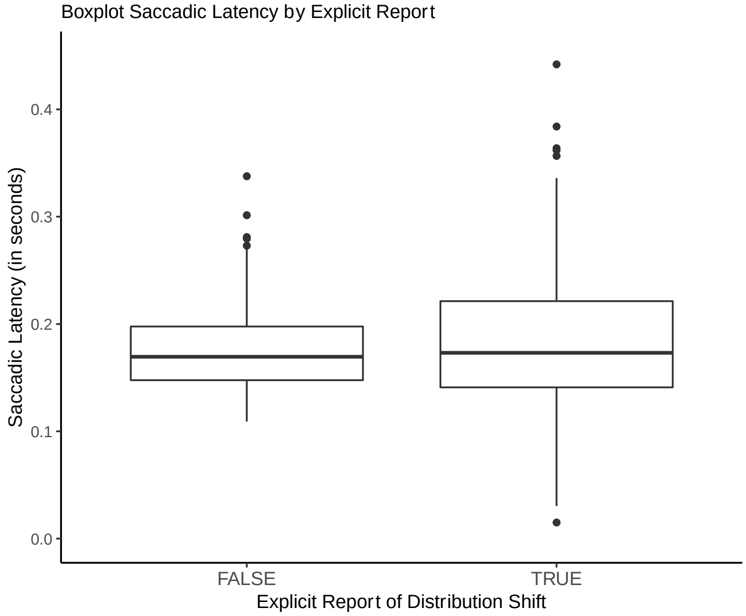

Figure 3. 7. Saccadic latency by explicit report. This box plot displays participants' saccadic latency (in seconds) on trials with reports of distribution shift, compared to trials with no reports in shift. Saccadic latencies were longer on trials with explicit reports of distribution shift.

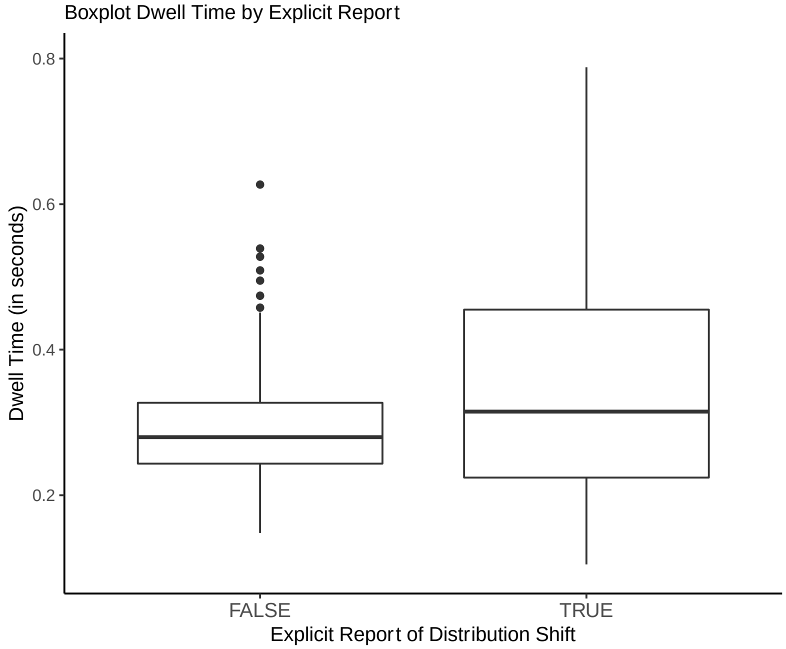

Figure 3. 8. Dwell time by explicit report. Boxplot of dwell time (in seconds) by explicit report of distribution shift condition. Dwell time spent on target fixation was contrasted between the conditions: no report of distribution shift (i.e., no mouse clicks) and explicit report of distribution shift (i.e., via mouse clicks). As seen in this graph, participants spent longer on trials with explicit report of distribution shift.

3.3.3 Uninformative Trials

Surprise

A linear mixed-effects model with interactions between two fixed effects: trial conditions (i.e., updating vs. uninformative trials) and block (i.e., block 1, 2, and 3), and a random effect of participants, predicting saccadic latency revealed a negligible effect, ($\text{estimate}_{\text{uninformative}}$= -0.007, SE = 0.009, *t*(352) = -0.79, *p* = 0.43) (Figure 3.9).

Updating

A linear mixed-effects model with interactions between fixed effects of trial conditions (i.e., updating vs. uninformative trials) and block (i.e., block 1, 2, and 3), and a random effect of participants, predicting dwell time revealed significantly shorter dwell times for uninformative trials, ($\text{estimate}_{\text{uninformative}}$= -0.071, SE = 0.025, *t*(352) = -2.81, $p < 0.01$) (Figure 3.10).

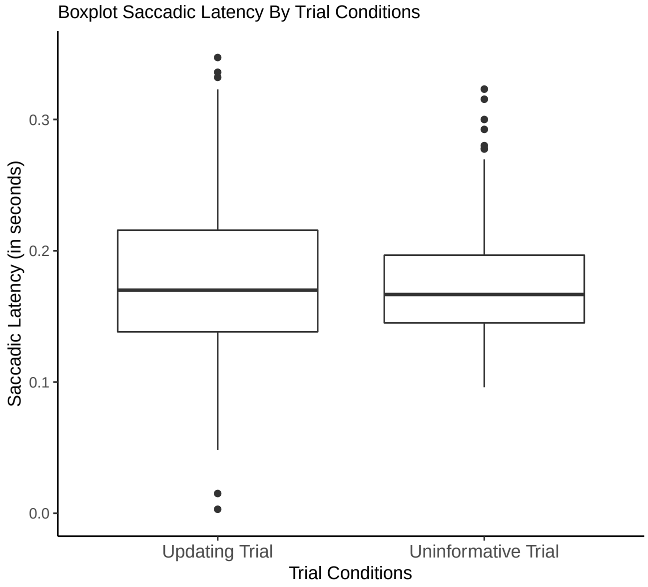

Figure 3. 9. Saccadic latency by trial conditions. Box plot of a comparison in saccadic latencies (in seconds) between uninformative trials (i.e., targets with no predictive value to future events, and the reported distribution shift trials (i.e., via mouse click). No differences in saccadic latencies were observed between the two conditions.

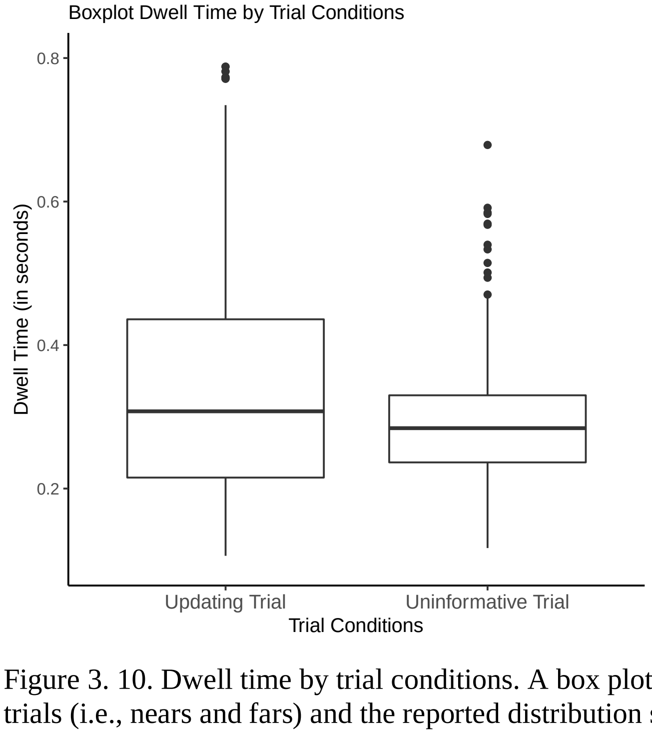

Figure 3. 10. Dwell time by trial conditions. A box plot with a comparison between the uninformative trials (i.e., nears and fars) and the reported distribution shift trials (i.e., update trial). Overall, the participants had shorter dwell times for the uninformative trials.

3.4 Discussion

Participants were poor at determining the true distributional adjustments manipulated in this Experiment, as the median number of explicit reports of distribution shifts in the task was 8. This aligned with the goal of this task, where distribution shifts were intended to be less transparent compared to the first Experiment. Although, there were only two real distribution shifts in this task, it is expected for individuals to construct the initial internal model of the spatial distribution of targets (which may not be perfect), and would need to fine tune or update one's model to capture a more accurate representation of future events. The high variation in the number of explicit reports, with a standard deviation of 13.61 reflects the discrepancies between observers' strategies of how their model captured future target locations. For instance, an observer with a much higher number of reported distribution shifts (e.g., 50) could be interpreted as having a model that predicts a much smaller spatial area of where future target locations can be expected, thus would have to update their model more frequently. Despite having poor explicit representations of when contingencies changed, implicit measures, saccadic latencies and dwell times were analyzed to determine if there were any behavioral costs associated with the distribution shifts.

When distribution shifted from narrow-to-wide, participants had longer saccadic latencies, but no differences in dwell times were observed. This abrupt shift in the distribution lead to targets appearing from novel areas with greater angular displacement from the mean of the previous distribution – which could be interpreted from the observer's point of view as targets appearing from the "extreme" tails of distribution 1. These surprising events were unexpected and low in probability. Although these events were unexpected and surprising, we inferred that no updating of internal model occurred since dwell times did not increase. While longer dwell times were expected to be observed when distribution shifted, it is important to note that surprising events do not necessarily lead to updating. Filipowicz et al., (2018) found that higher levels of surprise lead to poorer performance in updating. Whilst surprising events can provide participants with evidence for a change in the environment, highly surprising events can sometimes be treated as “outliers” (Filipowicz et al., 2018; Summerfield & Tsetsos, 2015).

For the wide-to-narrow shift condition, both saccadic latencies and dwell times were shorter when the distribution shifted from wide-to-narrow. While shorter saccadic latencies were predicted, the observation of shorter dwell times was not expected. Since the distribution shifted from a wide variance to a narrow variance, new targets would appear from high probability locations from the preceding distribution. Thus, these events were high probability under the model and would be well predicted by observers’ mental model. This was reflected by shorter saccadic latencies. However, no conclusion can be drawn from the shorter dwell times. Factors such as outliers – longer dwell times observed before the distribution shift could be driving this effect (see Figure 2.5). These longer dwell times observed prior to the distribution shift could be due to possible updating behaviors before the actual distribution shift had occurred. However, no conclusion can be made.

Lastly, observations from the control condition, no-shift in distributions confirmed our hypothesis, as no difference in saccadic latencies and dwell times were observed. Overall, in these three distribution shift conditions (i.e., narrow-to wide, wide-to-narrow, no-shift) saccadic latencies were responsive to target probabilities – shorter latency of saccades for expected/high probability events and longer saccadic latencies for surprising/low probability events. However, dwell times did not reflect the distribution shifts, as these changes in the distribution may not have been noticed by the observer. To determine that this was the case (i.e., absence of longer dwell time due to observer’s belief that no change has occurred), explicit reports were examined (i.e., observer believed they detected a change in the underlying distribution of target locations).

On trials where individuals reported a distribution shift, longer dwell times and longer saccadic latencies were observed. Longer saccadic latencies could reflect the between-trial processes of the possible need to update, and the behavioral reorienting processes induced by surprising events (O’Reilly et al., 2013). Furthermore, the linear mixed-effects model suggested that longer dwell times were better predicted by the explicit reports of distribution shift (compared to saccadic latencies). Which is consistent with our findings that longer dwell times reflect updating. It can be seen that surprise and updating are strongly associated. To dissociate the effects of surprise from updating, uninformative trials were analyzed.

In line with the hypothesis, participants did not update their model for the uninformative trials (i.e., events with no predictive value to future events) although these events were surprising (i.e., occurred 10% of the trials). Shorter dwell times were observed for uninformative trials when compared to trials where individuals updated. Furthermore, no difference in saccadic latencies were found between uninformative trials and updating trials (i.e., reported shift in distribution). This suggests that uninformative trials are just as surprising as trials where an observer perceived a change in the distribution (i.e., explicit reports). It is important to address that the saccadic amplitudes differed between uninformative trials (i.e., 6.0 and 10.0 degrees of visual angle) and updating trials (i.e., 8.0 degrees of visual angle). Although individuals generated different magnitudes of saccades for these different trial conditions, a study by Darrien and colleagues (2001) reported no difference in saccadic latencies for varying saccadic amplitudes for target displacements. In other words, saccadic latencies are not influenced by the angular distance the eye needs to travel.

Overall, our findings from this Experiment align with the findings from Experiment 1. Although, the participants were poor at determining the true distributional adjustments, saccadic latencies were sensitive to surprising target locations, moreover, dwell times reflect the updating of predictive internal models.

Chapter 4: General Discussion

In this thesis, the aim was to examine the capacity to update mental models when the change in distributions was not explicitly signaled by an accompanying change in perceptual characteristics of the observations to be represented, and whether eye movement metrics could reflect surprise and updating. This study demonstrated that individuals were able to learn spatial distributions of target locations and detect distributional changes, with the absence of explicit signals. However, when the environmental changes were far more subtle, it resulted in poor updating performance. Furthermore, this research demonstrated an important distinction between surprising events that are unexpected and expected, but low probability observations. Although the uninformative events were surprising, they were also expected, as the experimenter informed participants that they would occur. Thus, these expected surprising events did not lead to the updating of one's mental model. By contrast, the target locations following the distribution shifts were surprising and unexpected – a combination that led to the updating of participants' mental models. Supporting the evidence that surprise and updating are distinct constructs.

Across the two studies, there was an increasingly compelling evidence that saccadic latencies are responsive to target location probabilities. On the other hand, dwell times reflect the updating of one's mental model. The first Experiment demonstrated that individuals were able to detect statistical regularities dispersed spatially, where longer saccadic latencies were observed for both expected and unexpected surprising events, whereas longer dwell times were observed for surprising events that were unexpected. This pattern of results suggests that participants were able to build a mental model of the probability distribution to facilitate speeded eye movements towards more probable locations.

The second Experiment provided support of the findings from the first Experiment, as saccadic latencies were responsive to target probabilities and surprise, despite having distribution shift conditions that were very subtle. Additionally, the dissociation between surprise and updating was demonstrated. This was done by contrasting trial conditions that were both equally surprising, however longer dwell times were solely observed for the trials with predictive value to future events. Although the current study provided the dissociation effect with the evidence that surprise does not always lead to updating; the current research was not able to demonstrate the possibility that updating can be triggered with the absence of surprising observations. In other words, longer dwell times with the absence of longer saccadic latencies.

Overall, the current research has demonstrated that saccadic latencies and dwell times can reflect implicit representations of surprise and model updating. This research provides an insight into the way we process information from our environment and make appropriate updates to our model when contingencies change. Although surprise is associated with behavioral costs (e.g., slower saccadic latencies), it provides valuable evidence for a change in the environment, which facilitates the updating of mental models.

References

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of experimental psychology: Learning, memory, and cognition*, *19*(5), 1061.

Carpenter, R. H., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377*(6544), 59-62.

Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLoS biology*, *10*(3), e1001293.

Danckert, J., Stöttinger, E., Quehl, N., & Anderson, B. (2012). Right hemisphere brain damage impairs strategy updating. *Cerebral Cortex*, *22*(12), 2745-2760.

Darrien, J. H., Herd, K., Starling, L. J., Rosenberg, J. R., & Morrison, J. D. (2001). An analysis of the dependence of saccadic latency on target position and target characteristics in human subjects. *BMC neuroscience*, *2*(1), 13.

Filipowicz, A., Valadao, D., Anderson, B., & Danckert, J. (2018). Rejecting Outliers: Surprising Changes Do Not Always Improve Belief Updating. *Decision, 5*(3), 165.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243-18250.

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, *28*(47), 12539-12545.

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. Neuron, 84, 870 – 881. http://dx.doi.org/10.1016/j .neuron.2014.10.013.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience, 30*(37), 12366-12378.

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *110*(38), E3660-E3669.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions?. *Trends in cognitive sciences*, *19*(1), 27-34.

Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts?. *Cognition*, *87*(2), 55-71.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *science, 331*(6022), 1279-1285.

Wu, C. C., Kwon, O. S., & Kowler, E. (2010). Fitts's Law and speed/accuracy trade-offs during sequences of saccades: Implications for strategies of saccadic planning. *Vision research*, *50*(21), 2142-2157.