**Title**

Establishing an experimental control in 3D Simulations to predict human learning using machine learning algorithms [will probably rename this].

**Abstract**

In this thesis, we seek to understand the process of learning and how attention is allocated during and after it. The thesis has three parts. First, we built a 3D naturalistic environment using a game engine called Unity3D. We are able to develop a general-purpose experimental testbed for running structure and attention learning tasks by integrating a novel hardware and software setup. Second, we created a post-analysis data pipeline that classifies eye gazes collected by an eye tracker into sequences of fixations and saccades. In this thesis, we ignore the saccades and detect fixations to differently labeled stimuli in the environment. Last but not least, we quantify the contribution of fixations to different objects and use machine learning algorithms to predict behavioural learning. [Fill out the results].

**Introduction**

In cognitive science and psychology, eye movements are often studied because they help elicit and analyze cognitive processes. To interact with the world, humans shift their eyes towards salient events to select information. When we visually sample the surrounding environment, our eyes tend to move to specific locations of the visual field (saccades) and we rest our eyes on these locations for a few hundred milliseconds (fixations). By controlling these sequences of saccades and fixation, we can determine what information reaches our visual cortex. This alternation of saccades and fixation characterizes visual exploration and it helps understand how humans learn and control their attention. Learning can be in many forms and in this thesis, we seek to understand structure learning, representation learning and learned attention. It is believed that in complex learning scenarios, we must first learn the structure that governs the domain of interest. During structure learning, it is more efficient to only consider a smaller set of stimulus dimensions so that we help maximize the rewards [Need more clarification on this]. This reduction of dimensionality is known as representation learning and it helps enable faster learning. On the other hand, learned attention is the allocation of attention towards the optimal reduced-dimensionality representation for the current environment. In reinforcement learning, first one must find the policy that maximizes the reward for any given state and the reduced representation of the environment and stimuli before one can begin to allocate one’s attention effectively. This process may continue for far longer than the original structure learning process in order to obtain the most optimal representation of dimensions. [Need to clarify this and provide more resources]. In this thesis, we developed a natural 3D environment that can be used to test and measure a plethora of learning-based experiments. [Insert references to gaming environments]. Our testbed is also designed to accommodate other hardware devices such as a ECOG for collecting neural data. We are mainly interested in behavioural events such as eye movements, particularly fixations because in problem solving and learning tasks, they easily yield important clues to human behavior. Recorded sequences of gaze locations can represent behavioural activities at a temporal resolution on the order of 10 milliseconds [Insert references]. With this information, we can study how much time people need to process various stimuli and how often they need to look at them. By further understanding the mechanisms that underlie visual sampling such as observing fixations, we can look deeper into how attention works [Insert references].

**Hypothesis**

It is expected that the duration and counts of different fixation features will be different. Some predictions include subjects will fixate equally on both objects while learning and they will fixate more on the target object after learning. However, the duration of the rewarded object should decrease after learning [Find sources to back this up. Chen et al 2013]. There will also be between subject differences in fixation durations, particularly prior to picking up the object due to different searching styles and attention. Lastly, there should be a correlation between the duration of the time spent fixating the object after reward and learning speed (Watson & Blair, 2008).

Eyetracking studies of category learning show that fixation counts and durations to irrelevant stimuli do not begin to differ from those to relevant stimuli until several trials after category rules have been learned (Blair, Watson, & Meier, 2010; Rehder & Hoffman, 2005).

**Task**

*Task Design*

This is a context-based feature learning task. Depending on the context, particular stimulus feature is relevant and these stimuli are presented as objects. A sequence of trials is defined as a block and it dictates the ruleset that determines which object is correct or incorrect. To initiate a trial, the subjects need to fixate on the fixation keyhole for xxx milliseconds. The door then opens and exposes the arena that is to be explored. In the open area, there are two objects with different combinations of feature values and a specific floor texture that is the context. To complete a trial, the subject needs to use the joystick to navigate and pick up an object. The selected object is the one that is brought to the reward door which then ends the current trial and starts the next one.

**Materials and Methods**

Materials.

*Game Setup - Hardware.*

The game is integrated with various hardware components including a Tobii TX300 Eyetracker, a Logitech joystick (need the model), an Arduino (need the model), a pair of photodiodes, a display monitor, a host computer.

The Tobii Eyetracker allows head free and non-invasive eye tracking at 300 hertz and it comes with a software development toolkit (SDK) that provides the API for using the eye tracker functionalities such as calibration and illumination. The eye tracker records various information such as left and right eye coordinates (X, Y, Z) and pupil sizes. The eye tracker has its own internal coordinate system, [insert details… ADCS, track box, distance, illumination modes]. The joystick is used for navigating throughout the game, and is connected via USB. It outputs two coordinates, X and Y. The photodiodes are clipped on to the top two corners of the experiment monitor for timing analysis. They pick up the two flashing black and white panels which are switching alternatively multiple times per frame.

[Insert diagram of hardware setup]

*Game Setup - Software.*

The software components include a Python GUI app that interfaces with the eye tracker and transfers data to the main game client created with Unity3D. The app also displays the subject's last calibration and eye positions in real time. This is shown only on the experimenter's computer.

[Insert diagram of software setup]

*The Environment.*

The main scene of the game is a large square arena enclosed by walls. Each of the four walls has a distinct landmark placed beyond it to aid with navigation. In this experiment, there are two main objects that are placed randomly on the floor where one is the rewarded target and the other one the unrewarded distractor.

There can be other kinds of objects but they are never-relevant distractors.

[Insert diagram of environments]

*Object features.*

The object that appears in the arena is called the multi-part geons. These geons are defined by four features (body shape, body color, body texture and arm width) each with two possible values (e.g. the body shape can be rectangular prism or a cylinder), making 16 possible objects.

[Insert diagram of all 16 objects or maybe partial set]

*Context.*

The context is the floor of the arena which can take on x possible states. The floor indicates which particular object feature is the relevant and which particular value of the feature is rewarded. For example, for a given trial, an icy floor might indicate that objects with thick arms would be rewdared. During each block of trials, each trial would come from one of two contexts, so the main task is to learn which feature and which of its values are associated with each context.

[Insert diagram of contexts]

*Replayer.*

The replayer identifies the objects that are overlaid by the coordinates of the fixations [Expand on the post-analysis and replayer].

Methods*.*

*Participants.*

Eleven subjects (4 females and 8 males; age range from X to X; Tyler, Kenneth, Alisa, Kevin, Sophie, Vincent, Abrar, Andres, Tom, Ellis, Robert, Patrick) were recruited from York University)

*Procedure.*

Some subjects wore glasses and some didn’t, which affected the accuracy of the eye tracker. Some subjects required interpolation of their data, which will be discussed later. First, subjects were asked to sit on the a 50 to 60 centimeters away from the monitor at eye level. The game started with calibration, and it was only repeated if the calibration was not good enough. After the calibration, the tutorial started, which was designed to help the subjects to get familiar with the environment and the overall task. After the tutorial, each subject was asked if they had any questions. If so, they were answered. The main task then started, and was played for an average of xx minutes and played on average with xx trials per subject.

*Calibration.*

Subject's eyes are calibrated with a nine points calibration procedure to ensure accuracy in the gazes.

*Task.*

To initiate the trial, the player is put into an intertrial room where there is a fixation keyhole that needs to be looked at for 500 ms to open the door to the square arena. In the arena, two main objects are put randomly on the floor and the player can pick up objects by navigating close to it. After the player picks up a selected object, the player can bring it to any black reward door with a white asterisk shape on it. The door will open and the asterisk will turn into a green "+" if it is the correct object or a red "X" if it is the incorrect object. Shortly, the player is teleported to the small intertrial room again and the entire process is repeated. The task of the player is to learn to associate the value of specific object features to one of the two contexts. In each block, the two contexts are different. This context-depended feature selection is learned through a sequence of blocks. Blocks effectively train the player to generalize the relevant feature irrespective of the exemplar object. At the end of each block, we expose the player to trials in which context one and two are randomly interleaved.

**Data Analysis**

*Timing Analysis*

We implemented various methods to ensure the timing of the data collection is precise and consistent. First, the photodiodes tell us if the monitor frames are being refreshed without any latency. Second, UNIX timestamps are attached to every packet of data transferred by using a low-latency and loss-tolerating communication protocols called User Datagram Protocol (UDP).

*Eye movement data analysis.*

To identify fixations in the visual scene, we used a velocity threshold algorithm similar to Salvucci and Goldberg with [Insert the parameters].

*Detecting fixations to differently labeled stimuli in the environment.*

Next, we characterize fixation regions based on labeling visual scene features. We've defined the following features according to whether the fixations arrived at:

F1-REWOBJ rewarded object (target)

F2-NREOBJ nonrewarded object (distractor)

F3-HAND hand

F4-IRROBJ irrelevant object (never rewarded)

F5-REWDOO reward-door

F6-REWKEY reward door keyhole

F7-INSWAL inside-box wall

F8-INSGRO inside-box ground

F9-OUTBOX outside-box

F10-PROGRES progress bar

F11-OTH other (e.g. outside of monitor)

Since the experiment is a dynamic viewing instead of a traditional static-viewing task, the fixation will miss some information access.

*Analyzing Learning*

We used a Matlab-based software that quantitatively characterizes subjects’ learning through the analysis of behavioural response. We analyze individual subject and estimate a learning curve that characterizes the probability of a correct response as a function of trial number. This is implemented by using a state-space model and binomial expectation and maximization algorithm as outlined in the work of (Smith et al, 2004).

*Predicting behavior learning*

A feature matrix is created to be used as the training data for the machine learning algorithm. The matrix is composed of four independent variables (target fixation mean duration, target fixation count, distractor fixation mean duration, and distractor fixation count). The dependent variable that is being predicted is the EM learning curve. To obtain the EM learning curve, we used the EM algorithm [insert EM paper here], where we feed in all the trial accuracy of each block.

Our machine learning model uses linear regression where the training set is split with cross validation. This ensures we do not overfit the training data. [insert cross validation techniques]

We decided to use Linear regression because it allows us to compared the power and the correlation between the features and the predicted value.

*Data preprocessing*

First, we normalized the independent variables to ensure they are on the same scale.

*Data pipeline*

The raw gaze data and the frame data from all the trials are computed to obtain the classification of fixation and saccades. These classifications are then feed into the replayer and the replayer produces a table of data including the object that the fixation is fixated on and the duration of fixations. Separately, the trial outcome is fed into an EM algorithm which produces the EM for each trial. Together, the fixation characteristics and the trial EM are used for the machine learning models.

**Results and Discussion**

[Insert results]

*General Discussion.*

While eye movements certainly do not entirely reveal a person’s thoughts, their flexibility and informativeness make them an excellent data source for many studies and applications (Salvucci 1999).

It is generally agreed that qualitatively different cognitive states can be accompanied by qualitatively different patterns of attention. That is, in different circumstances, we pay attention to different things for different reasons. Recent work has suggested that the exploration and exploitation modes of action found in reinforcement learning models (cf. Sutton & Barto, 1998) may have corresponding attentional states. It is hypothesized that during exploration periods, agents will attend to the stimuli whose associated outcomes are most uncertain, in accordance with the attentional model originally proposed by Pearce and Hall (1980), while precisely the opposite is true during exploitation periods, in which agents attend to the stimuli which are most strongly predictive of their outcome, in accordance with the model originally proposed by Mackintosh (1975). Gottlieb (2012) refers to these two attentional states as attention for learning and attention for action, respectively.

*Future Direction.*

Due to the shortage of time, it would have been better if the sample size is more than 20 – 30. Also, some subjects were exposed to different sets of objects, which may influence stimulus discrimination. Also, some subjects had different length of test blocks (ask Marcus about this). [Talk about fixation saliency]

**Bibliography**

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely selective attention: Eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization.*Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1196-1206. doi:http://dx.doi.org/10.1037/a0016272

Salvucci, D. D. (1999). Mapping Eye movements to Cognitive Processes. In Proceedings of the Twentieth Annual Conference of the Cognitive Science Society, 923–928.

Smith, A. C., Frank, L. M., Wirth, S., Yanike, M., Hu, D., Kubota, Y., … Brown, E. N. (2004). Dynamic Analysis of Learning in Behavioral Experiments. *The Journal of Neuroscience*, *24*(2), 447 LP-461. Retrieved from http://www.jneurosci.org/content/24/2/447.abstract