Exploring Normative Eye Movement Patterns in Functional Tasks

by

Ewen B. Lavoie

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Faculty of Physical Education and Recreation
The University of Alberta

Abstract

When interacting with an object, humans are quite effective at navigating their hand to an object, grasping it, and acting on it. The level of ease with which we do this masks the complex interplay of sensory modalities that is occurring. This study utilizes a head-mounted eye-tracker and upper-limb motion capture markers to reveal how one of these sensory modalities, vision, enables efficient object interaction. Participants completed several trials of two tasks mimicking real-world demands. The first task involved turning and grasping a pasta box from an original position outside the participant's field of view and placing it onto two shelves before returning it to its starting location. The second task had participants move cups filled with beads four times over a partition. Both tasks show participants spend nearly the full duration of the trial fixating on objects relevant to the task, well in advance of their hand arriving at an object. As well, participants spend little time fixating on their own hand when reaching towards an object, and slightly more time, although still very little, fixating on the object in their hand when transporting it. Instead, during a grasp, participants make a saccade from the object to its drop-off location, and hold this fixation until the object is being released by the hand. Other sensory systems, likely proprioception and haptic feedback, allow participants to behave this way. When interacting with an object outside the field of view, slight changes in this behavior occur. Specifically, participants are unable to fixate on the object as far in advance of their hand, move slightly slower, and increase their maximum grip aperture. A possible explanation for these behaviours is a predictable interaction between covert and overt attention, Dorsal and Ventral Streams of visual processing, and proprioceptive and haptic feedback that allow individuals to carry out object interactions in a smooth, cyclical manner with the eyes leading the hand.

Preface

Some of the research conducted for this thesis occurred as part of the Hand Proprioception and Touch Interfaces (HAPTIX) project sponsored by the Defense Advanced Research Projects Agency (DARPA) BTO under the auspices of Dr. Doug Weber through the DARPA Contracts Management Office Grant/Contract No. N66001-15-C-4015. Dr. Jacqueline Hebert is the lead investigator at the University of Alberta, with co-investigators: Drs. Craig Chapman, Patrick Pilarski, and Albert Vette. The literature review in chapter 1, analysis of results in chapter 3, and discussion points in chapter 4 are my original work. The technical apparatus and task design in chapter 2 were made in collaboration with Drs. Hebert, Chapman, Pilarski, and Vette, with the subsequent data collection and analysis being carried out by myself.

All procedures were approved by the University of Alberta Health Research Ethics Board, the Department of the Navy Human Research Protection Program (DON-HRPP) and SSC-Pacific Human Research Protection Office (SSCPAC HRPO).

This thesis is an original work by Ewen B. Lavoie. No part of this thesis has been previously published.

Acknowledgements

Beginning graduate school after a hiatus from academics was a difficult decision to make, and would not have been possible without the efforts of many people along the way. I could not be more grateful for the opportunity that Dr. Jacqueline Hebert provided for me to work on the HAPTIX project. The balance with which Dr. Hebert challenged me, while also offering advice and guidance, will stick with me. I feel equal levels of gratitude for the opportunity Dr. Craig Chapman provided for me to be a part of the ACELab and work on a project that was new and exciting. I understand not all graduate students have such an opportunity laid before them. Along with being patient, as I took time to answer questions I'm sure he already knew the answer to, Dr. Chapman provided daily support at the knock of a door (usually when D-Lab played tricks on me), encouragement when it was needed (looking at pupils all day for weeks will necessitate the need for this), and guidance in logic during discussions. As well as being an excellent mentor, Dr. Chapman had no problem beating me at video games (except Mario Tennis), inviting me to his family's beautiful ocean-view cabin in the Washington wilderness, and conversing with me about space travel, science fiction novels, and hockey trades. All this and more makes working with, and learning from Dr. Chapman enjoyable, challenging, and formative, and makes the ACELab the best place to continue my graduate work as a PhD student.

I also wish to acknowledge the financial assistance provided to me by Drs. Jacqueline Hebert and Craig Chapman, the Faculty of Physical Education and Recreation, the University of Alberta, the Albert Government, and the Natural Sciences and Engineering Research Council of Canada, without which this work would not have been possible.

Aida Valevicius and Quinn Boser deserve special recognition for the work they put in planning, organizing, and conducting the data collection for these experiments. Aida, your optimism in the face of setbacks was much needed, and likely underappreciated in the moment. Maybe my hips will allow me to go on some short runs soon? Quinn, when something went wrong, you usually knew how to fix it. Can you could provide me with a list of rollercoasters I should ride?
The rest of the HAPTIX team at the U of A, Dr. Patrick Pilarski (for computer wizard things), Dr. Albert Vette (for biomechanic wizard things), John Luu, Ognjen Kovic, Riley Dawson, Tarvo Kuus, McNiel Keri, and Brody Kalwajtys, also need to be recognized for being great people to be around, and being patient when eye-tracking data was still not analyzed. As well, my fellow ACELab graduate students deserve to be credited for putting up with my opinionated rants in our shared office. Jenn, are you still in the lab? You're never around. Nathan, do you do any work, or just look at memes? Jeff, you're a first-liner in my books, no matter what Craig says. Some of the other inhabitants of the 4th floor of the Van Vliet Complex who have helped me along the way with early "good mornings" and impromptu discussions about science and life are Drs. Brian Maraj, Kelvin Jones, and Dave Collins.

I can pursue academics because I stand on the shoulders of giants. My giants are not aged scientists, but rather my remarkable family. I have been blessed with four grandparents who taught, and still teach me to have faith, persevere, put family first, and work hard, not for recognition, but because it is the right thing to do. For my entire life, my parents, Mitch and Cecilia, have provided me with all the love and support necessary for me to succeed in any arena of my choosing. The only struggle is living up to the expectations that I set for myself based on

the exemplary model that is their lives. My brother Malcolm has shown me what sustained, focused exertion and commitment to long-term goals can achieve in one's life. My brother Callum has become a role model of mine for his ability to persevere in the face of hardship and do the right thing when lesser men would falter. And, my sister Maura reminds me to seek out challenges for myself and trust my abilities to conquer them, all while having a bit of fun. I am indebted to you all more than you know.

Last, but certainly not least, I am immensely grateful for the love and support Ashley has given me. Ashley, you have been a calming presence in my often-turbulent life, and helped keep me sane while I was re-re-re-re-analyzing data. Thank you.

Table of Contents

List of Tables

List of Figures

Glossary of Terms

°: Degree
>: Greater than
”: Inch
<: Less than
#: Number
%: Percent
p: Probability value
2D: Two-dimensional
3D: Three-dimensional
ANOVA: Analysis of variance
AOI: Area of interest
C: Cup transfer
cm: Centimetre
DOF: Degrees of freedom
DON-HRPP: Department of the Navy Human Research Protection Program
HD: High definition
Hz: Hertz
IR: Infrared
mm: Millimetre
mm/s: Millimetres per second
ms: Millisecond
ORA: Object related action
oz: Ounce
PB: Pasta box transfer
QR: Quick response
ROI: Region of interest
SSCPAC HRPO: Space and Naval Warfare Systems Center Pacific Human Research Protection Office
s: second
V1: Primary visual cortex

1 - Introduction

Most ancient human civilizations had a symbolic fixation with the eye. The Greeks talked of the Evil Eye, while the Egyptians scribbled the Eye of Horus on the walls of the Pyramids. The eyes have been called any number of poetic names including, but not limited to, “the windows of the mind”, “nature’s looking glass”, “the lamp of the body”, and “the window to the soul”. However, this obsession with the eye is not limited to ancient civilizations, religious texts, or proverbs, as popular music of today is littered with references to “lyin’ eyes” (Frey & Henley, 1975), “angel eyes” (Anderson & Ulvaeus, 1979), “green eyes” (Martin, 2002), “father’s eyes” (Clapton, 1998), and even “blue eyes crying in the rain” (Rose, 1975). As one could guess from just these song titles, the eye is broadly referenced in popular culture and is used in many different circumstances. But, I would argue that the breadth of the scholarly research dedicated to the eye is greater than that of all the songs ever sung. In continuation of this tradition, this thesis is primarily concerned with eye movements, specifically those that occur during natural object interaction tasks, and I will begin with a brief overview of the first research techniques used on the eye and interesting observations made about the eye and its primitive movements.

1.1 – A brief history of eye and eye movement research

The ancient curiosity with the eye fueled scientific experimentation mostly in the fields of anatomy and physiology in the early days of the scientific revolution. Du Laurens mapped the paths of the optic nerves (1599); Kepler described how the lens of the eye focuses light onto the retina creating an image (1611); Scheiner refined the anatomical model of the eye and showed the formation of a retinal image on the excised eye of an ox (1630); Descartes developed an improved artificial eye which helped show its image-inverting properties (1637/1902).

Afterimage experiments were employed by Wells (1792), Ruete (1845,1857), Helmholtz (1867), and Wundt (1862) which helped illuminate the coordinating abilities of the paired eye muscles that enable both eyes to simultaneously fixate on a desired target. Hermann von Helmholtz wrote of the importance of eye movements as they "play an essential role in the formation of the perception of space" (1867), but due to the lack of technological prowess, the focus of research was on the eye after a movement had taken place, and not on the movement itself.

How does the eye actually move in a human being? One early account by John Hunter, a surgeon, described three types of eye movements: "the eye moving from one fixed point to another,…the eye moving along with an object in motion,… [and] the eye keeping its axis to an object, although the whole eye, and the head… are in motion" (Hunter, 1786, pp.209-210). These three types of eye movements have since been defined as saccades, smooth pursuits, and stabilized fixations, respectively. Observational analyses of eye movements are the oldest of their kind, but a conundrum arises when searching for the limitations of 'X' whilst employing the function of 'X'. Human eye movements have thus been recorded using several different techniques.

1.1.1 – Evolution of eye movement recording techniques

Attaching devices directly to the surface of the eye to record eye movements can be traced back to the late nineteenth century. Orschansky (1899) attached an aluminum cup with a mirror to the surface of the eye deflecting a beam of light to land and mark a photographic film, thus recording the position of the eye, while Dohlman (1925) modified this design to use a rubber eye cup. Some of the important discoveries using this technique are thanks to Alfred Yarbus. Yarbus revealed the need for the retinal image to move often to ensure proper visual system functioning, which explains the three tiny eye movements he observed during fixation on a stationary object: small involuntary saccades, drift, and tremor (1967). These movements, along with blinks and head movements keep the retinal image moving and ensure an empty field does not form. Additionally, Yarbus showed that the eye makes saccades to change the direction of the most developed area of the retina, the fovea. He also revealed that the speed of the saccade changes with the amplitude of the eye movement, is similar for different individuals, and cannot be altered consciously. Yarbus found that smooth pursuits can only occur when there is an object to fixate on in the visual field, and that an individual can alter the speed and direction of smooth pursuits in a few ways, depending on how the velocity of the object changes. Perhaps Yarbus' most impactful contribution to the literature was his confirmation that the objective of the observer dictates the pattern of fixations on a scene. In other words, depending on what the goal is, an observer will fixate on different areas of a scene and for different periods of time. Furthermore, observers show slightly different patterns of fixation on a scene when given the same objective (Yarbus, 1967). Figure 1 (a, b, c) shows one of several suction cup mirror models used by Yarbus, the way that participants' eyelids were held back during testing, and the recording apparatus requiring participants to place their chin in a chinrest (Yarbus, 1967).

The most accurate of the attachment eye movement recording systems use a scleral search coil embedded in a fitted contact lens (Collewijn et al., 1975; Findlay, 1997; Steinman & Collewijn, 1980). As the eye moves, the potential difference between the coils fluctuates, yielding a precise measure of the location of the eye in the head. In this set-up, a participant's head is in a fixed position surrounded by a metal cage containing magnetic-field producing coils. Figure 1 (d, e) shows a scleral search coil in a subject's left eye, and the corresponding metal cage where the testing would take place.

An arguably less invasive form of recording eye movements does not require the attachment of any devices to the surface of the eye to deflect light, but relies on the reflection of light directly from the surface of the eye. This is, without a doubt, the most common eye movement recording method used today. In 1901, Dodge and Cline invented the first corneal reflection eye-tracking device when they discovered that photographing the eye failed to provide sufficient contrast between the iris and pupil (1901). The recording device, the 'Dodge Photochronograph', shown in Figure 1 (f), sits in front of the user, making some tasks difficult for participants to accomplish. The most widely cited experiments of this recording type were carried out by Guy Buswell who used a modified version of the Dodge Photochronograph (Buswell, 1920; 1935). A series of mirrors were installed on the apparatus to deflect a beam of light that had been deflected off the cornea of a subject to a camera at a right angle to the participant. This allowed the participant to have a clear field of view during the experiment. Figure 1 (g) taken from Buswell's 1935 book, *How People Look at Pictures: A Study of the Psychology of Perception in Art,* shows a picture of this apparatus. He used this tool to study eye movements during reading aloud (Buswell, 1920), and viewing art (Buswell, 1935). The latter study found there are some features in images that most individuals fixate on during free-viewing, but that there are obvious differences from one individual to another (Buswell, 1935). And, when participants are given a specific objective, eye movement recordings reflect this substantially (Buswell, 1935). This idea of 'top-down' control was shown again by Yarbus in 1967, and will be discussed in more detail later.

The above eye movement recording techniques have required participants to be in relatively unnatural positions. Some require participants to have their heads clamped, while others involve attaching objects to the surface of the eye or the skin around the eye. One other technique, developed in 1974 by Merchant and colleagues, keeps the eye monitoring and illumination equipment much further away from the participant (Merchant et al., 1974). These types of remote eye-trackers have been used to study human-computer interaction (Farid et al., 2002) and driving (Wilkie and Wann, 2003), often allowing participants to shift their bodies and heads around in the capture space, but usually requiring participants to maintain head position in a chinrest. Remote eye-trackers rely on corneal reflection, pupil tracking, or both, and can give experimenters more freedom when designing studies. Figure 1 (h) shows Merchant and colleagues' remote eye-tracking set-up (Merchant et al., 1974).

Finally, the eye-tracking technique that we employ in our experiments, and what seems to be the direction that many researchers are leaning, is the head-mounted eye-tracker. These were first used in the early 90's, although not widely, and have seen a massive leap forward in comfort and mobility since then. Figure 1 (i, j) shows the Dikablis 2.0 Professional model which sports one infrared pupil camera for each eye, a centrally-mounted scene camera, a plastic nose-piece like that of a pair of glasses, and an elastic strap to hold the device steady.

As well, several models have been designed with wireless options, to enable more mobility to users and experimenters. These generally involve the user wearing a backpack to carry a laptop and power source. Although not as stable in terms of delivering the HD video at a sufficient and consistent frame rate, this technique currently offers the most freedom for researchers in experimental design.

1.1.2 – Summary of major findings from early eye movement research

Previous literature has shown that the eyes do three basic actions: saccades, smooth pursuits, and stabilized fixations (Hunter, 1786), but during stabilized fixations the eyes make small involuntary saccades, drifts, and tremors (Yarbus, 1967). As well, it was shown that the eye movement patterns during free viewing of pictures can differ substantially between individuals, but when participants are given a specific goal to accomplish, eye movement patterns of individuals tend to become more similar (Buswell, 1935; Yarbus 1967). This shows that humans have some control over the direction of fixation of the eyes, but what is really driving the eyes to make saccades? There are two competing schools of thought. The 'bottom-up' model states that eye movements are primarily driven by the characteristics of the image on the retina, and are not driven by executive functions (Land & Tatler, 2009). This would line up quite closely with the free-viewing conditions in the studies of Buswell (1935) and Yarbus (1967). The 'top-down' model argues that the eyes move based on the goals of the current task, and are therefore strategically deployed, and not just a function of the image hitting the retina (Land & Tatler, 2009). This fits with the participants' behaviour during the goal-directed tasks of the Buswell (1935) and Yarbus (1967) experiments. In all likelihood, the direction of saccades is determined by a combination of these two models, and varies between and within individuals based on numerous factors including, but not limited to colour, shape, and edge density of an image, the goals to be accomplished by the eye movements, and the informativeness of an area of the visual field. These will be discussed in the next section.

In addition to the push and pull from bottom-up and top-down factors, our eyes are also constantly responding to information that is at the current location of fixation (falling on the fovea) as well as information arriving from the periphery. But, how robust is our ability to discriminate information in our peripheral vision, and how does this impact the way in which we interact with objects? Posner defined 'overt attention' to be a shift in attention that is accompanied with a saccade of the eyes, while 'covert attention' is a shift in attention without the shifting of the eyes (Posner, 1980). I exhibit overt attention when my telephone rings, and I locate it with my eyes to pick it up. Overt attention occurs when I fixate on the object by moving my eyes so that an image of the object lands on the fovea of my retinas. In contrast, I demonstrate covert attention when my telephone rings and I keep my eyes fixated on my computer screen while I reach towards the telephone to pick it up. A shift of covert attention occurs when I attend to the phone without shifting my fovea to it. Klein argues that these two phenomena are completely independent of one another (Klein, 1980), whereas Remington believes that an object may attract both a subject's covert and overt attention, but these processes are governed by independent processes (Remington, 1980). These are all important questions which will be discussed with the details of some definitions and recent research findings below.

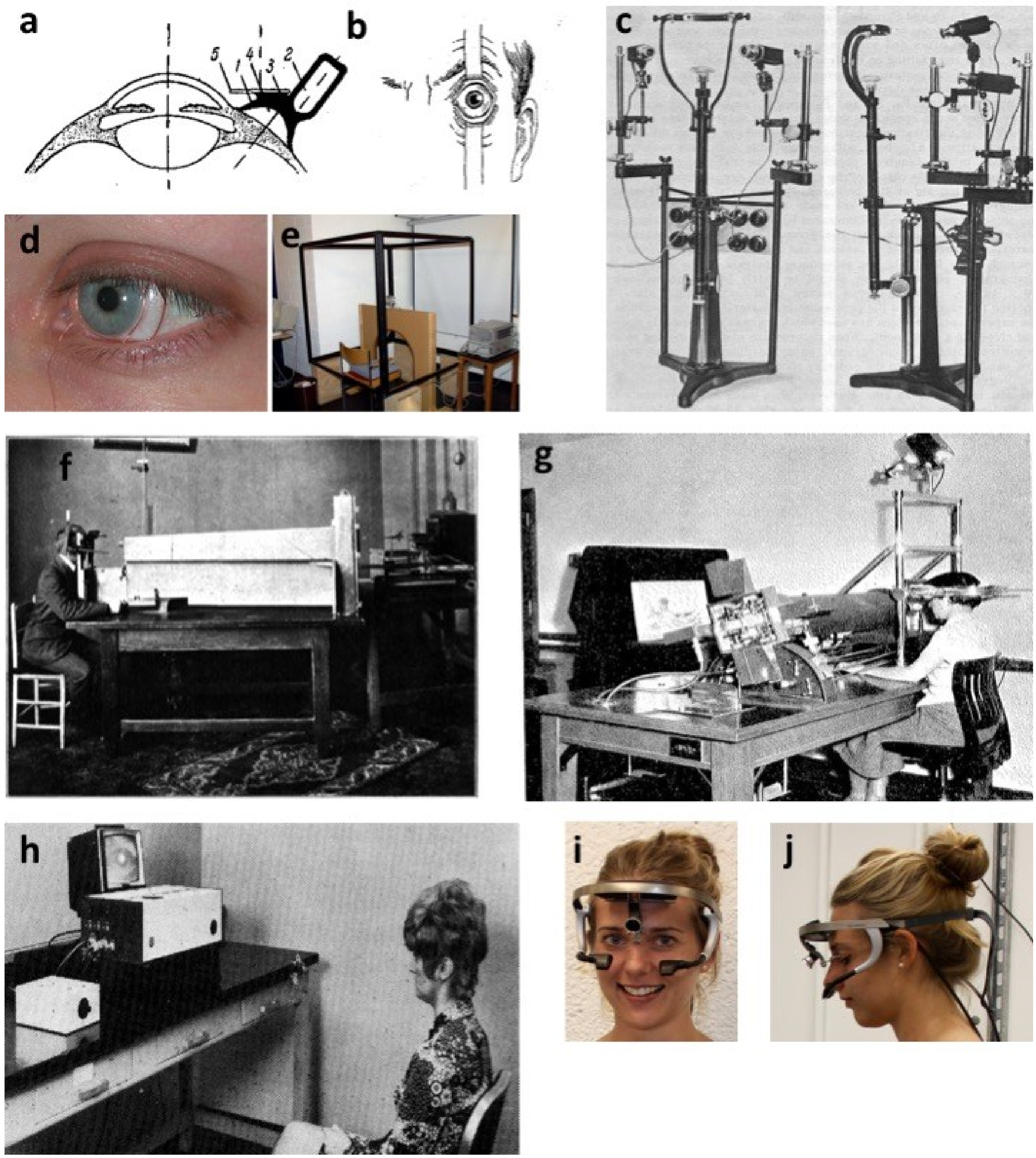

Figure 1: a) A suction cup mirror, b) an eyelid clamp, and c) a recording apparatus used by Alfred Yarbus (1967). d) A scleral search coil in a subject's left eye and e) the seated testing cage with magnetic-field producing coils. f) The Dodge Photochronograph recorded eye movements from seated, head-fixed participants (Diefendorf & Dodge, 1908). g) A participant sits at a right angle to the recording apparatus (Buswell, 1935). h) A remote eye-tracking set-up allows a participant to shift her head within a cubic foot volume space (Merchant et al., 1974). A head-mounted eye-tracker from i) the front and j) the side.

1.2 – Eyes and their role in attention

If we look more in depth at the distinction between the 'bottom-up, salience' model of visual fixation patterns and the 'top-down, goal-directed' model, we can gain a greater understanding of how the eyes function in the allocation of a person's covert and overt attention.

1.2.1 – Bottom-up: Eye movements driven by environmental salience

Most of the evidence for the 'bottom-up, salience driven' eye movement mechanism comes from visual search studies where participants are asked to locate visual objects on a screen in the midst of distractor objects. 'Bottom-up' refers to the inability of the user to control this mechanism. Building on Buswell (1935) and Yarbus' (1967) findings on free-viewing an image, it has been found that visual attention is involuntarily allocated to certain features of an image like colour, shape, edge density, contrast, and orientation (Krieger et al., 2000; Itti & Koch, 2001). As well, socially relevant stimuli have been found to cause involuntary allocation of visual attention. Laidlaw and co-authors showed that participants are unable to avoid fixating on the eyes of humans in portrait photos when asked to by an experimenter (2012). In fact, it has been proposed by Itti and Koch (2001), that, in humans, an image is subconsciously broken down into 'saliency maps' of different features like colour, intensity, contrast, orientation, direction and velocity of motion that are then processed simultaneously by the brain. These 'saliency maps' are reformed to yield an overall 'master' map of salience which directs visual attention to the most salient areas. In a 'winner-take-all' strategy, visual fixation will scan to the most salient areas of the image (Itti & Koch, 2001), while avoiding previously fixated areas through an 'inhibition of return' strategy (Klein & Hilchey, 2011). Itti and Koch created a computational model that was able to mimic both human and monkey visual image search behaviour (2000). Figure 2 (a) shows Itti and Koch's saliency map portrayal of 'bottom-up' visual attention in greater detail. Parkhurst and colleagues tested this model as well and found similar results concluding that "attention is indeed guided by stimulus-driven, bottom-up mechanisms under natural viewing conditions, even when top-down mechanisms are presumably operating" (2002).

1.2.2 – Top-down: Eye movements driven by task demands and participant goals

Itti and Koch state that a bottom-up mechanism might be able to account for the initial attentional division of an image, but that there must be a complementary top-down mechanism that plays a role in attentional selection as well. This is portrayed in the bottom-right corner of Figure 2 (a), which has been replicated from Itti and Koch (2000). There are several models for a complementary type of top-down mechanism (Rybak et al., 1998; Schill et al., 2001; Deco & Zihl, 2001). However, others argue that top-down control of eye movements is much more than just an 'add-on' to the primary bottom-up control. Stark and Choi state that, based on the illusion of our entire field of view being in great clarity even though we only really have detailed information from the tiny sliver of the visual world that hits our fovea, what we believe we are seeing is only loosely connected to what is actually hitting our retina (1996). And, our brain fills in information to make us believe that what we are seeing is clear. This 'fill-in' information is based on what we are expecting to see and forms a cognitive model which informs our eye movements when analyzing a scene. As well, Henderson and colleagues carried out studies similar to those of Itti and Koch (2000) and Parkhurst et al. (2002), and completely refute their bottom-up visual saliency computational models, as they found areas that were fixated were found to be more meaningful to participants than areas that were not (2007). "Any observed correlations between fixation locations and image statistics could be due to the informativeness of fixated locations rather than to differences in the image statistics themselves" (Henderson et al., 2007). In other words, the areas that were found to be highly salient were also areas of meaning for the participants. This was echoed in the world of socially relevant stimuli by Birmingham and colleagues who showed that the saliency map model performed poorly at predicting the fixation locations of participants on photos of human faces (2009).

Whatever the relationship between bottom-up and top-down mechanisms, both have been shown to exist, and affect one another. An arguably more accurate representation of the relationship between bottom-up and top-down visual control is shown in Figure 2 (b), which has been reproduced from Itti and Koch (2001) having been originally published in Schill et al. (2001). Referring back to Yarbus (1967), we can see that the free-viewing behaviour of individuals differed significantly from task-directed viewing, and, eye movement patterns differed significantly between tasks on the same image. Figure 3 shows how one individual's eye movements differed depending on the task given to them by the experimenter. Seven viewing sessions were conducted of the image shown in Figure 3 (a), each three minutes long, and each with a different goal given by the experimenter. When free-viewing the picture (control condition), the participant fixated on a number of areas, but was biased towards the faces of the people, shown in Figure 3 (b); when asked to "estimate the material circumstances of the family", the participant fixated more on the clothes worn by the people and items in the room, shown in Figure 3 (c); when asked to "give the ages of the people", the participant fixated almost exclusively on the faces of the people, with few saccades from one location to another, shown in Figure 3 (d); when asked to "surmise what the family had been doing before the arrival of the unexpected visitor", the participant made fixations to different locations on each person in the picture, shown in Figure 3 (e); when asked to "remember the clothes worn by the people", the participant made fixations mainly to the bodies and faces of the people in the picture, shown in Figure 3 (f); when asked to "remember positions of people and objects in the room", the participant made many fixations to the objects and people in the room, with little regard for the faces of the people, shown in Figure 3 (g); and, when asked to "estimate how long the visitor had been away from the family", the participant fixated almost entirely on the faces of the people, but with many saccades from one face to the next, shown in Figure 3 (h) (Yarbus, 1967). This experiment has been replicated with more modern, head-free eye-tracking technology and has yielded similar results (DeAngelus & Pelz, 2009), indicating that humans have some form of top-down executive control over the locations, sequences, and durations of their visual fixations, when gathering information in a specific goal-directed manner.

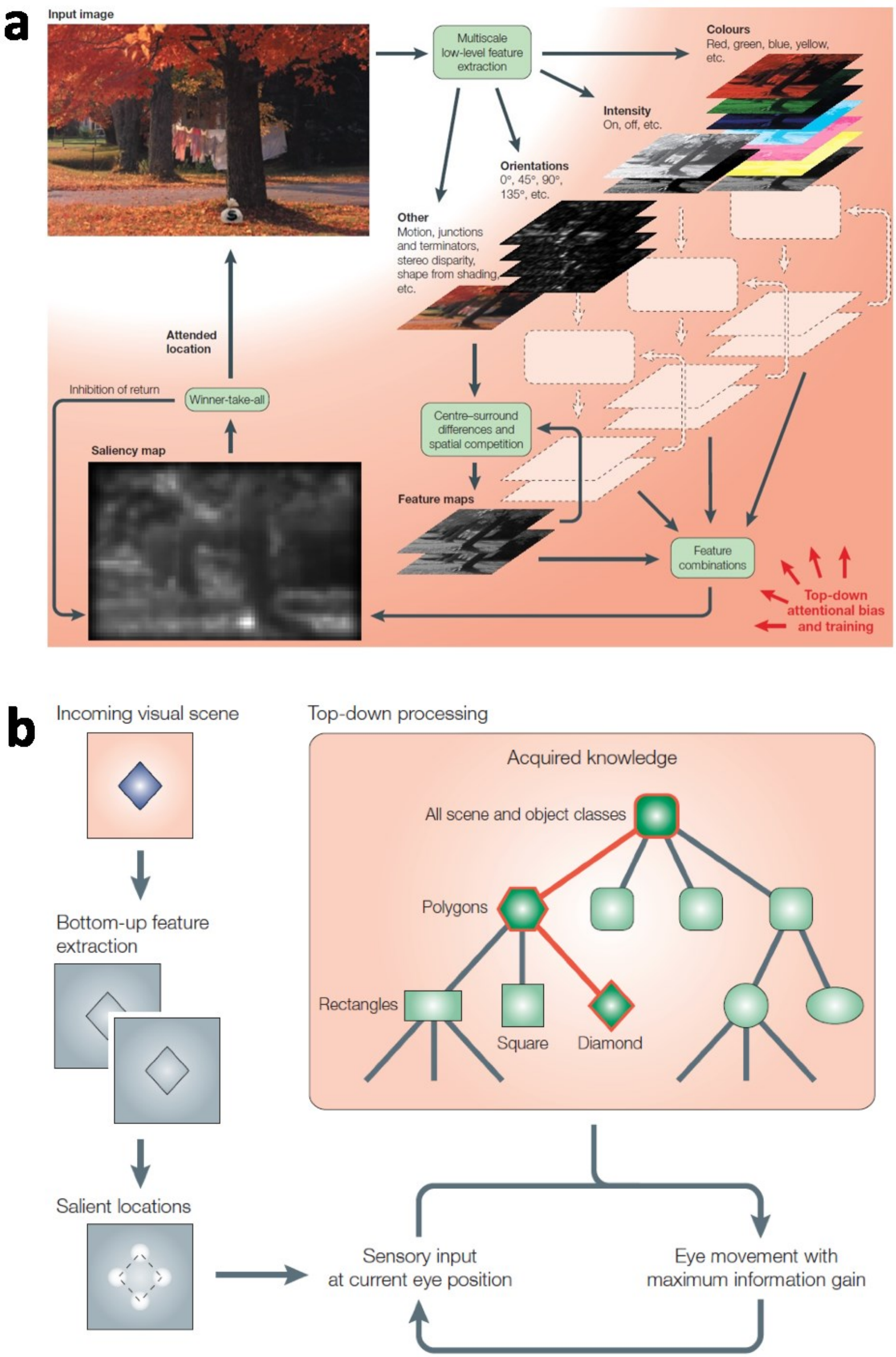

Figure 2: a) Control of bottom-up attention modelled as saliency maps (Itti & Koch, 2001). b) One representation of the relationship between bottom-up and top-down mechanisms of visual control (Itti & Koch, 2001).

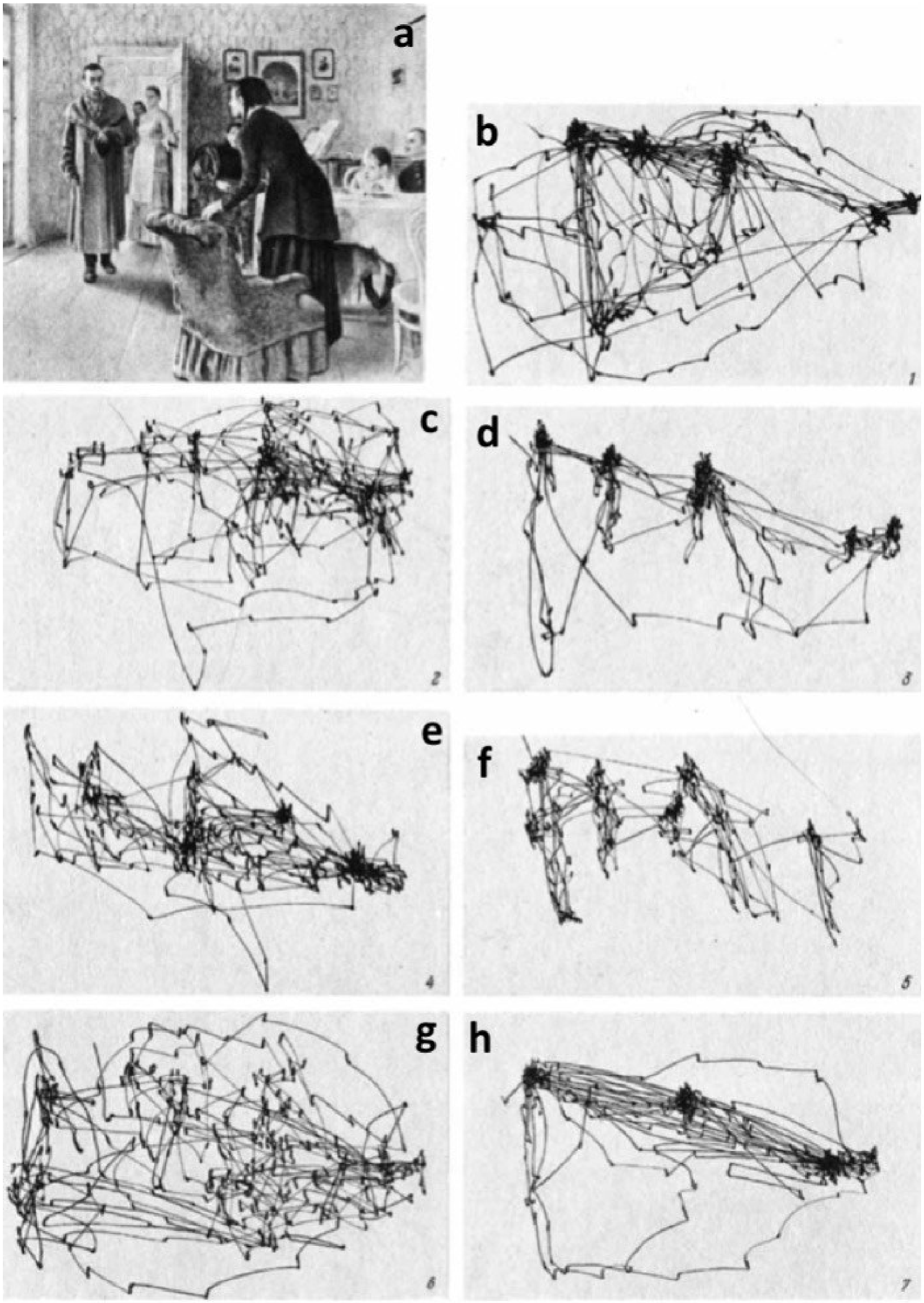

Figure 3: Eye movements from the same subject on the same image (a) when given seven different task conditions: b) Free-viewing, c) Estimate material circumstances of family, d) Deduce ages of people, e) Determine what the family is doing, f) Remember the clothes being worn, g) Remember positions of people and objects, and h) Estimate how long the visitor had been away (Yarbus, 1967).

1.2.3 – Covert and Overt: Directing attention to locations

The human retina has greater processing power at the centre of gaze, the fovea, than in the surrounding areas. Because of this, we often orient our fovea to areas at which we require greater visual information. This has been shown in the previously discussed topics on top-down visual control (Buswell, 1935; Yarbus, 1967; DeAngelus & Pelz, 2009; Henderson et al., 2007; Parkhurst et al., 2002; Birmingham et al., 2009). The circumstances under which this occurs are still being uncovered. Many studies focus on defining the 'conspicuity area', that area around the centre of gaze where an individual can detect a target (Geisler & Cormack, 2011). It is quite evident that during visual search tasks, the eyes saccade to fixate the fovea at different regions to be able to process that region to complete the task efficiently.

It has been shown by numerous studies that, although humans have the ability to shift covert attention to different regions in their field of view while their fixation is maintained, if a saccade of the eyes occurs, covert attention will involuntarily shift to the target location of the saccade immediately before the saccade (Deubel & Schneider, 1996; Henderson, 1992; Kowler et al., 1995). Deubel and Schneider (1996) used a task where participants were asked to discriminate between two images as one of these two images would appear on the screen before the eyes made a saccade but after the signal to make this saccade was given. It was found that participants had greater success at discriminating when the images appeared at the target location of the upcoming saccade, meaning that participants shifted their covert attention just before the saccade took place. This has been elaborated on to show that even when a third, distractor stimuli appears after the 'go' signal, but before the eyes make a saccade, covert attention will shift to this distractor stimuli before shifting to the target location of the saccade, only if the eyes also fixate on this distractor before arriving at the target (Peterson et al., 2004). This is evidence for a tight coupling of covert attention location and upcoming saccade location. Hoffman and Subramaniam found that as a shift of visual fixation occurs, subjects are unable to attend to a location that is not the target saccade location (1995). What follows covert attention at the location of a saccade, then, necessarily, is overt attention, at least briefly.

We will now discuss in-depth the connection between covert and overt attention. The most popular theory is that they are linked, specifically that covert attention leads overt attention on objects of interest in the visual field. This would make sense in the natural everyday tasks like tea and sandwich making which will be discussed, but leads to some rather interesting questions. Does covert attention always run at the same level in the entire visual field, like a spotlight (Posner, 1980), or are we able to manipulate the attentional intensity at specific regions of the visual field like a zoom lens (Eriksen & St. James, 1986)? Previously discussed evidence shows covert attention is first focused on a target in the periphery before the eyes saccade to the target to facilitate overt attention. In other words, overt attention relies on covert attention. In this way, it is possible that we can orient covert attention to an object, and not overt attention. But, we cannot orient overt attention without first having oriented covert attention. Yet another theory argues that covert attention should not even be called 'attention', because it is only a side-effect of overt attention as its only purpose is to plan for an upcoming saccade of the eyes (Rizzolatti et al., 1987). This 'premotor theory of attention' states that shifts of covert attention necessarily occur immediately before a saccade, and therefore are simply just a preparation for a saccade (Rizzolatti et al., 1987). This has been strengthened by other research teams, who showed massive overlap in the anatomical regions of the brain associated with the processing of attention and eye movements (Corbetta et al., 1998; de Haan et al., 2008). This also has been echoed in a study combining electroencephalography and head-mounted eye-tracking showing similarities in the neural responses of covert and overt attentional shifts (Kulke et al., 2016). Additionally, Kulke and colleagues concluded that the slightly larger measurements in the frontal brain areas during covert attentional shifting could be due to the inhibition of saccades, further evidence for a subconscious connection between attention and eye movements. Rizzolatti and colleagues extended the premotor theory of attention to the preparation of motor outputs other than eye movements (Rizzolatti et al., 1994). More recently, Gherri and Forster have shown using electroencephalography that there is likely a common generator for attentional shifts of the eyes and hands (2012). Further studies on goal-directed hand movements will be discussed in the next section.

1.3 – Attention and eye movements while moving and interacting with the world

Researchers continue to dispute the extent of connection, if any, between bottom-up and top-down eye movements, and covert and overt attention (Hunt & Kingstone, 2003). These distinct terms, while useful at times, feel a bit artificial. In the studies discussed, experiments are performed using unchanging images, on seated, head-fixed participants in laboratory environments. Perhaps these divisions in types of visual attention are more a reflection of the artificial and forced nature of the tasks and environments in which a behaviour is studied, rather than a legitimate, universal human behaviour. Perhaps if tasks emulated more natural actions of the human species, we would see a more vivid breakdown of the functions of each type of visual attention. Below we discuss studies that have tried to bridge this gap to more natural human actions and the role that attention plays in the selection of an action.

1.3.1 – Attention as an action selection mechanism

In studies that have focused more on the overt attention of subjects using eye-tracking, subjects tend to fixate at the goal of a 'top-down' directed action either with, or without a tool (Ballard et al., 1992, Ballard et al., 1995; Johansson et al., 2001; Land et al., 1999). In fact, subjects have been observed making more reaching errors on goal-directed actions when they are not permitted to look at the goal of the movement (Bekkering et al., 1995; Henriques et al., 1998). But, how does covert attention play into object interaction tasks? Extending their work on shifts of attention prior to eye movements, Deubel and colleagues showed that during the planning of a reach movement to a goal, attention is paid to the goal of a task up to the onset of movement (Deubel et al., 1998). This was further tested using electroencephalography and found that covert attention shifts to the hand immediately before the onset of movement towards a goal (Eimer et al., 2006; van Velzen et al. 2006). And, Baldauf and Deubel showed that covert attention also shifts to objects that are of priority to the subject in the specific goal-directed movements, concluding that hand movements rely on covert "visual preparation" (Baldauf & Deubel, 2008a).

Baldauf and colleagues extended their results by testing participants making more complex, multi-step reach movements (Baldauf & Deubel, 2008a; Baldauf & Deubel, 2008b; Baldauf et al., 2006). In the 2006 study, participants were instructed to keep visual fixation on a central cross while making rapid reaches to two or three targets located in their peripheral vision on the screen in front of them. Participants were cued with an arrow to the target they were to point to first, and were to continue to the next clockwise target after touching the screen with their index finger at the first target. Participants were cued with an auditory signal to begin their first movement and were to complete the task as fast and accurately as they could. A secondary letter discrimination task was used to assess covert visual attention at the peripheral targets. After the auditory cue but before the onset of movement, letters or numbers would be shown for 50ms at each of the peripheral targets. Two of the peripheral targets would flash as either E's or Ǝ's, and participants were instructed to indicate after the trial if the symbols flashed were the same symbol or different. The timeline of a trial is shown in Figure 4 (a) from Baldauf and Deubel (2010). It was found that participants were far more successful at discriminating the symbols flashed at the areas of the upcoming movement goals compared to areas that were irrelevant to the upcoming task. This is shown in Figure 4 (b), where participants were able to answer correctly far more often at locations of future manual action than at locations irrelevant to that specific trial (Baldauf & Deubel, 2010). This figure also shows the 'splitting of attention' that Baldauf and Deubel describe in their 2010 publication. When participants knew in advance that they were to move to two targets in quick succession, they split their attention between the two targets immediately before movement onset, shown by the higher percentage of symbol discrimination at the second target. Interestingly, it was not just at the first movement goal that participants were more successful, but at the second movement goal, and in subsequent experiments, the third movement goal. Remember, this was all before the first movement even began. This is strong evidence that the covert visual attention system dedicates more resources to locations of future movements. Baldauf and Deubel strengthened this theory by finding similar results in electrophysiological data (2008a). Finally, Baldauf and Deubel show that greater visual attention is paid to earlier movements in a sequence. In other words, when completing three sequential manual reaching tasks, before the onset of movement, the greatest magnitude of covert attention is dedicated to the first target, the next highest level to the second target, and the lowest level of covert attention is given to the third target in the sequence. This is deduced from the percentage of correct discriminations made at these locations, shown in Figure 4 (c) from Baldauf and Deubel (2010) which is adapted from their 2006 study. Of interest, the third target, although much lower in correct discriminations than the first, is still much greater than other areas in the visual field.

Baldauf and colleagues also extended their work by manipulating movement complexity, finding greater levels of covert visual attention being paid to the second movement if the task difficulty for this movement was increased (Baldauf et al., 2008b). The task difficulty in this experiment was increased on the second target by introducing longer delays before the second movement and removing the target of the movement completely. It was found that more parietal region resources were dedicated to the second target than the first. In another 2008 study, Baldauf and Deubel designed a similar experiment to probe covert visual attention during a bimanual reaching task (Baldauf & Deubel, 2008c). The results of this task showed that participants will divide their attention equally to targets that are equidistant from their respective hands and from a central cross on which their fixation must be maintained. More surprisingly, when targets are placed at unequal distances from each hand and the central fixation cross, more covert attention is dedicated to targets that are further away from the hand that will reach to it, perhaps extending their findings with respect to movement difficulty.

Thus, Baldauf and Deubel (2010) have moved the theory of visual attention even further from the spotlight and zoom-lens models by showing that not only does covert attention help plan subsequent hand movements to future goals, but it is also divisible. And, at that, it is differentially divisible based on the necessary accuracy of the goals of the movement, the time of each subsequent movement, and, in bimanual tasks, the distance between the hand and the target. Baldauf and Deubel describe this as the Attentional Landscape Model, a visualization of which is shown in Figure 4 (d) which has been reproduced from Baldauf and Deubel's landmark 2010 paper "Attentional landscapes in reaching and grasping" (Baldauf & Deubel, 2010).

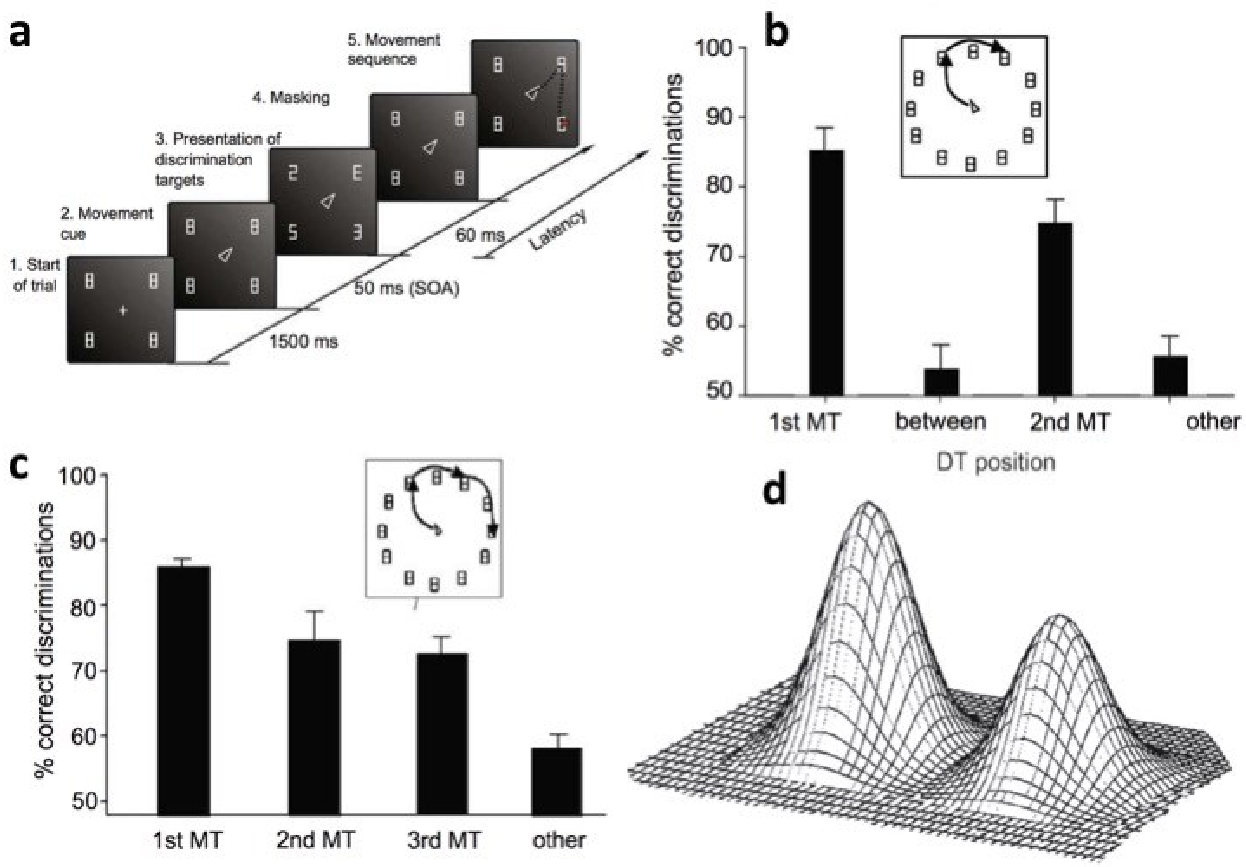

Figure 4: a) Timeline of a trial to test for strength of covert attention at areas of future manual action. b) Participants discriminate letters better at locations of future sequential manual actions than at irrelevant locations. c) Percentage of correct discriminations at three future locations of manual action. d) Illustration of an attentional landscape divided unequally between two locations. All from Baldauf & Deubel, 2010.

In summary, we contend that covert attention is used to select the targets of action. Action targets are those things which are most behaviorally relevant to the user. And, relevance is a combination of both bottom-up salience and top-down goals. Once a behaviourally relevant area has been fixated, overt attention enables the user to carry out the selected action. But, in a natural interaction task, goals do not need to be arbitrarily defined, or explicitly stated. Instead they emerge naturally. If I am making tea, then the teapot is relevant. If I am making a sandwich, then bread is relevant. This does not need to be stated by the experimenter.

1.3.2 – Eye movements in object interaction tasks

Although the studies of the likes of Buswell and Yarbus are important, the participants of these experiments had their heads held in a clamp while they were seated at a desk. Other tasks like typing (Butsch, 1932) and playing the piano (Weaver, 1943) were carried out in a similar fashion. Without detracting from their importance, these experiments were limited in their design, as participants were unable to move their head and body in a natural fashion. Since then, head-mounted eye-trackers have empowered researchers to design experiments that allow participants to move their heads and bodies, along with their eyes, in a more normal fashion. This has amounted to an increase in literature about human behaviour in more natural, everyday tasks like making a sandwich or a cup of tea.

1.3.2.1 – Object related actions: The work of Land and Hayhoe

As we have seen, there are many studies exploring the eye movement patterns of humans in the laboratory or in relatively unnatural tasks, but how do the eyes behave when doing everyday tasks in a natural setting? Pelz and colleagues moved toward addressing this question by recording the eye, head, and hand movements of ten subjects during a block-copying task in which they were to view a 3D model of blocks and replicate it (2001). Subjects were fitted with a monocular, headband-mounted, infrared eye-tracker, and a magnetic field tracker for their head (6 DOF) and thumb. Gaze location of the eyes was obtained from eye-in-head and head position signals. Pelz and colleagues found that although there were varied patterns of movement for participants and different strategies implemented, there was a regular, cyclical rhythmic pattern of eye, head, and hand movements, in that order, taking an average of 1.5s (2001). When picking up or putting down a small object with a precision grip, the eyes move first, then the head, then the hand. Additionally, the eyes were the limiting factor of both the head and the hand moving.

For example, if the eyes were occupied by fixating a relevant area in the task, the head and hand would stall their movements. As the blocks required precise manipulation, researchers inferred that putting a block down required both visual feedback and proprioception, while pick-up only required proprioception, as the eyes would saccade to the next target just before the pick-up had occurred (Pelz et al., 2001). Finally, this was the first instance that showed there is little use of spatial memory in object manipulation tasks, as the eyes lead the head and hand on nearly every manual manipulation of an object. This had previously been popularized as the 'do it where I'm looking strategy' seen in a 1992 study by Ballard and colleagues (Ballard et al., 1992). In that study, however, subjects' heads were held still by a bite-bar and the tasks were done on a computer screen using a mouse.

Up until the 1990s the available technology had not allowed for eye movements during natural tasks to be recorded. The first attempt to define this type of eye behaviour was done by Land and colleagues in 1999 (Land et al., 1999). In this experiment, three subjects had their eye movements recorded by a head-mounted, monocular eye-tracker while they made tea in a small kitchen. The device recorded the view of the scene directly ahead of the subjects (first-person view) while also recording the eye by using a concave mirror. After calibrating these two images together, the experimenters were able to determine the foveal direction of the eye relative to the head to an accuracy of 1° (Land et al., 1999). A white dot representing the subject's foveal direction was then superimposed onto the first-person view of the scene, allowing for extensive analysis of gaze to take place (Land et al., 1999). In addition, a third-person view of the experiment was recorded, allowing the researchers to determine when each subject moved their body. Another landmark study with the aim of understanding how the eyes behave during everyday tasks in a natural environment was carried out by Mary Hayhoe in 2000. Seven subjects were seated at a table, with a similar head-mounted eye-tracking device as described above, and asked to make a peanut butter and jelly sandwich and pour a glass of cola (Hayhoe, 2000). Both studies allowed participants to move their heads freely, although the tea-making study is arguably closer to a 'natural' task as participants could move their bodies freely around the kitchen, while those in the sandwich-making task were required to stay seated. In both studies, participants were not given directions on how to carry out the task, which lead to variability in the order of some movements (eg. one participant in the tea-making task poured milk into the mug before the tea, another did the opposite). This did, however, allow for a natural 'top-down' selection of eye and hand movements by each participant, which may not have occurred in a defined sequence task. As well, both studies included irrelevant objects, as would occur in a natural environment. Land and Hayhoe analyzed their studies together to obtain a better understanding of how the eyes enable ongoing motor tasks to occur successfully (Land & Hayhoe, 2001). Because there is a hierarchy in the segregation of the motor tasks (making tea can be separated into 'filling the kettle', 'pouring the milk' etc., and 'filling the kettle' can be separated into 'locating the kettle', 'picking the kettle up', 'removing the lid', 'turning on the tap' etc.), Land and Hayhoe described any actions and eye movements performed on an object without an interruption as an 'object related action', or ORA (2001). So, picking up a cup (reaching and grasping it, terms we will use in this thesis), moving it across the counter (transport), and putting it down (release) would be one ORA. In both studies, the average ORA took just over 3s.

There are some striking similarities in the findings of these two different tasks, which is strong evidence for a model that must exist that controls the interactions between the visual and motor systems during everyday tasks. The eyes fixated on an object before the hands manipulated it (Land & Hayhoe, 2001). The average time between the eyes fixating an object and the hand manipulating it in the tea-making task was 0.56s, while only 0.09s in the sandwich-making task (Land & Hayhoe, 2001). As well, a latency of the hands at an object was seen at the end of manipulation, as there was about half a second between the eyes leaving an object and the hands following (Land & Hayhoe, 2001). Land and Hayhoe attribute this to a half-second buffer of visual information in the brain (Land & Hayhoe, 2001).

Objects were generally fixated for the duration of manipulation, except for some extended tasks like filling the kettle, where the eyes would sometimes fixate on irrelevant objects or objects to be used in the future during the task (Land & Hayhoe, 2001). Other than during wait times, the average of both tasks saw fixation time less than 5% on objects that were irrelevant to the goal of the task. This is some evidence for a 'top-down' driven approach to the visual system when performing a specific task. In fact, in a portion of the sandwich-making trials, the number of irrelevant objects was increased to half of the total number of objects on the table. Before a trial began, fixation on these task irrelevant objects was 52%, but once the trial began, this dropped to 18%. Land and Hayhoe hypothesize that this represents a shift from 'bottom-up', 'salience-driven' to 'top-down', 'task-driven' eye behaviour. In both tasks, there was strong consistency across participants providing even more evidence that the visual system is strongly influenced by the specific objectives of a task (2001), matching with what the Attentional Landscape Model would call behavioral relevance.

Some of the differences between the tea and sandwich tasks were: the eye-hand latency in the tea-making task was longer, there were more, shorter fixations in the sandwich-making task, and there were more unguided releases of objects in the sandwich making task (Land & Hayhoe, 2001). Land and Hayhoe have attributed these differences to the closer proximity of the objects in the sandwich-making task compared to the tea-making task, and the fact that participants were seated during the sandwich-making task while in the tea-making task they were standing. The tea-making task also required participants to turn and walk across the room to retrieve some objects, while the required objects in the sandwich-making task were laid out in front of the seated participants.

From these studies, Land and Hayhoe describe four functions of visual fixations when interacting with objects: Locating, Directing, Guiding, and Checking (2001). Fixations during 'locating' are so described as they are used to find an object that will be interacted with soon. In both tasks, participants spent time at the beginning of their trial finding the location of some objects before beginning a manipulation. There is no associated motor activity with 'locating' fixations. Fixations during 'directing' actions on an object occur immediately before the hand manipulates the object. Generally, the eye will leave the object before the hand arrives creating the assumption that the visual system is providing several types of information (object position, shape, etc.) to the motor system that does not require constant visual feedback for successful motor activity. Similar 'directing' fixations also occur when objects are put down. Fixations during 'guiding' are more complex than 'directing' fixations as they usually involve more than one object being brought together (a kettle and its lid). There are often a few 'guiding' fixations made in close succession as the eyes shift foveal direction from one object to another (from the kettle to its lid, then back to the kettle). Finally, fixations during 'checking' usually involve the completion of an action, and can either be one of long duration, or several of shorter duration (Land & Hayhoe, 2001). Fixating on the 'fill line' of a kettle while the hand turns the knob of the tap would fall into the category of 'checking'.

Land and Hayhoe agree that when picking up or putting down an object, the hand is rarely, if ever, fixated as the fixation lies on the object, or its target location, respectively (2001). When an object has been reached by the hand, the eyes do not fixate on it. And, when pouring a liquid, fixation lies on the object being filled, rather than the pouring object. These rules lead to the conclusion that "vision is a scarce and valuable resource, and it is disengaged from a particular aspect of an action as soon as another sense is available to take over" (Land & Hayhoe, 2001).

Land and Hayhoe created a diagram showing their proposed flow of information during ORAs, shown in Figure 5 (a). In this top-down model, the schema provides the visual system with information about the next object to be interacted with, the oculomotor system with information about the location of this object, and the motor system with specific actions to be performed on this object. Once the object's location has been located by the visual system, the motor system initiates the movements of the hand towards where an action will be performed. Monitoring of the actions is done through a combination of visual and oculomotor systems (Land & Hayhoe, 2001). It can be argued, however, that this is a simplistic view of the monitoring of the action, by failing to consider the feedback from the motor system, and, that vision could be used more in a confirmatory role, rather than supervisory role, during monitoring. In any case, when the monitoring system deems an action is complete, the action is terminated and the schema then selects the next object to be interacted with (Land & Hayhoe, 2001).

Land and Hayhoe conclude that the findings of Underwood and Everatt (1996), who argue unconscious automatic actions do not require feedback while consciously controlled actions do, are not entirely true (Land & Hayhoe, 2001). In this study of easy, automatic ORAs, there are many instances where the eyes are monitoring an action, even actions on objects that have previously been fixated. In fact, the eyes very strongly precede and predict almost every action carried out in these tea and sandwich making tasks, providing little evidence that the visual system creates a model of the environment while performing these tasks. The required information to successfully complete a manual task seems to be obtained from the visual system every time, suggesting that the visuomotor system solves object interaction problems in real-time. (Land & Hayhoe, 2001).

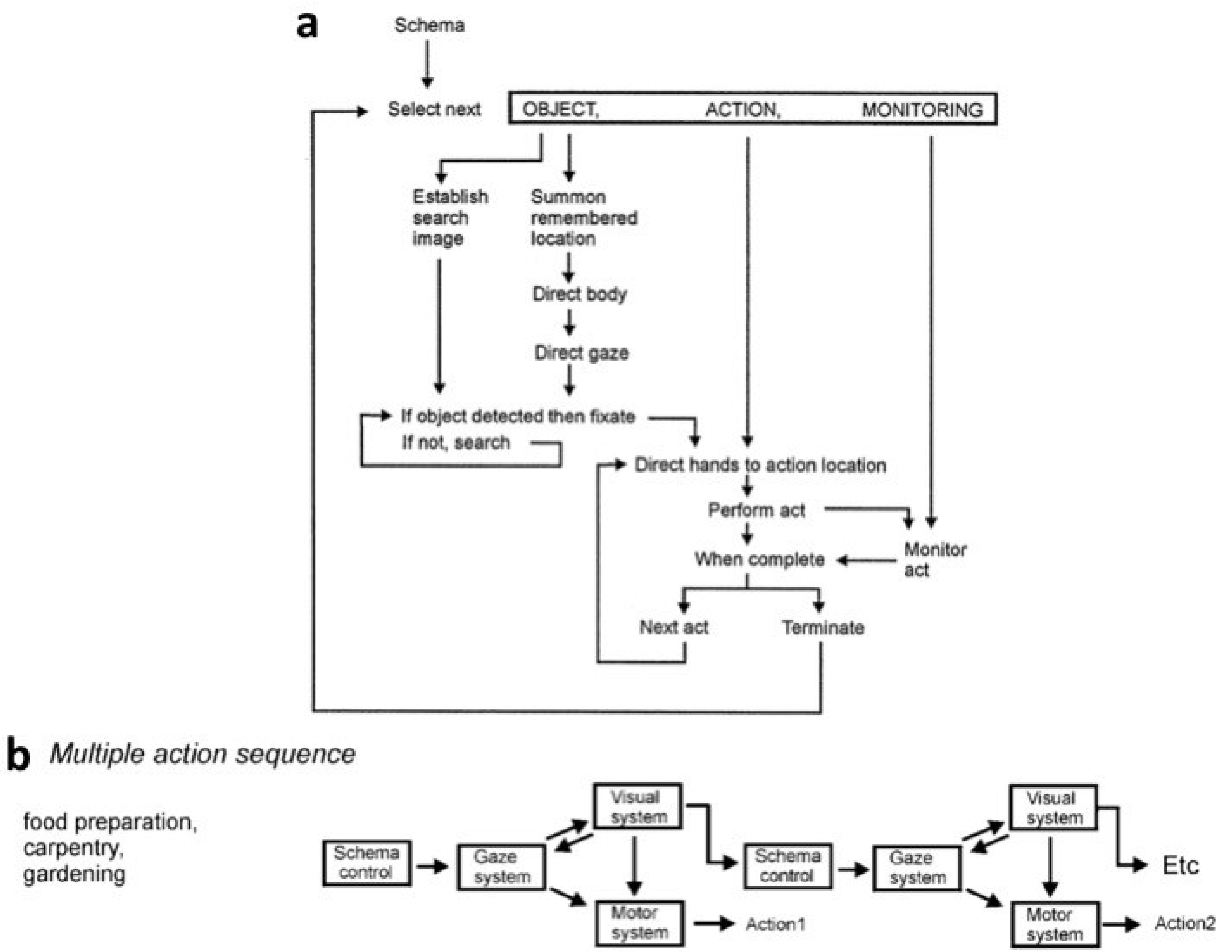

Figure 5: a) The path of information processing during an object related action (Land & Hayhoe, 2001). b) Sequence of control during complex, natural tasks like making a sandwich or a cup of tea (Land, 2009). The schema system plans the overall task, the gaze system finds and fixates objects, and the visual system relays information to the schema and gaze systems. This has been replicated from their 2001 paper, “In what ways do eye movements contribute to everyday activities?”.

1.4 – Current Study

1.4.1 – Motivation

Recent advances in eye-tracking technology have allowed researchers to move from testing stationary and seated individuals to testing fully-mobile participants engaging with real objects. However, even with these advances, the vast majority of studies using eye behaviour still use restrictive lab-based tasks that are not representative of the demands on the eye movement system in the real-world (Kingstone et al., 2008). Here, we take a track akin to the Cognitive Ethology research approach and feel it is necessary to connect lab-based tasks to real-world tasks in order to gain an understanding of natural human behaviour (Kingstone et al., 2008). The experiments of Mary Hayhoe and Michael Land were some of the first with this aim, however, in these experiments body movements were not precisely and accurately recorded (Hayhoe, 2000; Land et al., 1999). Without a doubt, the most defining characteristic of these previous studies is the freedom given to the participant. The only instruction given to participants was to either "make a peanut butter sandwich" or "make a cup of tea". This freedom, although most similar to everyday tasks, does not enable the researcher to compare one participant's eye movements to another's easily, as participants could carry out tasks in a slightly different order. By providing participants with the manner and order of movements, more consistent eye movement behaviour can be observed. As well, in previous studies, participants are provided distractor objects that are not needed to carry out their objective. Although interesting and relevant to normal, everyday life, this creates another dimension of variability that complicates the study and creates difficulty when trying to parse out normal behaviour. Therefore, our goal was to create goal-based tasks involving clear object interactions, representative of real-world tasks, requiring overt attention. As well, we aimed to use the least restraining eye-tracking technology available to allow participants to perform these tasks in the most natural way possible.

Thus, we recorded eye movements in two tasks mimicking real-world demands, establishing a normative data set for functional eye gaze behaviour during standardized tasks. The first task emulates moving a box of pasta from a countertop into a cupboard (the Pasta Box Transfer Task - PB) and the second task emulates moving cups across a table (the Cup Transfer Task - C). See Figures 6 and 7 for images of the respective task setups. Participants were to stand relaxed in front of the tasks, and were free to turn their bodies and move their heads, however, they were asked not to walk around, and to use their right arm to perform the tasks. The tasks were designed to mimic everyday activities and each had unique requirements. The PB task required the individual to turn their head and body and interact with an object at different heights. The participant turned to the right and grasped a pasta box at the height slightly lower than that of a typical kitchen counter and turned back to face forward and placed the box into a shelf at typical kitchen-height. Then, the participant proceeded to move the pasta box from the current shelf to a second, higher shelf located in front of them but on the left side. Finally, the participant moved the pasta box back to its original starting position. Thus, there were 3 object related actions (ORAs), each comprised of a reach, grasp, transport and release. The C task has smaller objects (compliant wax paper cups) filled with beads requiring more precision in grasp modulation with a consequence of possible spillage. The C task was arranged in front of the participant at typical kitchen counter height, with 2 cups in the right side of a partitioned box. Participants moved the first cup from the right side of the box over a partition to a specific placement target on the left. Due to the location of the cup in the box, this required grasping from the top of the cup. They then immediately moved a second cup over the partition from right to left, requiring grasping the side of the cup. Next, participants moved the second cup back to its starting location using a side grasp, and immediately moved the first cup back to its starting location using a top grasp. Thus, for the C task, each trial had 4 ORAs.

We used head-mounted eye-tracking and kinematic movement tracking systems to characterize both eye and body movements. As has been mentioned earlier, distractor items can affect eye behaviour in object movement tasks. To check if the wearing of motion capture markers would draw eye gaze, we collected data on both the PB and C tasks with a full upper-body kinematic marker set, as well as a significantly minimized marker set, to discover any differences in eye behaviour that could be attributed to the presence of the markers. This will be discussed in detail in the Methods section.

Our objective was to find task similarities and differences in eye movement patterns during functional object related actions. Specifically, we intended to demonstrate the amount of time people spend reaching, grasping, transporting, and releasing objects during everyday tasks, and, during each of the abovementioned actions, what objects are visually fixated and for how long. In addition, we aimed to uncover the temporal relationship between an object being visually fixated and the hand beginning a manipulation of the object.

1.4.2 – Predictions

For our experiments, we expect to see the eyes fixate on objects that are relevant to the objectives of the task and very little, if any, fixating on irrelevant objects (Land & Hayhoe, 2001). Yarbus (1967), Buswell (1935), DeAngelus and Pelz (2009), and others have showed that the participant's behaviour can be changed by the objective of the task. So, because our task requires participants to interact with objects, in general, we expect to see each participant fixate on the objects that they plan to interact with. This has been shown in many studies already (Land, 1999; Hayhoe, 2000; Belardinelli, 2015). As well, we expect to find at least one visual fixation towards each object immediately before the participant's hand arrives at the object, and expect the participant to fixate on the target drop-off location of each object before the hand arrives with the object. This follows the findings of Hayhoe (2000), and Land (1999) that show the eyes leading the hand in a strong correlative manner, although the exact time in advance that the eyes lead the hand differs between the two studies. In fact, we expect to see a very strong predictive pattern of the eyes leading the hand; so strong, that if we were only given the periods of time that each object was fixated, we would be able to predict where the hand was in space.

We expect the amount of time that the eyes fixate on each object to differ depending on the location of the object, with objects placed in front of participants at the start of the task in their field of view having longer fixation times, and objects outside of their field of view at the start of the task having shorter fixation times. As well, we predict that larger, easier-to-move objects like the pasta box will have shorter fixation times, while smaller, riskier-to-move objects like the cups filled with beads will have longer fixation times, in general. In the C task specifically, the two different grasp patterns may require different precision when grasping, which, in turn, may require different periods of visual fixation. For example, it is possible that a top grasp requires more precise placement of the index finger, resulting in longer fixations than performing an action with a side grasp.

During the Reach phase of the hand moving towards an object, we expect to see the participant fixate on the object the hand is moving towards. At some point, either at the end of the Reach phase or during the Grasp phase we predict the eyes will saccade to the next target –

the drop-off location for the object. The eyes will maintain fixation on the drop-off location until the object has contacted the drop-off location, at which point the eyes will saccade to the next object of interaction.

The period of visual fixation will correlate with the amount of time it takes a participant to complete that phase of movement. So, if a participant reaches towards an object more slowly, they will fixate longer on the object because it will take longer for their hand to arrive at the object. Or, if the participant moves at the same speed, but has more distance to cover, the fixation towards the object during the reach will be longer.

Importantly, we do not expect to find any substantial or consistent visual fixation towards a participant's own hand. We expect that participants will use other sensory modalities like proprioception and haptic feedback to accomplish the goals of movement of the hand.

Finally, we predict an approximate 0.5s Eye-Hand Latency in our two tasks, based on the values obtained in the tea-making task from the Land and Hayhoe studies (2001). As stated earlier, the sandwich-making task resulted in much shorter Eye-Hand Latency values possibly because participants were seated doing a task requiring finer motor movements with targets much closer together resulting in the lag of the hand behind the eyes to be shorter. Our tasks are more like the tea-making task, as participants are standing, which will force similar kinematic, and biomechanical movements, and likely elicit similar eye movement behaviours.

The basic pattern of fixations and actions during our two tasks should follow the sequence provided in Figure 5 (b).

2 - Methods

2.1 – Participants

A group of 24 adults, who had no upper body pathology or any history of neurological or muscoskeletal injuries within the past two years were recruited to participate in our study. Of these, 4 data sets were dropped due to apparatus and/or software issues. The remaining 20 participants (11 male) had an average age of 25.8 ± 7.2 years, an average height of 173.8 ± 8.3 cm, and were made up of 18 self-reported preferred right-handed users, and 2 self-reported preferred left-handed users. As well, 18 participants had normal or corrected to normal vision, while all participants were naïve to the purposes of the experiments. All procedures were approved by the University of Alberta Health Research Ethics Board, the Department of the Navy Human Research Protection Program (DON-HRPP) and SSC-Pacific Human Research Protection Office (SSCPAC HRPO).

2.2 – Apparatus

2.2.1 – Pasta Box Transfer Task Apparatus

For the PB task (see Figure 6), a table 36” high and 32” wide, with two 16”-wide shelves set 9” back from the front edge of the table, at heights of 7” (Mid Shelf) and 12” (High Shelf) from the top of the table on the right and left, respectively, was used. As well, a table was placed to the right perpendicularly to the table at a height of 30”. Three 3.5” x 4.5” rectangular targets were located on the apparatus. The first (Start/End) target was placed on the side cart with its centre 13.5” away from the edge of the table and 7.75” to the right of the close edge of the cart.

The second (Mid Shelf) target was placed on the right shelf on the table with its centre 12.5” from the front edge of the table and 8” to the right of the midline of the table. The third (High Shelf) target was placed on the left shelf on the table with its centre 12.5” from the front edge of the table and 8” to the left of the midline of the table. As well, a 3.25” x 2.5” ‘Home’ area was placed on the right corner of the table at the front edge, with its centre 12.5” right of the midline. The ‘Neutral eye position’ marker was placed on the midline of the table 9” from the front edge of the table and 18.5” above the top of the table. Finally, the ‘Pasta box’, a standard-sized Kraft Dinner box, weighed 225 grams with dimensions 7” x 3.5” x 1.5”. See Figure 6 (a), (b), and (c) for visual descriptions of the PB apparatus.

2.2.2 – Cup Transfer Task Apparatus

For the C task (see Figure 7), a box with interior dimensions of 30” wide and 14” depth with 3” high edges, and a 6” high central divider sat on top of a table, 36” above the ground and 32” wide, 2.5” from the front edge. Four 3.15” x 3.15” targets with 2” diameter circles around their centres were placed on the interior of the box. Two green (Near) targets had their centres placed 3” from the front edge of the box, one with its centre 3” from the right edge of the box (Near 1), and the other with its centre 3” to the left of the midline of the box (Near 2). Two blue (Far) targets had their centres placed 3” from the back edge of the box, one with its centre 3” to the right of the midline of the box (Far 1), and the other with its centre 3” to the right of the left edge of the box (Far 2). As well, a 3.25” x 2.5” ‘Home’ area was placed on the right corner of the table at the front edge, with its centre 12.5” right of the midline. A ‘Neutral eye position’ marker was placed on the back edge of the 6” high divider, 16.5” from the front of the table. Two standard 5 oz. Wax Treated Paper Cold Cups (58PATH, Dixie Consumer Products, LLC) were filled with beads (Soft Plastic Pellets A4155 – Phase 2, Patterson Medical Holdings, Inc.) to a weight of 85 grams (including the weight of the cup). The 'Green' cup has a coloured green stripe along the top rim of the cup (made with a permanent marker), while the 'Blue' cup has a coloured blue stripe around the centre of the cup (made with a permanent marker) for cueing the participant to the type of grasp required. See Figure 7 (a) and (b) for visual descriptions of the C apparatus, and Figure 7 (c) and (d) for visual information of the cups and beads used.

2.2.3 – Eye and Motion Tracking Apparatus

Participants were fitted with a Dikablis Professional 2.0 head-mounted, binocular eye-tracker. The eye-tracker rests on the bridge of the nose (like a pair of glasses), the forehead, and the sides of the head above the ears. Participants were asked to position the headset comfortably before experimenters tightened the built-in elastic strap on the back to hold it steadily in place. The Dikablis headset records each eye in infrared at 60 Hz, and is equipped with a forward-facing, high-definition scene camera which records the first-person view of the participant. All three of these cameras can be moved before they are calibrated together, enabling experimenters to position the cameras for the best data collection specific to the task. Figure 1 (i) and (j) show the Dikablis headset on a subject from a front and side view, respectively. 12 infrared cameras and accompanying motion capture markers were used, along with the Vicon Nexus 2.0 software, to track each participant's movements. Figure 6 (c) shows a Pasta box with motion capture marker placements, Figure 7 (c), (d), and (g) show the motion capture marker placements on the Green cup, Blue cup, and hand and digits, respectively. Markers were placed on these objects in only one of the conditions of our experiment, which will be discussed in detail later. Since the primary focus of this thesis is on eye-tracking, we will not be going into additional details about the motion tracking parameters.

Figure 6: a) The Pasta Box (PB) task set up with relevant dimensions and locations of the targets and the neutral eye position marker. b) A top-view of the PB task, without the side table. c) A Pasta box with marker placements used in the PB task during the Both Condition. No markers were on the Pasta box during the Eyes Only Condition.

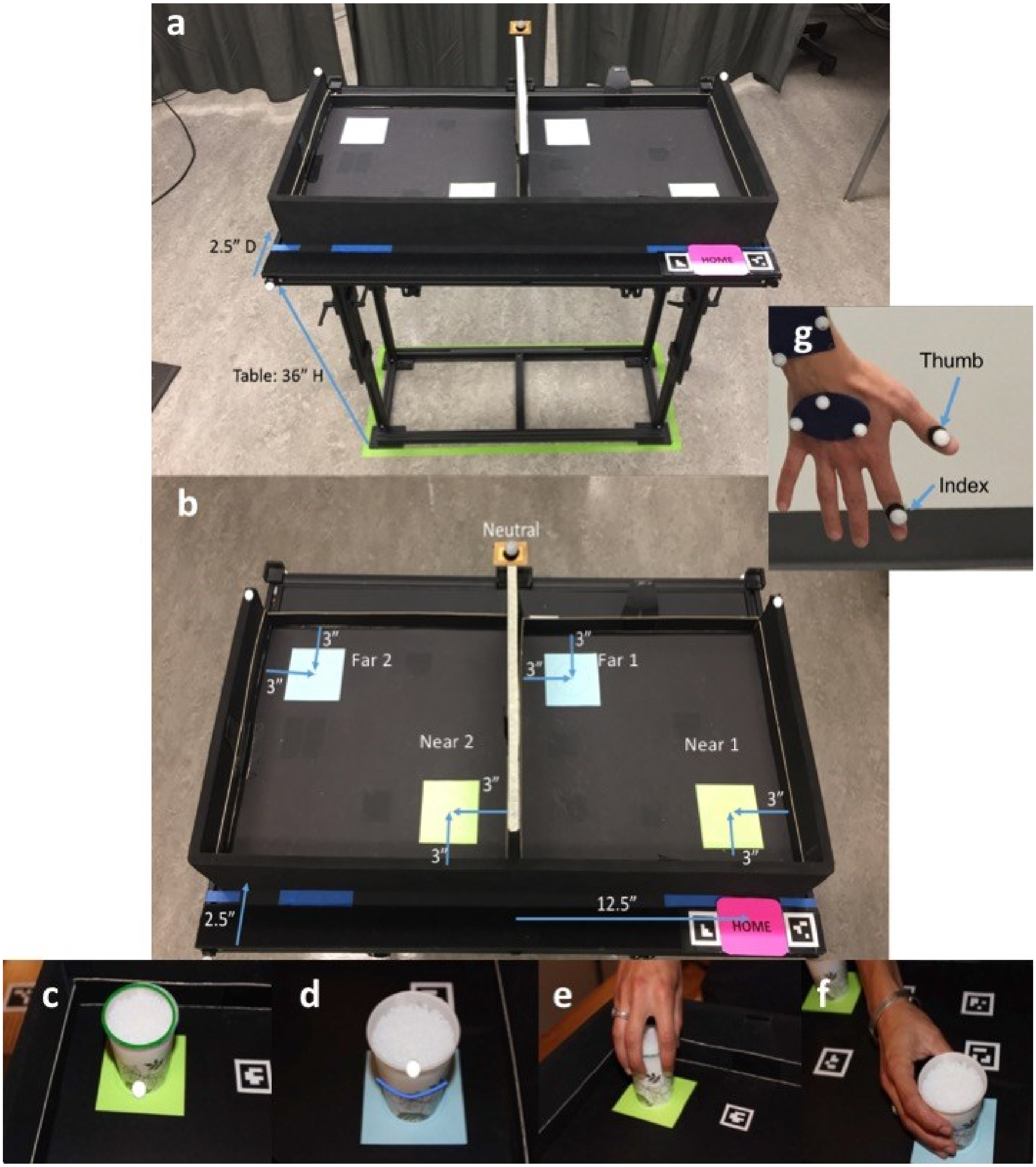

Figure 7: a) The location in space of the apparatus used in the Cup (C) task. b) The locations of the targets, and the Neutral eye position in the C task. c) The Green cup on one of the two green targets, with motion capture marker placement during the Both Condition. d) The Blue cup on one of the blue targets with motion capture marker placement during the Both Condition e) A participant grasping the Green cup with a top grasp. f) A participant grasping the Blue cup with a side grasp. g) Motion capture marker placements on the hand, thumb, and index finger of a participant during the Both Condition. Motion capture markers were not placed on the cups, thumb, or forefinger during the Eyes Only Condition.

2.3 – Procedure

2.3.1 – Experimental Setup

For our experimental design, we had each participant complete each task (PB and C) under three different conditions. The first condition (Eyes Only Condition) had participants wear only the head-mounted eye-tracker and one marker located on the back of their hand (required for data segmentation). In the Eyes Only Condition, markers were not placed on the pasta box or the cups. The second condition (Both Condition) had participants wear the head-mounted eye-tracker, and a full set of 57 upper-body motion capture markers, which included markers on the forefinger and thumb, and a plate with three markers on the back of the hand. In the Both Condition, additional markers were placed on the pasta box (4), and the cups (1 each). The third condition (Motion only), which will not be discussed in detail in this thesis, had participants wearing only the full set of upper-body motion capture markers, without the head-mounted eye-tracker. The order of these conditions was randomized in a way that a third of the participants started with the Eyes Only Condition, a third started with Both Condition, and a third started with motion capture only condition. However, due to the nature of the motion capture system setup, the motion capture only condition and Both Condition were always collected one after the other. The order of the tasks was also randomized within each condition. Participants were required to wear tight fitting compression garments for the experiments for motion capture purposes. The top was sleeveless with narrow shoulder straps to allow for good range of motion at the shoulder, while the bottoms were tight-fitting compression shorts.

2.3.2 – Eye-tracker Calibration

After the initial donning and adjustment to fit the eye tracker, the angle of the first-person HD scene camera was adjusted to ensure all relevant eye movements of the task would be captured. As both tasks were directly in front of the participant's eyes or lower, the scene camera was usually angled down. Next, the two pupil cameras were positioned to ensure the built-in D-Lab software could recognize the participant's pupils when fixated at all relevant locations during the task. Figure 8 (a) shows a poor configuration of the pupil cameras (e.g. the pupils are not recognized by the software), while 8 (b) shows an optimal configuration. The participants were then asked to stand in the task zone and look to regions of interest important to the upcoming experiment. As they did so, experimenters monitored their eye movements in a video visualization on the D-Lab software, to ensure their pupils were being recorded at all positions. If the pupils were incorrectly detected by the built-in software, at any point, the experimenters modified the position of the eye cameras. Once these three cameras were in appropriate positions, the experimenter measured the distance from the scene camera to the Neutral eye position marker (in the C task), or the centre of the High Shelf Target (in the PB task) when the participant was standing in the task location. The eye-tracker was calibrated to the eye movements of each participant from the distance measured. Participants were instructed to stand still and fixate, without moving their head, on 4 points on the wall highlighted by bright red circles at the experimenter's call while the experimenter captured these locations in the D-Lab software. Figure 8 (c) shows the visualization the experimenter sees during the 4-point calibration.

This calibration process occurred between 2 and 4 times for each data collection session, contingent upon the order of the three conditions within the two tasks the participant was assigned to perform. If any of the cameras were moved, or if the headset shifted, the calibration process outlined above was repeated.

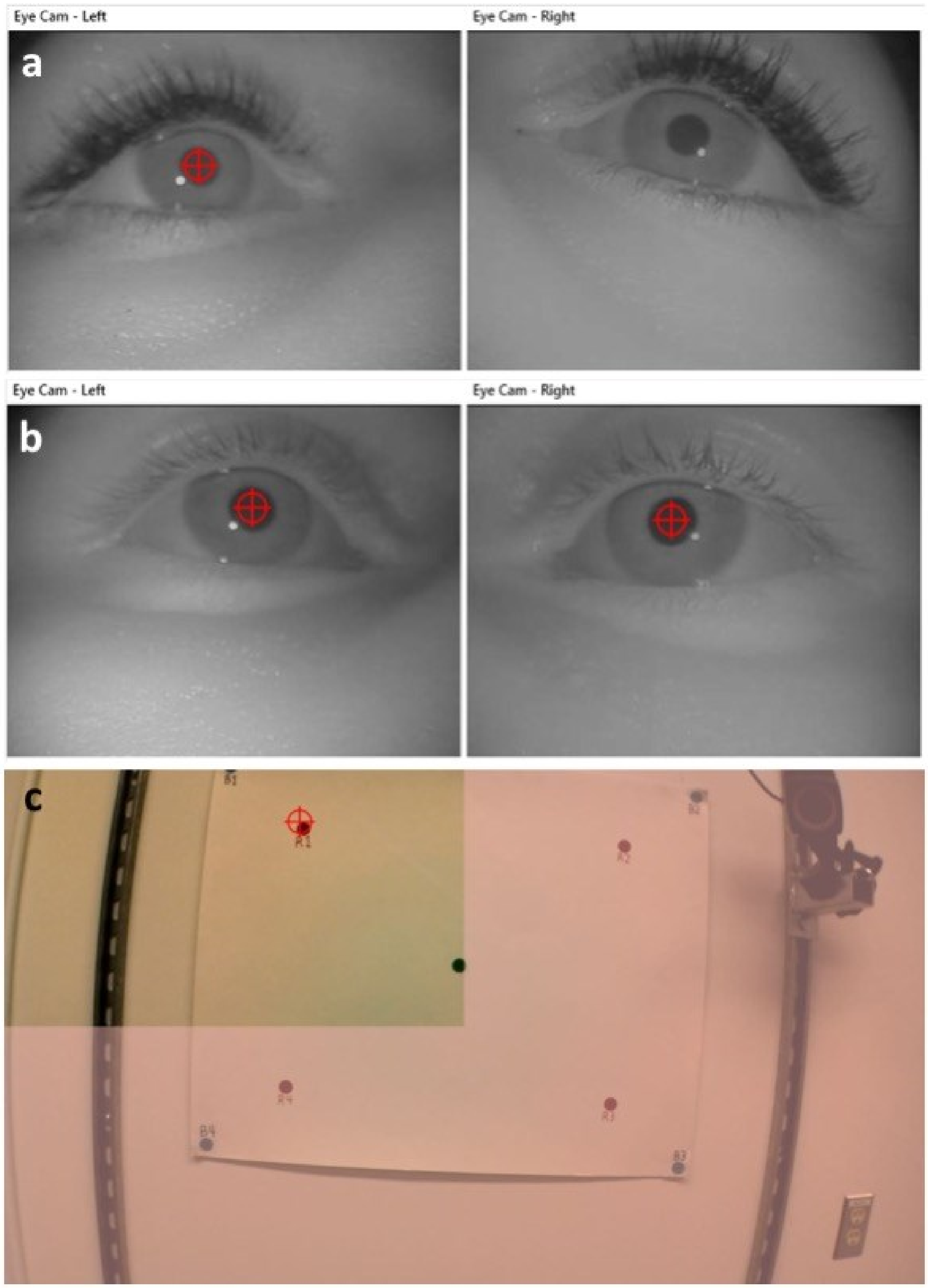

Figure 8: a) IR visualization of a poor configuration of the pupil cameras. b) IR visualization of a good configuration of the pupil cameras. c) First-person view from the HD scene camera during the built-in 4 point calibration of the D-Lab software. Notice the color differentiation which indicates to the experimenter which area the participant should be instructed to fixate in.

2.3.3 – Tasks

Three conditions of two tasks were performed by each participant. Each combination of condition and task was completed as many times as necessary to obtain 20 trials without errors. Therefore, each data collection session was a minimum of 120 trials. Outlined below is an in-depth description of each of the Pasta Box (PB) and Cup (C) transfer tasks. Both tasks were performed by the right hand only, and participants were asked to keep their left hand in a relaxed position.

2.3.3.1 – Pasta Box Transfer Task

The Pasta Box Transfer Task requires participants to perform an initial grasp of a Pasta box from the Start/End Target on the side cart at the right side of the body and move it to the Mid Shelf Target in front of them (Movement 1, Figure 9a). From there, participants move the box to the High Shelf Target by crossing the body's midline (Movement 2, Figure 9b). Finally, the Pasta box is picked up again from the High Shelf Target and placed back at its initial position on the Start/End Target on the side cart (Movement 3, Figure 9c). Between each movement, and at the end of the third movement, participants were required to place their hand at the 'Home' position to allow for greater movement standardization and proper task segmentation.

Participants were instructed to perform the task at a comfortable, but efficient pace, as to avoid making errors. As well, all grasps were to be side grasps, so that the thumb contacts the left side of the Pasta box with the greatest surface area, the fingers contact the right side of the box with the same surface area, and the palm either contacts or lies directly adjacent to the long edge of the Pasta box. Participants were instructed to complete the movements in the sequence outlined above, place the box on the short edge within the boundaries of each target, and avoid dropping the box, contacting the apparatus, hesitating, or making undesired movements (like scratching one's leg). If a rule was violated, participants were told to complete the trial to the best of their ability and that an extra trial would be added at the end of that group of trials. For example, if a participant violated one of these rules in 3 separate trials, 23 trials would have been collected.

2.3.3.2 – Cup Transfer Task

The Cup Transfer Task, requires participants to move two compliant cups filled with beads from an initial position to a final position by lifting the cup and clearing a partition. The first cup (Green cup) was moved from the near right corner of the right portion of the box (Near Target 1) to the near right corner of the left portion of the box (Near Target 2) and had to be moved with a top grasp (Movement 1, see Figure 10a). The second cup (Blue cup) was moved from the far left corner of the right portion of the box (Far Target 1) to the far left corner of the left portion of the box (Far Target 2) and it had to be moved with a side grasp (Movement 2, see Figure 10b). Once the cups were moved to their target locations, participants returned their hand to the 'Home' position and proceeded to transport the cups back to their initial positions by inversing the order, therefore moving the Blue cup first (from Far Target 2 to Far Target 1, Movement 3, Figure 10c) and the Green cup second (from Near Target 2 to Near Target 1, Movement 4, Figure 10d), returning their hand to the 'Home' position after Movement 4.

As with the PB task, participants were asked to perform the C task at a relaxed but efficient pace, to avoid errors. Along with the grip instructions that have been outlined above, participants were asked to complete the task in the sequence outlined, and avoid dropping or deforming the cup, contacting the partition or any other portion of the apparatus, spilling any beads, hesitating, or making undesired movements (like scratching one's leg). Again, if a rule was violated, participants were instructed to complete the trial to the best of their ability, and that an extra trial would added at the end of that group of trials.

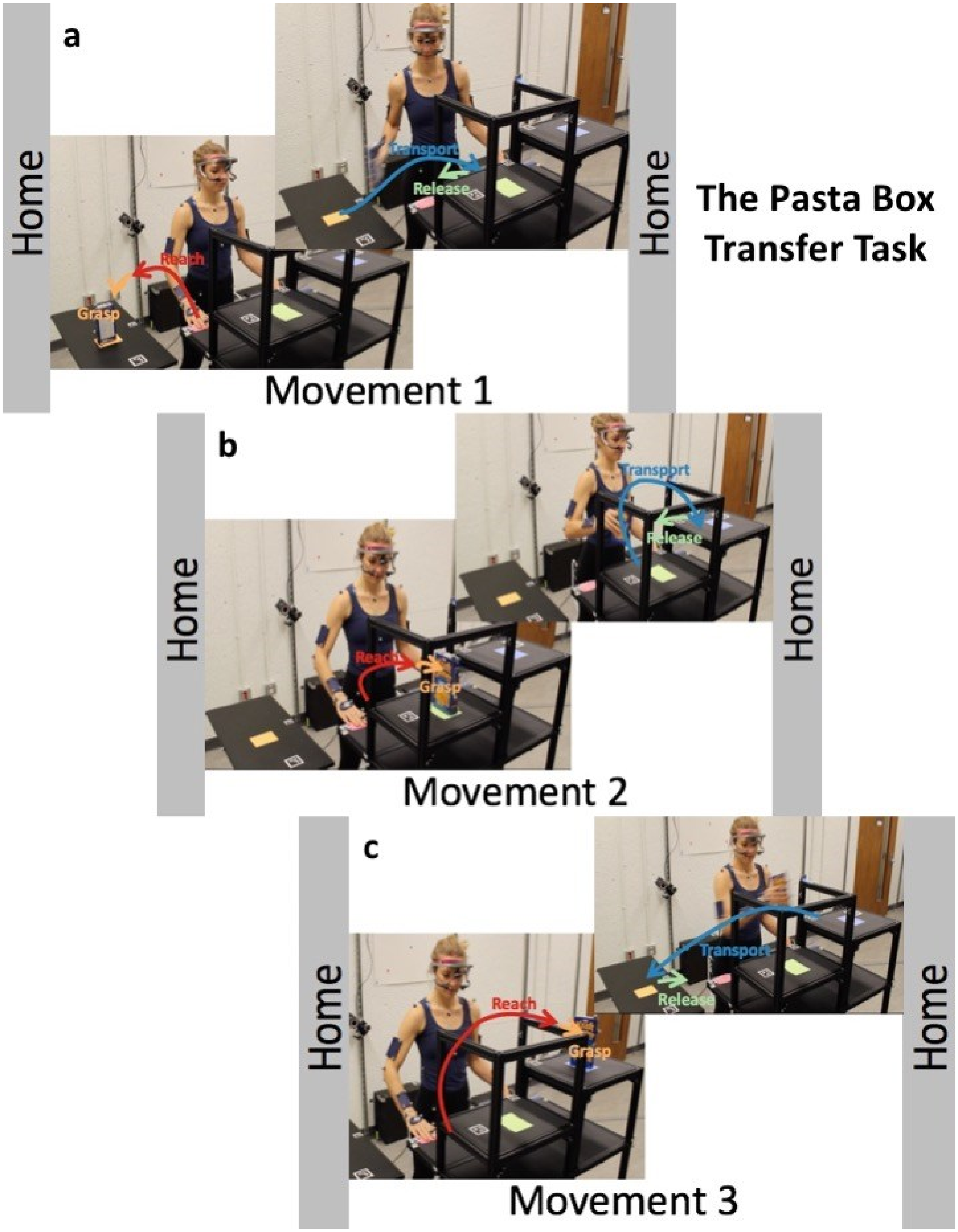

Figure 9: The Pasta Box Transfer Task includes Reach, Grasp, Transport, and Release of a Pasta box at 3 targets. a) Movement 1: pick-up from side cart Start/End Target and put-down on Mid Shelf Target. b) Movement 2: pick-up from Mid Shelf Target and put-down on High Shelf Target. c) Movement 3: pick-up on High Shelf Target and put-down on Start/End Target.

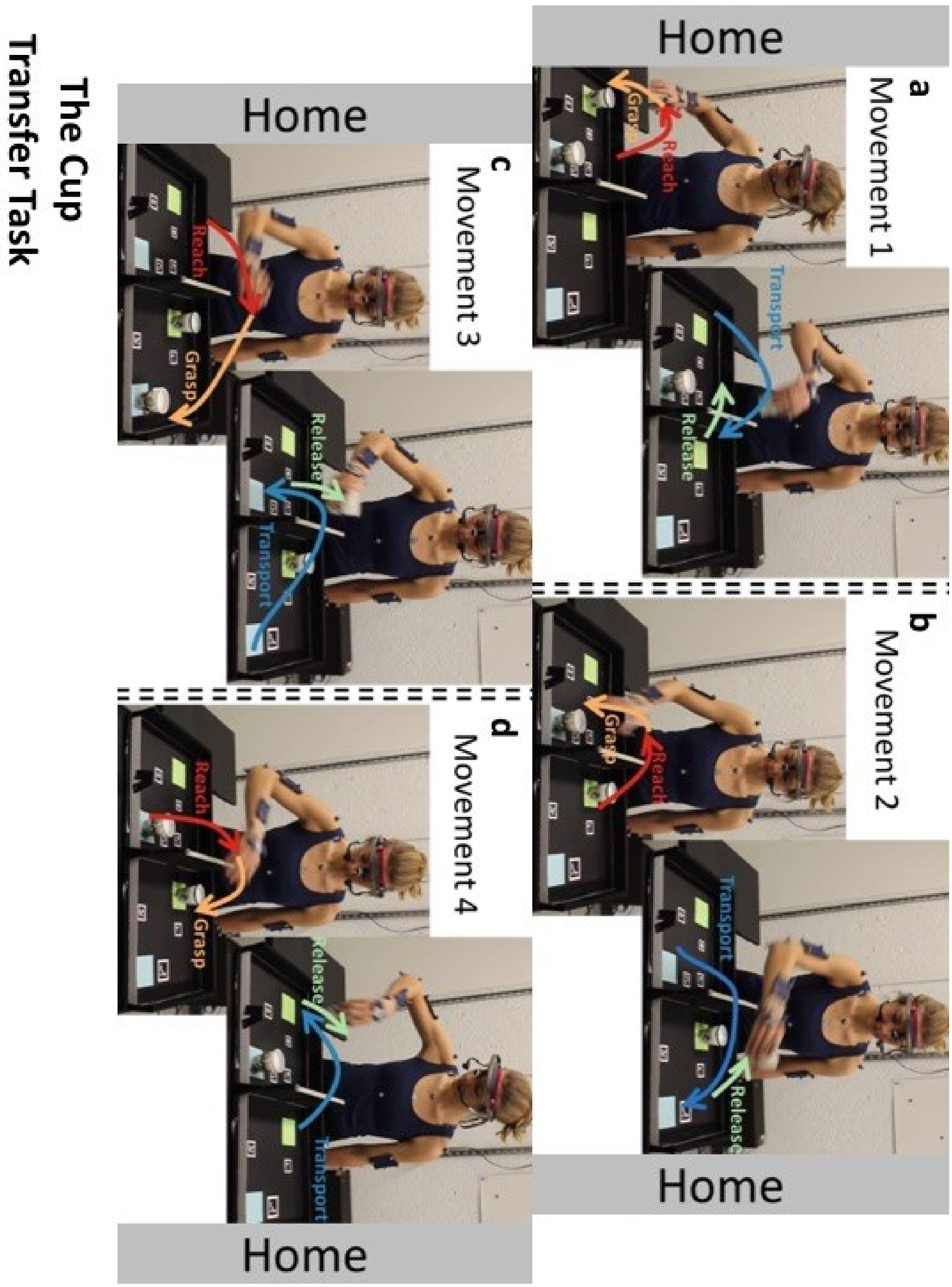

Figure 10: The Cup Transfer Task includes Reach, Grasp, Transport, and Release of 2 cups at 4 targets. a) Movement 1: pick-up of the Green cup with a top grasp at Near Target 1 and put-down at Near Target 2. b) Movement 2: pick-up of the Blue cup with a side grasp at Far Target 1 and put-down at Far Target 2. c) Movement 3: pick-up of the Blue cup with a side grasp at Far Target 2 and put-down at Far Target 1. b) Movement 4: pick-up of the Green cup with a top grasp at Near Target 2 and put-down at Near Target 1.

2.4 – Data Processing

To ensure accurate synchronization of the eye and motion tracking software programs, custom software was created to trigger the start and end of their recordings at the same time. The eye and motion tracking software programs recorded slightly different durations, likely due to the eye-tracking software starting its recording sooner than the motion tracking software. This was resolved by deleting the appropriate number of frames from the start of each eye-tracking recording to match that of its corresponding motion tracking recording. The average difference for correction was 0.124s. As well, any trial with a difference in the durations of the eye and motion tracking recordings greater than 0.400s was discarded (11 trials discarded). Before analyses on the recorded eye data could take place, post-processing was done. An algorithm in D-Lab that detects Quick Response (QR) codes (the black and white patches visible in Figures 7 c, d, e, f, and Figure 11 b, c, d, e) that had been placed in the task space enabled automatic visual fixation detection to areas or objects set by the experimenter, thus decreasing manual processing time. In Figures 6, and 7 (a) and (b), the QR codes have been removed for clarity. Following this, segmentation of the eye-tracking data based on the motion tracking data was carried out, allowing us to calculate measures of visual fixation to different Areas of Interest in the scene at different phases of movement.

2.4.1 - Post-processing of video-based eye data

2.4.1.1 – Pupil Autodetection Correction

The first post-processing step of the eye-tracking data was the adjustment of misdetected pupils. The D-Lab software offers quite accurate automatic detection of the user's pupils, however, periodically it will incorrectly assign the pupil to be in the user's eyelashes, or not assign a pupil at all. An example of this can be found in Figure 11 (a). Using the D-Lab software, the experimenter scanned through each trial of each participant to locate and correct any of these mistakes. As would be expected, the minimization of misdetected pupils was maximized by optimal placement of pupil cameras, described in the experimental setup. See Figure 8 (a) for a visual description of a pupil which has not been detected, and Figure 8 (b) for an accurate pupil detection.

2.4.1.2 – Calibration Adjustment

Although the calibration process described in detail above is quite accurate, to attain our best estimate of fixation location we also corrected the initial fixation position on every trial. In other screen-based eye-tracking studies, this is known as "drift-correction" and can occur automatically on every trial. In both tasks described above, participants were instructed to begin and end each trial with their eyes fixated on the Neutral eye position. This enabled the experimenters to notice any offset, or drift, from where each participant's fixation was in relation to the Neutral eye position, and correct it if necessary. Correction was carried out by shifting the fixation location of each participant to the Neutral eye position at the start of every trial, which would offset the rest of the trial by the same distance and direction. See Figure 11 (b) for an example of an offset fixation calibration, and the corresponding adjustment of this calibration in Figure 11 (c).

2.4.2 – Data Segmentation

2.4.2.1 Both Condition: Motion Tracking Segmentation with full Motion Capture

To identify each Object Related Action (ORA), and its component Reach, Grasp, Transport and Release phases, we used the motion capture data to conduct the following steps.

First, the velocities of the hand and objects (Pasta box for the PB task, and Green and Blue cups for the C task) were calculated, as well as the grip aperture of the hand (distance between the thumb and index finger markers), for all trials of all subjects. The earliest and latest peaks of the entire hand velocity profile for each trial were used to calculate the beginning and end of each trial, respectively. To be considered a peak, there must be a prominence of at least 300mm/s on both sides of the peak without encountering the end of the signal or another, larger peak. From the first peak in the trial, working backwards, the first instance that the velocity profile drops below 5% of this peak value marks the beginning of the Reach phase of Movement 1. It is assumed the last peak in the entire velocity profile of the trial is the participant's hand returning to Home. So, from this peak, working forwards, the first instance that the velocity profile drops below 5% of this peak value marks the end of the final movement (3 in PB, 4 in C).

Next, the object velocity profiles are analyzed. For the PB task, three object velocity peaks correspond to the three transport phases, while four object velocity peaks (2 for each cup) correspond to the four transport phases in the C task. The start and end of each Transport phase is calculated by moving forward and backward from these object velocity peaks to the first instance in either direction where the object velocity profile falls below 5% of this peak.

For the PB task, the second Reach phase is calculated by searching for two hand velocity peaks between the end of the first Transport phase and the beginning of the second Transport phase. These two hand velocity peaks correspond to the hand returning to the Home area after the first object movement, and leaving the Home area to begin the second object movement. Working backward from the second of these two hand velocity peaks to the first instance that the velocity falls below 5% of this peak marks the beginning of the second Reach phase. Similarly, for the third Reach phase of the PB task, the same process was used, but the two hand velocity peaks between the end of the second and the beginning of the third Transport phases were analyzed.

For the C task, the same process as above was employed, using the two hand velocity peaks between the end of the second and beginning of the third Transport phases to calculate the third Reach phase, as the hand returns to the Home area after the second object movement before beginning the third object movement.

To calculate the Grasp and Release phases, a two-step process was used. Step 1 used peak grip aperture. Specifically, working backwards from the beginning of each Transport phase a peak in the grip aperture profile marked the beginning of the Grasp. That is, from this point on to the start of Transport, the hand was closing on the object. Similarly, searching forward from the end of each Transport phase, the next peak in the grip aperture profile marked the end of the Release phase. That is, from the end of the Transport to this point represents the hand opening after letting go of an object. Step 2 of the segmentation process calculated the location of the hand at each of the Grasp and Release positions, relative to the centre of the Area of Interest (AOI) on which the object was either being picked up from or being dropped off at. An average of these distances was calculated for each participant, and then these were averaged to create a set of values defining Grasp and Release for each instance in both tasks. These values can be found in Table 1. Finally, these values were used to recalculate the Grasp and Release phases in both the PB and C tasks in the following manner: From the start of the Transport phase, the point at which the distance of the hand relative to the centre of the AOI is greater than the set distance for this phase, the Grasp phase begins. The Release phases were calculated in a similar manner, except that moving forward from the end of each Transport phase, the point at the hand relative to the centre of the AOI is greater than the set distance for this phase, the Release phase ends.

The reason we implemented this two-step process – using the grip aperture to define distances, but then only relying on distances to segment the data – is that it allows us more consistency in dealing with participants who may not have grip aperture data. While not the focus of this thesis, we also work with prosthetic limb participants where grip aperture is not reliable, so needed to create a version of segmentation that would work in all cases.

2.4.2.2 Eyes Only Condition: Motion Tracking Segmentation with reduced Motion Capture

Because the Eyes Only Condition did not have markers on the objects, we were unable to use the same segmentation algorithm as in the in-depth analysis of the individual tasks. So, to compare the Eyes Only Condition to the Both Condition, we used a simpler strategy. Using custom Matlab scripts, the location of the hand marker (Eyes Only) or centre of the hand plate (Both) at certain phases of each task were recorded for each participant, and were then used to segment each participant's data into Reach, Grasp, Transport, and Release phases for each of the PB and C tasks for both the Eyes Only and Both Conditions.

For the PB task, on the first of each participant's trials, the location of the hand marker was noted when the marker was stationary at all three of the grasp locations, as well as at the Home location. Then, a spherical Region of Interest (ROI) was created around each of these saved locations with a 10cm diameter. These ROI locations were then applied to all the other trials for that particular condition for that participant, and were used in conjunction with the hand marker, to segment the data into phases of movement. For example, for Movement 1, the hand marker would begin each trial inside the Home ROI, and when the participant moved their hand at the beginning of the trial, the hand marker would leave the Home ROI, thus starting the Reach phase of Movement 1. When the participant's hand marker entered the ROI at the Start/End Target, the Reach phase of Movement 1 would end and the Grasp phase would begin. When the participant had grasped the Pasta box, and moved their hand to begin transporting it to the Mid Shelf Target, the hand marker exited the ROI at the Start/End Target, thus beginning the Transport phase of Movement 1. When the participant's hand marker entered the ROI at the Mid Shelf Target, the Transport phase of Movement 1 ended, and the Release phase began. Finally, the Release phase ended as the participant's hand marker exited the ROI at the Mid Shelf Target to return to Home. This pattern continued for the other two movements in the PB Task.

Similarly, in the C task, ROIs with diameters of 10 cm were created at each grasp location from the first trial of each condition for each participant. These locations were then applied to the remaining trials for each participant, and used to segment the data into the Reach, Grasp, Transport, and Release phases for each of the four movements.

To ensure that our comparison of Eyes Only to the Both Condition would not be affected by differential segmentation, we also used this ROI strategy to segment the Both data (for this comparison only). While carrying out this ROI segmentation, four participants were not able to be included as the Grasp and Release locations did not consistently fall within the ROI boundaries during at least one of the four Task X Condition combinations. For this portion of the study, a custom visualization and analysis tool was employed that used the X and Y coordinates from the eye-tracking data, and the 3D motion capture coordinates to create a 3-dimensional gaze vector.

To pre-empt our results, there were no major differences found between the 16 participants who remained in the Eyes Only versus Both data. Thus, the majority of this thesis examines the manually corrected, video-based, Both data from all 20 participants and relies on the velocity-based segmentation.

Table 1: Defining Distances for Grasp and Release			
	Movement	Phase	Distance (mm)
Pasta	1	Grasp	44.31
	1	Release	16.29
	2	Grasp	25.48
	2	Release	21.24
	3	Grasp	24.01
	3	Release	16.25
Cups	1	Grasp	13.92
	1	Release	14.72
	2	Grasp	26.98
	2	Release	27.89
	3	Grasp	27.66
	3	Release	18.37
	4	Grasp	20.44
	4	Release	13.23

Table 1: The distances in the PB and C tasks of the Both Condition used to define the Grasp and Release phases of movement.

2.4.3 Areas of Interest (AOIs)

2.4.3.1 – QR Code Detection

As can be seen in Figures 7, 9, 10, and 11, Quick Response (QR) codes were placed in the scene. The D-Lab software was built to detect different QR patterns, which could then be used to assign AOIs in the scene, which drastically decreases the manual data analysis labour time. Before this, however, an "Exhaustive" QR detection was performed on every trial, for optimal automation of data analysis. Notice in Figure 11 (b, c, d) the red squares surrounding the black and white QR codes signifying that the QR code has been detected by the D-Lab software.

2.4.3.2 Creation of AOIs

The D-Lab software enables the experimenter to assign Areas of Interest (AOIs) relative to one or more QR codes placed in the scene. This is done by assigning specific QR codes to use for a specific AOI in the scene, and clicking with a computer mouse to create the desired shape of the AOI. This was carried out for the Neutral eye position, the Home area, each drop-off location, and objects during pick-up for both tasks. Next, experimenters used the built-in calculation of fixations algorithm in the D-Lab software to these AOIs, which yielded any instance that a participant's fixation location in the scene intersected any of these AOIs for 120ms or more. Figure 11 (d) shows a fixation registered to an AOI during the PB task. As well, more fine-tuned algorithms like elimination of eye blinks (any time where both pupils are missing for more than 300ms) and filling of short gaps (120ms or less) between fixations to the same AOI were used (D-Lab 3.0 Manual, 2017). Although this saved some manual labour, the experimenters still had to manually verify and correct many occurrences where the software failed to automatically register AOI fixations, as any instance in which the QR codes were obscured or blurred (eg. during head movement, see Figure 11e) failed to yield accurate fixation values. In addition, the D-Lab software enabled manual coding of AOIs that were not able to be applied to specific QR codes, like a moving object. This was done when a participant fixated on their own hand or an object in their hand. These were rare in most participants, but still required experimenters to go through each frame to verify if any hand fixations had occurred.

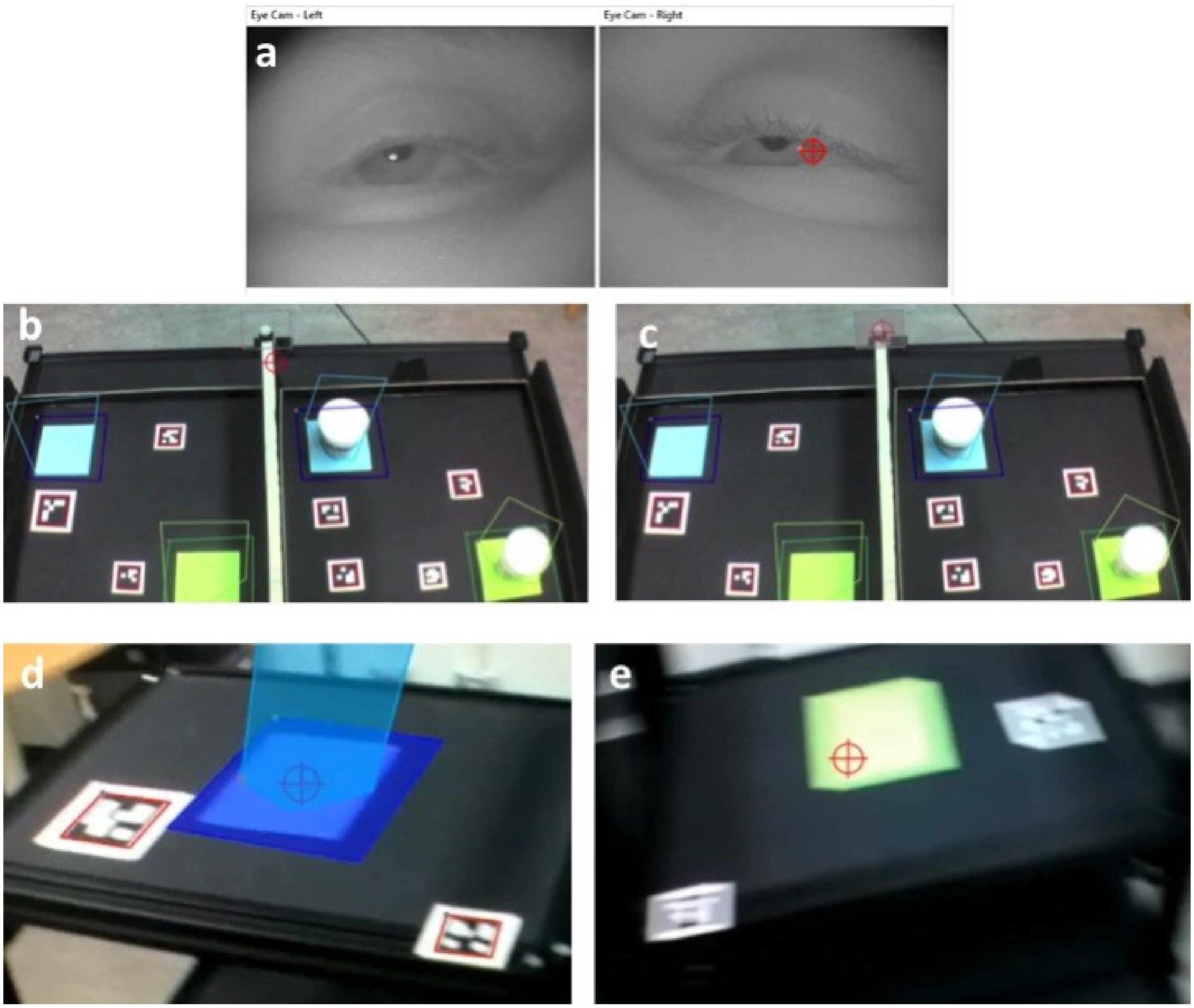

Figure 11: Camera images from the D-Lab software. a) Frames from the pupil cameras where the pupil has not been detected (left), and a dark spot on the iris has been misdetected as the pupil (right). b) Crosshairs of fixation direction, before it has been aligned to the neutral eye position. c) Crosshairs of fixation direction, after it has been aligned to the Neutral eye position. d) Crosshairs within an AOI registering a fixation in the D-Lab software. e) Blurred QR codes due to head movement, causing a fixation to go unregistered.

As we are interested in overt fixations to areas relevant to object interactions, we limited our data analysis to specific regions during each ORA. Thus, the AOIs within each phase are the current location being acted on by the hand (Current), the future location that the hand will act upon when it has completed its current action (Future), and the hand itself or an object being moved by the hand when no other AOI is being fixated (Hand in Flight). Given that each task had several discrete object interaction movement phases, Current and Future AOIs are not static, but are specifically assigned to that movement phase, as outlined in Tables 2 and 3. Additionally, because the participant's hand is in close proximity to the Current AOI during the Grasp and Release phases, the hand is included in the Current AOI during these phases.

Table 2: Pasta Box Transfer Task Segment AOIs				
	Segment	Current	Future	Hand
Movement 1	Reach	Start/End Target, Pasta Box	Mid Shelf Target	Hand
	Grasp	Start/End Target, Pasta Box, Hand	Mid Shelf Target	
	Transport	Mid Shelf Target	Home sticker	Hand, Pasta Box
	Release	Mid Shelf Target, Pasta Box, Hand	Home sticker	
Movement 2	Reach	Mid Shelf Target, Pasta Box	Top Shelf Target	Hand
	Grasp	Mid Shelf Target, Pasta Box, Hand	Top Shelf Target	
	Transport	Top Shelf Target	Home sticker	Hand, Pasta Box
	Release	Top Shelf Target, Pasta Box, Hand	Home sticker	
Movement 3	Reach	Top Shelf Target, Pasta Box	Start/End Target	Hand
	Grasp	Top Shelf Target, Pasta Box, Hand	Start/End Target	
	Transport	Start/End Target	Home sticker, Neutral eye position	Hand, Pasta Box
	Release	Start/End Target, Pasta Box, Hand	Home sticker, Neutral eye position	

Table 2: Definitions of which relevant objects make up the Current, Future, and Hand in Flight AOIs during each of the Reach, Grasp, Transport, and Release phases of the 3 ORAs of the PB Task.

Table 3: Cup Transfer Task Segment AOIs				
	Segment	Current	Future	Hand
Movement 1	Reach	Near Target 1, Green cup	Near Target 2	Hand
	Grasp	Near Target 1, Green cup, Hand	Near Target 2	
	Transport	Near Target 2	Far Target 1, Blue cup	Hand, Green cup
	Release	Near Target 2, Green cup, Hand	Far Target 1, Blue cup	
Movement 2	Reach	Far Target 1, Blue cup	Far Target 2	Hand
	Grasp	Far Target 1, Blue cup, Hand	Far Target 2	
	Transport	Far Target 2	Home sticker	Hand, Blue cup
	Release	Far Target 2, Blue cup, Hand	Home sticker	
Movement 3	Reach	Far Target 2, Blue cup	Far Target 1	Hand
	Grasp	Far Target 2, Blue cup, Hand	Far Target 1	
	Transport	Far Target 1	Near Target 2, Green cup	Hand, Blue cup
	Release	Far Target 1, Blue cup, Hand	Near Target 2, Green cup	
Movement 4	Reach	Near Target 2, Green cup	Near Target 1	Hand
	Grasp	Near Target 2, Green cup, Hand	Near Target 1	
	Transport	Near Target 1	Home sticker, Neutral eye position	Hand, Green cup
	Release	Near Target 1, Green cup, Hand	Home sticker, Neutral eye position	

Table 3: Definitions of which relevant objects make up the Current, Future, and Hand in Flight AOIs during each of the Reach, Grasp, Transport, and Release phases of the 4 ORAs of the C Task.

2.5 Dependent Measures

Given the specific objectives of our study, for dependent measures we chose to use the duration of each of the phases outlined in the previous section, the number of fixations to each of the Current, Future, and Hand in Flight AOIs in each phase, along with the percentage of time fixated on each of these AOIs in each phase. As can be seen from Tables 2 and 3, during Grasp and Release phases, any fixations to a participant's own hand were included in the Current AOI, while during Reach, the hand was its own Hand in Flight AOI, and during Transport the hand and the object being transported made up the Hand in Flight AOI. Any dependent measure involving AOIs is calculated using the timeline of fixations to these AOIs exported from D-Lab and the triggers from the motion tracking data defining each phase of movement.

In addition to these three phase-specific dependent measures, another measure was used to compare the difference in time between the eyes beginning a fixation on an object, and the participant's hand arriving at the object. This was calculated for each ORA in both tasks (3 in PB, 4 in C), but only for the Both Condition, as the Eyes Only Condition used a less exact form of segmentation.

2.5.1 – Duration

The duration (s) of each phase was calculated using simple arithmetic on the time points which indicated the transition from one phase to the next from the motion tracking software.

2.5.2 – Number of Fixations

The number of fixations (#) to an AOI is the number of distinct continuous fixations to this AOI, without shifting away from it. This was calculated for each of the Current, Future, and Hand in Flight AOIs as outlined above.

2.5.3 – % Fixation Time

The percent fixation time (%) to an AOI is the amount of time fixated on an AOI in a phase divided by the total duration of that phase, multiplied by 100 to be represented as a percentage. This was calculated for Current, Future, and Hand in Flight AOIs as applicable. Note that, the results presented here are averages of the trials, for each participant, where a fixation occurred.

2.5.4 – Eye-Hand Latency

The difference in duration (s) between the first fixation to Current in the Reach or Grasp phase and the beginning of the Grasp phase. This was calculated for each ORA in both tasks in the Both Condition.

3 - Results

3.1 – Overview of Statistical Analysis

For each subject, each of the dependent measures was calculated on every trial, then averaged. For all measures except Eye-Hand Latency, this meant that each participant had one value for each combination of Condition (Eyes Only and Both), Task (PB and C), Movement (3 in PB, 4 in C), and Phase (Reach, Grasp, Transport and Release). For Eye-Hand-Latency, each participant had one value for each ORA. When reported, repeated measures ANOVAs were analyzed with the Greenhouse-Geisser sphericity correction, with significance being marked at $p < 0.05$. Following any significant ANOVA result, post-hoc pairwise comparisons were conducted using a Bonferroni correction with a corrected $p < 0.05$ marking a significant effect.

Four broad analyses were conducted. First, we examined for any effects of wearing additional motion capture markers by comparing the Eyes Only to the Both Condition. Then, to understand the specific eye gaze patterns within each task, using only the Both Condition, each of the PB and C tasks were analyzed individually. Finally, to derive any task specific commonalities or differences we again analyzed the Both Condition and compared the PB and C tasks. Statistical comparisons were run on each dependent measure outlined in the previous section: Phase Duration, Number of Fixations to Current, % Fixation Time to Current, Number of Fixations to Future, % Fixation Time to Future, Number of Fixations to Hand in Flight, % Fixation Time to Hand in Flight, and Eye-Hand Latency. Note that during any Grasp and Release phase, a fixation to the participant's own hand is included in the Current location, leaving Hand in Flight measures to be included only in the Reach and Transport phases.

3.2 – Eyes Only versus Both Conditions

In the first analysis, the Eyes Only Condition was compared to the Both Condition for each dependent measure (excluding Eye-Hand Latency) using three-factor (Condition, Movement, Phase) repeated measures ANOVA tests for each of the PB and C tasks. The purpose of this comparison is to highlight differences due to the addition of motion capture markers, therefore, only those results with either a main effect or interaction effect involving Condition are reported. Few differences were found to be significant. Most notably, participants took slightly more time when wearing additional motion tracker markers. That is, the Both Condition saw participants take 0.024s (4.4%) longer for each phase of movement than the Eyes Only Condition. Interestingly, this increase is seen in both tasks to the same magnitude which can be viewed in Figure 12, with full results displayed in Table 6.

As well, in the PB task, it was found that both the Number of Fixations and the % Fixation Time to the Future targets were statistically significantly less in the Both Condition (0.069 Fixations; 1.3%) than in the Eyes Only Condition (0.086 Fixations; 1.7%). Perhaps due to the larger movements in the PB task, and the slightly longer duration of movements in the Both Condition, participants initiated saccades towards the Future targets at the same time as in the Eyes Only Condition, but their eyes did not arrive at the Future targets soon enough to register a fixation to the target before the phase of movement changed and the Future target became the Current target. This can be understood as participants moving slightly less freely when wearing the motion tracking equipment resulting in fewer numbers of fixations, and less time of fixation to Future targets. Full results of this analysis can be found in Table 6 as well. It is worth stating that, although these results are statistically significant, the magnitude of these values are quite small (0.4% difference in % Fixation Time of Duration values approximately half a second), and should not be taken as an indication that great behavioural changes were caused by the motion tracking equipment.

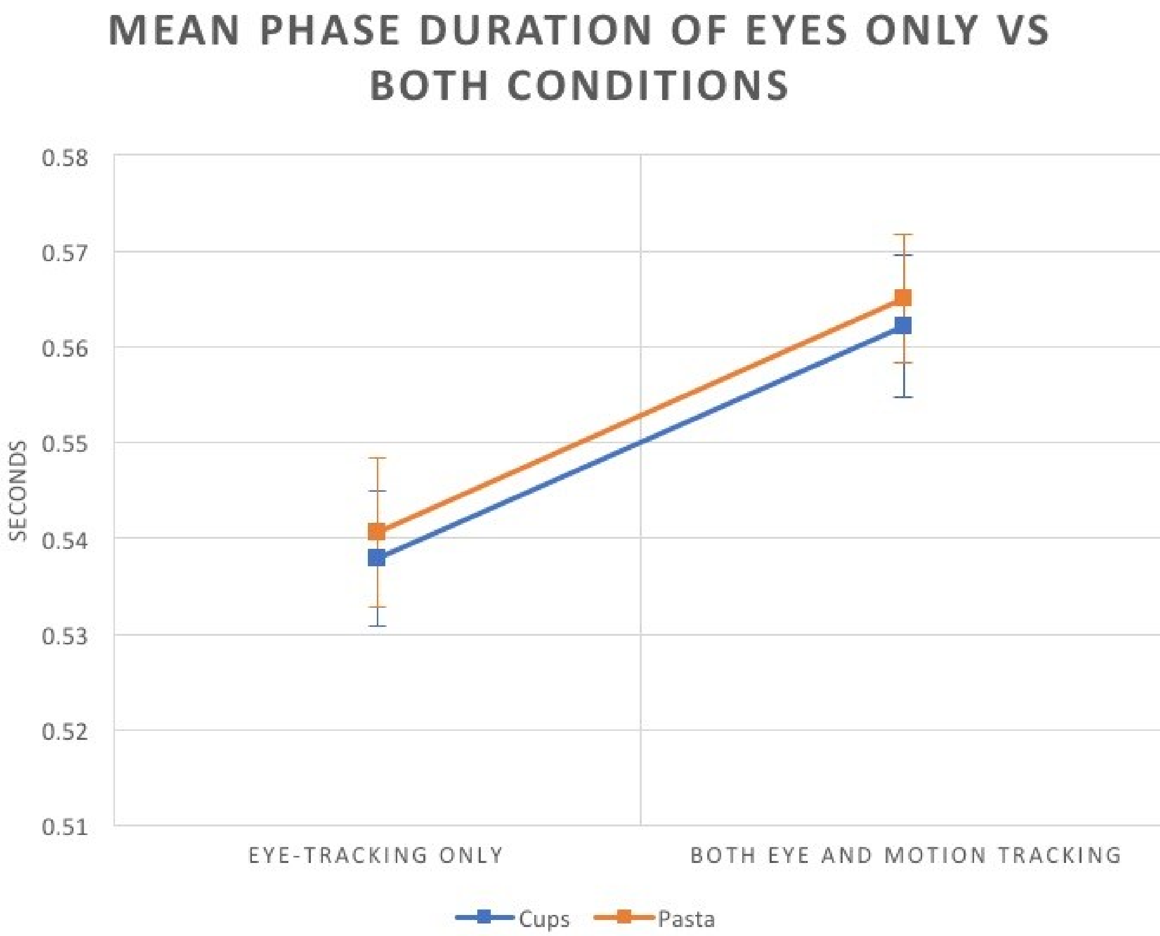

Figure 12: The average duration of phase of movement for the Eyes Only Condition, and the Both Condition. For both the Cups and Pasta tasks, wearing motion capture markers resulted in slightly longer duration of movements. The error bars shown here were calculated by first obtaining each participant's standard error, then calculating a mean value across all participants.

3.3 – Normative Eye Behaviour during Sequential Object Movement

3.3.1 – Pasta Box Transfer Task

A general description of the eye movement behaviour was able to be summarized from the results (as shown in Tables 4 and 8, and Figures 13 and 15), and is instructive to review in relation to our hypotheses.

3.3.1.1 – Description of eye gaze behaviour

ORA 1: Initially, at the start of the trial, the hand is on Home and the eyes are on Neutral. When the go cue is presented, the participant takes a moment to plan the complex movement sequence, before initiating a body turn toward the first Pasta box location on the side table. Due to this movement planning delay and lack of momentum, the participant takes longer to complete this first reach to the Start/End Target than other reaches. The participant fixates on the Pasta box for roughly a quarter of a second before their hand arrives. This short Eye-Hand Latency can be explained by the Start/End Target requiring the participant to turn and look down, which often caused the pupils to be lost by the eye-tracker and/or the fixation point to be offset from the Start/End Target AOI. Also, since the Start/End Target location is initially out of sight, there is more uncertainty in object Grasping, prolonging the Grasp phase. In addition, the turning the body to the side table itself takes additional time, thereby reducing the length of time a participant is fixating on the Pasta box in the first Reach and Grasp. From the turned body position at the end of this first Grasp, it is relatively easy for the participant to turn their head back to forward-facing and fixate on the Mid Shelf Target as they start to Transport the box to that location. As they bring the Pasta box onto the Mid Shelf Target, their eyes stay fixated on the target until the end of the Release, at which point they look down toward the Home area while reaching there. Often, given the low task demands of the return of the hand to Home, this fixation is brief (or absent) and the eyes quickly shift back to the Pasta box at the Mid Shelf Target to plan to pick it up.

<u>ORA 2</u>: The second Reach, from Home towards the Pasta box, is the shortest in this task. When reaching towards the Pasta box, the participant's eyes will begin their fixation on the Pasta box well ahead of the hand starting to Grasp it (roughly half a second earlier). The eyes stay fixated on the box as the hand approaches it and then shift to fixate on the High Shelf Target during the Grasp phase. The close proximity of these two targets (Mid Shelf to High Shelf), and the lack of a required body turn mean that the participant is able to fixate sooner on this next drop-off location than in other ORAs. The second Transport takes longer than the first one because the participant must move the Pasta box out of a confined location around a barrier, and therefore takes more care by moving slower. During this second Transport, the participant also fixates slightly more on their own hand (and/or the Pasta box), which may be due in part to the careful movement, but also because the hand is higher in their field of view and moving toward their body. This particular alignment of the hand moving toward the head naturally results in more fixations (and/or more recorded fixations) toward its location. As the participant Releases the Pasta box on the High Shelf Target, their gaze lingers at this location longer than other ORAs. We attribute the prolonged fixation at this drop-off location to the configuration of their body: their arm is across their body at a high location making it difficult to turn their body and head to fixate on the next location of their hand, the Home area. As a result, we also see a corresponding dip in the fixations to the Home area during this Release phase, which also reflects its total distance away from this drop-off location (largest Release to Home distance).

ORA 3: The third Reach from Home back to the High Shelf is the longest Reach of this task and as a result, takes longer and results in less time fixating on the pick-up location. When reaching towards the Pasta box, the participant's eyes will begin their fixation on the Pasta box well ahead of the hand starting to Grasp it (roughly half a second earlier). During Grasp, we find that participants spend more time fixating on the Pasta box, likely for the body configuration reason outlined above. That is, they cannot easily turn their body and head towards the drop of location on the side table since their arm is high and across their body. The final Transport of the Pasta box from the High Shelf, high on the left, to the Start/End Target, low on the right, takes the longest because it is the furthest movement in the entire experiment (across both tasks). This long duration, and corresponding body turn, results in substantially fewer fixations toward the drop-off location. Interestingly, during the body turn and subsequent sweep of eye gaze across the entire workspace, the scan path crosses near to both the Neutral eye position and the Home area. For this final movement, these count as the Future (and final) eye gaze locations, and occasionally participants will briefly fixate on them during this last Transport. For the same reasons as outlined before (both that the line of sight is aligned with the hand that is moving toward the body, and, that it is difficult to turn the head and body while reaching across it) the participant also fixates much more on their own hand at the start of this Transport. The final Release of the Pasta box on the Start/End Target is longer, likely because it is lower and there is more uncertainty about its exact location since it is largely out of sight. This is accompanied by less time fixating on the box during this Release, in part because participants can easily turn their head to fixate on the nearby Home area as their arm is down and rotated laterally from their body.

3.3.1.2 – Detailed statistical results

For each dependent measure in this task, except those of Hand in Flight and the Eye-Hand Latency, a two-factor (Movement, Phase) repeated measures ANOVA was carried out with three levels coding for Movement (1, 2, 3), and four levels coding for Phase (Reach, Grasp, Transport, Release). For each dependent measure for Hand in Flight, a two-factor repeated measures ANOVA was carried out with three levels coding for Movement (1, 2, 3), and two levels coding for Phase (Reach, Transport). Wherever significant, an interaction of Movement and Phase was followed up with simple main effects single-factor ANOVAs of Movements for each Phase. Post-hoc tests were run for significant main effects and simple main effects which compared all possible pairwise comparisons of the relevant factor (Movement or Phase for main effect, and Movement for simple main effects). For Eye-Hand Latency, a one-factor (Movement) repeated measures ANOVA was carried out with three levels coding for Movement (1, 2, 3), with a post-hoc test run due to a main effect of Movement. Many significant differences were found which can be viewed in Table 4 (Phase Dependent Measures) and Table 8 (Eye-Hand Latency). This section will be structured by first looking at the Duration of each phase of movement, then by looking at the metrics (the Number of Fixations and % Fixation Time) within each of the AOIs (Current, Future, and Hand in Flight) during each phase, and finally discussing the Eye-Hand Latency in each Movement. Graphical representations of each measure can be found in Figures 13 and Figure 15 (a).

In the next section, we will discuss in detail the multitude of reasons for the effects found in the PB task. Here is a brief description of each of the possible causes of the effects, which are referenced in italics in parentheses below. Some effects were caused by the distance over which the participant was required to move (*Movement distance*). As well, some effects can be attributed to a lag in movement speed due to the complexity of movement planning. Near the beginning of both tasks, participants take more time to complete movements, as they begin from a sedentary position and must compute movement information about the entire movement sequence. Thus, some of this computation time spills into the first movement, resulting in it taking longer. By comparison, subsequent movements (second, third etc.) can have their planning completed during the first (or earlier) movements and thus do not have any resulting planning delays (*Complexity of movement planning*). The location of the Start/End Target caused participants to exhibit intriguing behaviours. The exact causal element of which is not easily deducible, as this target was lower than the other two targets, was outside the field of view of the participants, and required turning of the head and body to interact with (*Start/End Target location*). Variations in biomechanics caused differences in eye movement behaviour, that are most notable when participants were interacting with the High Shelf Target as it forces their arm to cross their midline, hindering the range of motion of their neck (*Biomechanics*). The grasp and release apertures used by participants varied at different target locations, which likely changed the duration of these phases of movement somewhat (*Apertures*). Furthermore, the Mid Shelf Target forced participants to move their hand into a confined location as there were barriers on either side of the shelf, and horizontally above the shelf which could have caused behavioural effects (*Confined location*). The height of the shelves and the proximity of certain areas of interest to one another are potential causes of some of the effects (*Shelf height; Proximity rule*), as is the path that the eyes traveled from one area to another, and the reconfiguration of the location of certain areas of interest after the body turned to the Start/End Target (*Eye movement path; Reconfiguration of body location*).

3.3.1.2.1 – Duration

As can be seen in the full results of the PB task, Figure 13 (a), there is an obvious trend of Transport taking participants the most time, followed by Reach, with Grasp and Release taking the shortest time. Within the Reach phase, further distances traveled by the hand correlate with larger durations (*Movement distance*). In addition to longer distances resulting in longer times, we also find that the first movement is disproportionally longer (e.g. Movement 1 > Movement 2). This is attributed to the complexity of movement planning, due to the Reach of Movement 1 being at the start of the trial (*Complexity of movement planning*). Interestingly, the movement planning burden being highest at the start of the task, aligns with previous literature as it is known that planning a longer, more complicated movement, takes longer to initiate (Henry & Rogers, 1960). Interestingly, this could extend as far as the Grasp phase of Movement 1, as it is significantly slower than the other two phases. However, in addition to being first, the Grasp of Movement 1 is likely slower because the Start/End Target is out of the field of view of participants at the beginning of the Movement (*Start/End Target location*), and requires participants to turn their body (*Biomechanics*). Notably, this also results in the largest grasp distance (see Table 1) which of course also contributes to the lengthening of Grasp 1 (*Apertures*). Grasp during Movement 3 takes longer than that of Movement 2 likely because it forces participants to reach across their body in a more strained manner, while Movement 2 is on the same side of the body as the hand (*Biomechanics*). The Durations of the Transport phases of the Movements are heavily dictated by the distance the object is moved (*Movement distance*), while also being influenced by the risk of contacting the cart during Movement 2 (*Confined location,* for example Movement 3 > Movement 2 > Movement 1). Finally, the Release phase of

Movement 3 is longer than the other two likely due to the drop-off location of the Pasta box (*Start/End Target location*).

3.3.1.2.2 – Current

The Number of Fixations to the Current AOIs hover around 1.0 for all Phases of Movement. Keep in mind that, where Current stays the same between consecutive phases (e.g. Reach→Grasp and Transport→Release), a single fixation will be counted once for each phase. The Start/End Target being out of the field of view likely causes the Reach of Movement 1 to trend towards having significantly fewer fixations than the other Movements (*Start/End Target location*). This is stated because the omnibus ANOVA yields a significant difference, but no pairwise test does. For the same reason, and to a much more significant degree, this leads to the % Fixation Time to Current during Reach of Movement 1 to be less than half as large as the other two movements. This reflects an important general principle – in cases where the head will need to turn to fixate a target, the target is unlikely to be fixated for as long prior to grasping it. The Reach of Movement 2 has the highest % Fixation Time and is significantly longer than that of Movement 3. We believe this is because the eyes arrive earlier at the Mid Shelf Target when leaving the Home area than they do to the High Shelf Target due to the closer proximity (e.g. distance between Mid Shelf Target and Home is smaller than the distance between the High Shelf Target and Home*, Proximity Rule*). Being more proximal allows the eyes to arrive earlier in a phase, increasing the overall % Fixation Time. Similar to the reasoning for Reach, the location of the Start/End Target leads to % Fixation Time of the Grasp at Movement 1 being lower than that of Movement 3, as well as Transport of Movement 3 (which also ends at the Side Table) being less than the other two (*Start/End Target location*). Finally, the % Fixation Time during Release of Movement 2 is larger than the other two movements, a finding we ascribe to the head being unable to turn towards the Home area as the arm is across the body and up high in a strained manner (*Biomechanics*). This is an interesting pattern we see across all the data in this thesis – when the arm is reaching across the body, and it makes it more awkward to turn the head, the eyes will linger at their current location of fixation – as if they are waiting for the arm to move, making the required head turn easier. See Figure 13 for graphical representations of the full values for the Number of Fixations (b) and the % Fixation Time (c) to Current.

3.3.1.2.3 – Future

As defined and segmented, fixations to Future locations happen rarely. Those that we do identify are often the result of the path of line of sight during specific movement phases crossing close to Future locations, and may not actually count as looks "ahead" but rather, coincidental "crossings through". For example, during the Transport phase of Movement 3, as the hand and eyes sweep down from the High Shelf Target to end at the Start/End Target, this trajectory passes very close to the location of the Future targets (Neutral and Home). It is difficult to say whether the resulting fixations to Future we recorded here are intentional, but, the result is a higher Number of Fixations and larger % Fixation Time (*Eye movement path*). The Number of Fixations and % Fixation Time to Future in the Release phase of Movement 3 is larger than that of Movements 1 and 2 which we believe are due to the proximity of the Current, Start/End Target, to the Future, Home area (*Proximity rule*), the ease with which the head can turn due to the arm being low and on its own side of the body (*Biomechanics*), and the need to re-centre your reference frame on the Home area (Future) after having turned it toward the side table (*Reconfiguration of body location*). See Figure 13 for graphical representations of the full values for the Number of Fixations (d) and the % Fixation Time (e) to Future.

3.3.1.2.4 – Hand in Flight

The Number of Fixations and % Fixation Time toward the Hand in Flight in the Transport phases are larger than those in the Reach phases. While low (<10% time fixating hand on average during Transport) this is consistent with the idea that there is some added benefit of getting visual feedback about the reliability of your grasp on an object while you are moving it. Within the Reach phases specifically, Movement 1 trends toward having significantly fewer fixations than in Movement 3, presumably because it is lower in the field of view and requires a turn of the body (*Shelf height, Start/End Target location*). This is stated because the omnibus ANOVA yields a significant difference, but no pairwise test does. Looking at the Transport Phase specific results, the general pattern is that the Number of Fixations are fewer and shorter for Movement 1 and more frequent and longer for Movement 3, with Movement 2 falling in the middle. We believe this is almost entirely driven by possible looks to the hand at the time of object pick-up. Specifically, these metrics are low for Movement 1 due to the Start/End Target location where the next movement easily sweeps across the body (*Biomechanics*), and high for Movement 3 because the shelf height for this movement brings the hand and object to be more central in the field of view (*Shelf height*) and moves somewhat toward the participant during early Transport. In addition, as mentioned earlier, at the start of Transport 3, the arm is across the body making it difficult for the participant to move their head and eyes to the drop-off location (*Biomechanics*), and as a result, the eyes linger at the pick-up location, sometimes long enough to be counted as a look during Transport. The Transport Phase of Movement 2 falls somewhere in the middle, with the shelf at a slightly lower height (*Shelf height*). It is also possible that any effects at this shelf location are due to it being a confined space (*Confined* *location*). See Figure 13 for graphical representations of the full values for the Number of Fixations (f) and the % Fixation Time (g) to the Hand in Flight.

3.3.1.2.5 – Eye-Hand Latency

As stated earlier, the Eye-Hand Latency values were calculated for each ORA of the PB task using the difference of the time of the first fixation to the Current AOI during the Reach or Grasp phase of the ORA and the time of the start of the Grasp phase of this ORA. As is clear from Figure 15 (a), the Eye-Hand Latency of Movement 1 (0.216s) is significantly shorter than those of Movements 2 and 3 (0.427s and 0.460s, respectively). The most obvious reason for this difference is the fact that the Current AOI in Movement 1 is the Start/End Target, which is out of the field of view of the participant (*Start/End Target location*). The eyes fixated on this AOI much later in the Reach phase than in the other two Reach phases. In addition, some participants were shorter, which necessitated the angle of the scene camera to be closer to 90° resulting in the Start/End Target being completely missed by the scene camera recording (*Start/End Target location; Shelf height*). Another likely reason that the eyes fixate later on the Pasta box in the first Reach phase than in other Reach phases is due to the difficulty the eye-tracker had with registering accurate fixations to targets requiring turning of the body (caused blurring of the scene camera disabling QR code detection by the D-Lab software and changing of the depth of objects resulting in offset fixation locations). Moreover, in Movement 1, the eyes and hand are starting at the same time, as it is the start of the task, which reduces the amount of time the eyes can fixate on the object before the hand arrives (*Complexity of movement planning*). In the other two ORAs, the eyes are starting ahead of the hand. Finally, the eyes are moving from a much further location (Neutral eye position) compared to the hand during Movement 1 than in Movements 2 and 3 (*Movement distance*). These taken together, result in the significantly shorter

Eye-Hand Latency of Movement 1. Ultimately, however, we expect that the Eye-Hand Latency for Movement 1 would be similar to Movement 2 and 3 (approximately half a second), if the eye-tracker was able to reliably capture the eye movements to this location.

Table 4: Pasta Box Transfer Task Dependent Measure Means

	Movement	Duration (s)			Number of Fixations to Current (#)			% Fixation Time to Current (%)			Number of Fixations to Future (#)			% Fixation Time to Future (%)			Number of Fixations to Hand in Flight (#)			% Fixation Time to Hand in Flight (%)		
		Mean	$F_{(6, 114)}$	Pair	Mean	$F_{(6, 114)}$	Pair	Mean	$F_{(6, 114)}$	Pair	Mean	$F_{(6, 114)}$	Pair	Mean	$F_{(6, 114)}$	Pair	Mean	$F_{(2, 38)}$	Pair	Mean	$F_{(2, 38)}$	Pair
Reach	1	0.686			0.934			31.382			0.031			0.441			0.003			0.003		
	2	0.526	**	2<<1; 3<1; 2<<3	0.999	*		79.299	**	3<<2; 1<<2; 1<<3	0.028	ns		0.445	ns		0.104	*		1.670	ns	
	3	0.656			0.994			68.632			0.000			0.000			0.146			1.746		
Grasp	1	0.284			0.974			80.793			0.025			0.394								
	2	0.152	**	2<<1; 3<<1; 3<<2	0.919	ns		86.040	*	1<3	0.075	ns		4.223	ns							
	3	0.181			0.980			93.221			0.000			0.000								
Transport	1	1.074			1.033			78.292			0.000			0.000			0.278			1.905		
	2	1.173	**	2<<3; 1<<3; 1<<2	1.015	ns		78.647	**	3<<2; 3<<1	0.000	**	1<<3; 2<<3	0.000	*	1<3; 2<3	0.678	**	2<3; 1<<3; 1<<2	8.560	1<<3; 1<<2	
	3	1.284			0.942			41.869			0.055			0.659			0.935			11.855		
Release	1	0.290			0.988			75.823			0.042			0.824								
	2	0.276	**	1<<3; 2<<3	0.990	ns		84.129	**	1<<2; 3<<2	0.008	**	1<<3; 2<<3	0.039	**	1<<3; 2<<3						
	3	0.396			0.901			63.144			0.425			9.106								

Table 4: The means of the dependent measures calculated for each Phase of the PB task, with statistical effects from a Movement X Phase interaction indicated by: ns = not significant, * = $p < 0.05$, ** = $p < 0.005$. As well, any significant pairwise contrasts of Movements within a Phase, and their direction are indicated by: < = $p < 0.05$, << = $p < 0.005$.

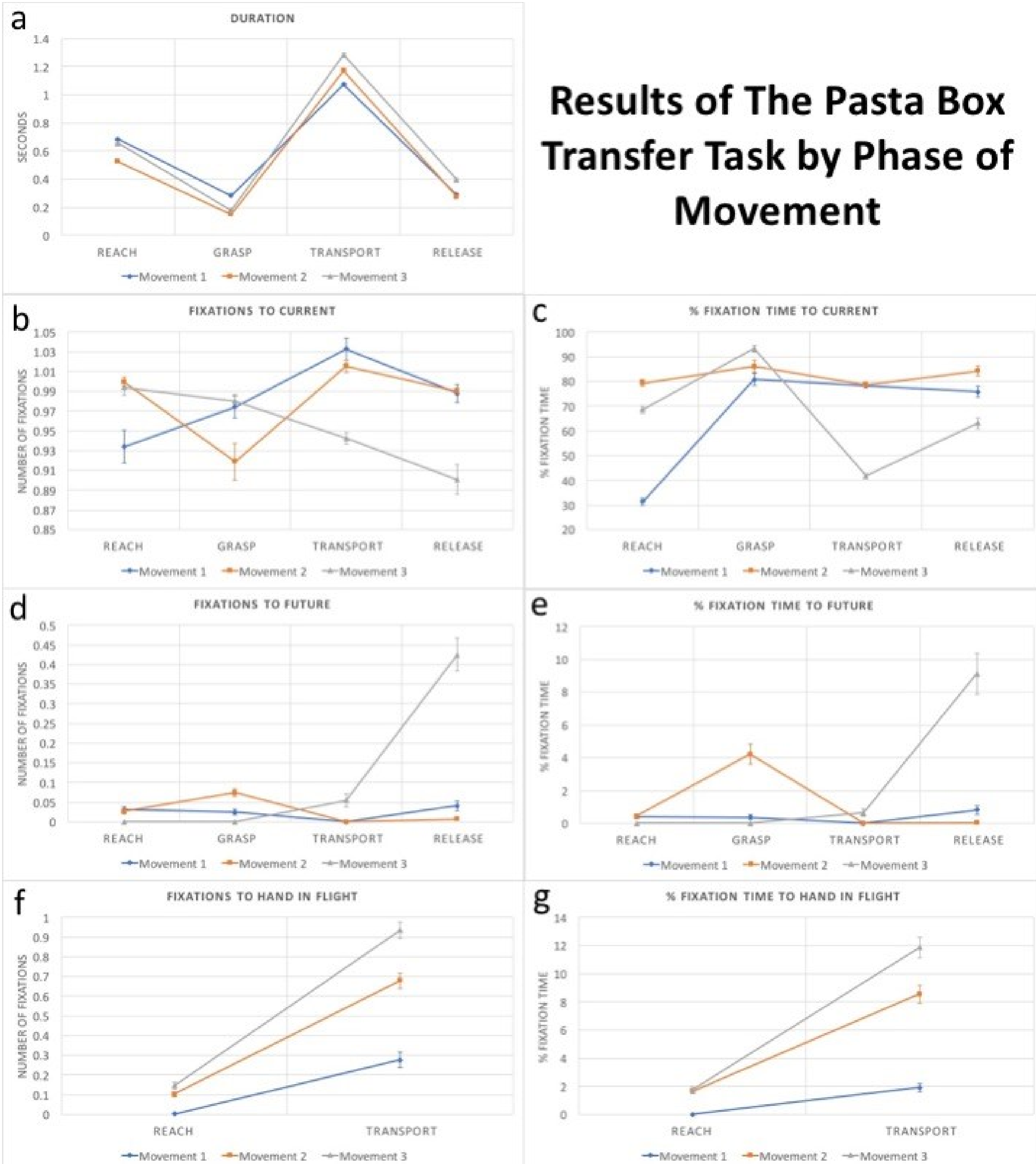

Figure 13: Data from the PB task comparing each Phase of Movement by a) Duration (s), b) Number of Fixations to Current, c) % Fixation Time to Current, d) Number of Fixations to Future, e) % Fixation Time to Future, f) Number of Fixations to Hand in Flight, and g) % Fixation Time to Hand in Flight. The error bars shown here were calculated by first obtaining each participant’s standard error, then calculating a mean value across all participants.

3.3.2 – Cup Transfer Task

3.3.2.1 – Description of eye gaze behaviour

Before delving into the in-depth statistical analysis and results (as shown in Tables 5 and 8, and Figures 14 and 15), here I provide a description of the possible drivers of eye movement behaviour for each ORA during the C task.

<u>ORA 1:</u> Participants start with their hand on the Home area and their eyes on the Neutral eye position. At the sound of the go cue, there is a slight delay as participants plan their movements. This delay transfers into the first Reach and even Grasp movement to the Green cup on the Near Target 1 which are longer than in the other ORAs. Additionally, since the eyes are forced to start on Neutral in Reach 1, which is a longer distance from their hand than in later Reaches, this results in less time fixating on the Green cup. When Reaching towards the Green cup in preparation of moving it, the participant's eyes will fixate on it well before the hand arrives at it (about half a second). Next, the participant Transports the Green cup over the partition efficiently with a top grasp and Releases it on the Near Target 2. The Near Target 2 showed difficulty in registering fixations because of its low and lateral location to the scene camera, resulting in fewer and shorter fixations. Interestingly, during the Release of the Green cup onto Near Target 2, the participant fixates a fair amount to the Future target of action, the Blue cup on Far Target 1. This can occur because these targets are quite close together, the closest of any two sequential targets in this entire experiment, enabling the fixation to land before the Releasing of the Green cup is complete. As well, participants need enough time to plan the upcoming Grasp of the Blue cup, and so, cut short their fixation on the Release of the Green cup to fixate on the Blue cup.

ORA 2: The second Reach phase, from the Near Target 2 to the Blue cup at Far Target 1 was short, because the participant is already moving and the distance the hand must travel is small (again, this was the shortest distance between any two consecutive targets in the entire experiment). The proximity of these targets also leads to a high % Fixation Time to the Blue cup during this Reach as the eyes have a very short distance to shift. The % Fixation Time to Current is also high here because participants were likely fixated on the Blue cup before the Reach phase even began. When Reaching towards the Blue cup in preparation of moving it, the participant's eyes fixate on it before the hand arrives at it (about half a second). The Grasping of the Blue cup takes slightly longer, either because the cup is further away from the participant's body, or because they were using a side grasp. The Transporting of the Blue cup over the partition to Far Target 2 takes slightly longer than the first Transport likely for the same reasons of distal movement and grasp type. Unlike the first Transport, the participant spends no time fixating to the future location of the hand (Home area) during this second Transport because the Home area is far from the Far Target 2 and the body is stretched over in a way that makes it impossible to fixate there. While Releasing the Blue cup on the Far Target 2, participants fixate for a long period on the drop-off location, and, like the Transport phase, they spend no time fixating on the Home area, as they are biomechanically unable to turn their head to the right while their right arm is stretched across their body.

After Releasing the Blue cup, the participant retracts their hand and touches the Home area, sometimes without even shifting their gaze from the Blue cup, as the next object to be picked up is again the Blue cup. Technically, this return Home doesn't count as an ORA, but has consequences for how the third ORA develops.

ORA 3: The Reach from the Home area to the Blue cup is the longest in duration, due to it being the longest distance that the hand travels in this task. Because of this, the eyes arrive at the Blue cup almost three quarters of a second before the hand does. The eyes and hand leave from the same spot (the Home area), and in some instances, the participant did not take their eyes off of the Blue cup from the previous Release. And so, the eyes land on the cup at the same time or sooner than in other Reaches, but the hand travels a much further distance, prolonging the time it takes the hand to arrive at the cup. The participant Grasps the Blue cup again, in about the same amount of time as the first Grasp of the Blue cup (during ORA 2), which is not unexpected as these Grasps are comparable distances from the body with the same grasp type. The Transporting of the Blue cup back to its original Far Target 1 location takes about the same amount of time as it took the first time it was Transported. However, there is slightly less fixation time at this drop-off location, and slightly more fixation to the participant's own hand. Several reasons contribute to this difference. First, biomechanically, the participant cannot as easily fixate to the target on the right when their right arm is stretched across their body to the left, resulting in a slight lag of fixation on the Blue cup at the start of the Transport phase. As well, as the eyes shift left to right from Far Target 2 to Far Target 1 and since the hand is on the right side of the cup, the fixation point falls on the hand for slightly longer than during the opposite movement direction (Transport of ORA 2 where the eyes are moving right to left). During the Release of the Blue cup back onto its original location, the Far Target 1, there is a reduction of fixation at the drop-off location. This is because the participant is already shifting their fixation to the very proximal next location of action, the Green cup at Near Target 2.

ORA 4: The final cup to be moved is the Green cup back to its original starting position. The participant Reaches towards the Green cup from its previous location, the Blue cup at Far

Target 1, in the shortest amount of time of any Reach. This is due to a combination of the short distance over which the hand moves, and because the hand is moving towards the body. Also, because the previous target is so close to the Green cup, the participant already has their eyes fixated on the Green cup during the Release of the Blue cup, leading to a fixation to the Green cup during most of the Reach. And, when Reaching towards the Green cup in preparation of moving it, the participant's eyes fixate on it well before the hand arrives at it (about half a second). There are few fixations made to the next target, the Home area, during this Reach. The Grasp of Movement 4 is also shorter than the previous ones, due to the use of a top grasp, and because the hand has moved towards the body. Because of the location of this Near 2 Target, the scene and pupil cameras have difficulty registering fixations to it. The participant then Transports the Green cup over the partition to the Near Target 1, which takes about the same time as the first time the Green cup was Transported. However, due to the drop-off location and the angle of the scene camera, a decrease in the time fixated on the drop-off location is registered along with an increase in the time fixated to the participant's own hand. This is also partly due to the hand being positioned on top of the object while doing a top grasp. When Releasing the Green cup on Near Target 1, the participant spends most of his time fixated on the drop-off location, and a bit of time fixated to the next targets (the Home area and/or Neutral eye position).

3.3.2.2 – Detailed statistical results

Similar to the PB task but differing in the number of levels for Movement, for each dependent measure except those of Hand in Flight and the Eye-Hand Latency, a two-factor (Movement, Phase) repeated measures ANOVA was carried out with four levels coding for Movement (1, 2, 3, 4), and four levels coding for Phase (Reach, Grasp, Transport, Release). For each dependent measure for Hand in Flight, a two-factor repeated measures ANOVA was carried out with four levels coding for Movement (1, 2, 3, 4), and two levels coding for Phase (Reach, Transport). Wherever significant, an interaction of Movement and Phase was followed up with simple main effects single-factor ANOVAs of Movements for each Phase. Post-hoc tests were run for significant main effects and simple main effects which compared all possible pairwise comparisons of the relevant factor (Movement or Phase for main effect, and Movement for simple main effects). For Eye-Hand Latency, a one-factor (Movement) repeated measures ANOVA was carried out with four levels coding for Movement (1, 2, 3, 4), with a post-hoc test run due to a main effect of Movement. Many significant differences were found which can be viewed in Tables 5 (Phase Dependent Measures) and 8 (Eye-Hand Latency). Similar to the PB task, the C task saw several significant differences in participant behaviour from one phase of movement to the next. This section is structured similarly, with the Duration measure being discussed first, followed by a description of metrics within each AOI, and finishing with an analysis of the Eye-Hand Latency measure. Graphical representations of each measure can be found in Figures 14 and 15 (b).

As in the PB task, there are a number of reasons for the effects found in the C task, which will be referenced in each section of the detailed analysis. Here is a brief description of each of the possible causes of the effects, which are referenced in italics in parentheses below. Participants' eye behaviour was affected by the distance over which their body had to move to complete a specific action (*Movement distance*). As well, phases of movement at the beginning of the task were slightly slower than actions later in the task, due to the complexity of movement planning (*Complexity of movement planning*). Also, in agreement with the PB task, participants in the C task could fixate sooner on objects that were closer to the last object that was interacted with by the hand (*Proximity rule*). And, biomechanical impediments, like when the right arm is stretched out across the body, caused the eyes to linger at objects longer than if the arm was in a less cumbersome position (*Biomechanics*). The path of the movement of the eye was also a cause of certain behavioural effects, especially where the path of the hand and path of the eye were similar, or where the path of the eye passed close to areas of interest (*Eye movement path*). Additionally, the type of grasp used, the distance the grasp was from the participant's body, the direction of hand movement, and the aperture of grasp and release seemed to affect some of the values collected (*Grasp type; Distance from body; Direction of hand movement; Apertures*).

3.3.2.2.1 – Duration

The trend of Phases of Movement seen in the PB task are also seen in the C task, with Transport being the longest, followed by Reach, and Grasp and Release being the shortest. Of the Reach phases, the short distance to travel in Movements 2 and 4 cause these to be the smallest, and the long distance in Movement 3 causes it to be the largest (*Movement distance*). Although the Reach of Movement 1 is a short distance, we observed a longer duration than would be expected based on movement distance, similar to that of the reach of the first movement of the PB task (*Complexity of movement planning*). The Reach of Movement 2 takes longer than 4, a finding we attribute to participants moving their hand away, rather than toward their body (*Direction of hand movement*). Grasps during Movement 2 and 3 are longer than those of Movement 1 and 4, which could be due to some combination of at least three reasons: the grasp type was different (top grasp for 1 and 4, side grasp for 2 and 3); the distance away from the body was different (near for 1 and 4, far for 2 and 3); and the size of the grip aperture defined distances was different - see Table 1 (*Grasp type; Distance from body; Apertures*). Finally, we see the Grasp of Movement 1 being longer than that of Movement 4, despite having the same grasp type (*Complexity of movement planning*). For the same possible reasons as outlined for

Grasps, the Transport phases of Movements 2 and 3 are longer than those of Movements 1 and 4 (*Grasp type and/or Distance from body*).

3.3.2.2.2 – Current

One of the prominent results we find in the eye movement behaviour during this task is driven by the Near Target 2 location (e.g. end of Movement 1, start of Movement 4). At this location, we experienced some signal drop out as it was hard for the eye-tracker to consistently record the pupil when participants were looking down and to the left. In addition, at this particular location, the exact position of the eye with respect to the cup was compromised as the hand was often in the line of sight (*Grasp type*). For this reason, the Number of Fixations to Current during the Release phase of Movement 1 and the Grasp phase of Movement 4 are measured as significantly lower. In other results for the Current AOI, the Reach phase of Movement 1 has a lower % Fixation Time because the eyes begin the task fixated on the Neutral eye position, creating an unnatural lag of the eyes behind the hand. Similar to the PB task, this lag of the eyes causes the participant's fixation to land on the Current AOI later in the Reach phase, thus registering less % Fixation Time than in later Reach phases (*Complexity of movement planning*). The % Fixation Time of the Reach phases of Movement 2 and 4 are significantly larger than those of Movements 1 and 3, likely because these targets are in close proximity to the previous drop-off location, and so the eyes were likely fixated on these targets during the Release phases of Movements 1 and 3, respectively (*Proximity rule*). The % Fixation Time to Current of the Reach phase of Movement 3 is low possibly due to the previous location of the eyes (Home area) being far from this target, increasing the time for the eyes to fixate on the target relative to the hand leaving the Home area (*Proximity rule*). The % Fixation Time to Current of the Transport phases of Movements 1 and 2 are larger due to the object being moved from right to left (*Direction of hand movement*). In Movements 3 and 4, participants transport the cups from left to right, starting with their right arm across their body, making it difficult to turn their head to fixate on the drop-off location on the right (*Biomechanics*). The % Fixation Time to Current of the Transport phase of Movement 3 is greater than that of Movement 4 likely because Movement 4 requires the participant to fixate low in their field of view and to the right, often causing the scene camera to miss the drop-off location. During the Release phases of Movements 1 and 3, participants were often ending their fixation to the Current AOI earlier than in Movements 2 and 4, to ensure their fixation to the Future AOI began with enough time to guarantee an efficient movement could be planned to this next target of action. This is supported by, not only the decreased % Fixation Time to the Current in Movements 1 and 3, but also the increased % Fixation Time to Future during these Movements. So, the eyes are more likely to shift their gaze away from the current drop-off location if the next target of action is near it (*Proximity rule*). This is evidence for the idea that the there is a minimum amount of time that the brain needs to compute visual information about the next target of action before the hand arrives at the target. This will be discussed more in the section discussing Eye-Hand Latency, and follows from the idea that there is an optimal amount of fixation prior to object grasping that the brain tries to achieve. See Figure 14 for graphical representations of the Number of Fixations (b) and % Fixation Time (c) to Current.

3.3.2.2.3 – Future

The Number of Fixations to Future during Grasp phases trends toward significance because the Grasp phase of Movement 1 shows higher values than the other Movements. This is stated because the omnibus ANOVA yields a significant difference, but no pairwise test does. This could be attributed to the ease with which the head and body can turn to the left when the right hand is grasping close to the body on the right (*Biomechanics*). The Number of Fixations during the Transport of Movement 2 is lower than that of Movement 1 because the distance to the future location is much further (*Proximity rule*). In the Transport of Movement 1, the Future AOI (Far Target 1) is a small saccade of the eyes away from the Current AOI (Near Target 2), while in the Transport of Movement 2, it is a physical impossibility that a person could fixate on the Home area (the Future AOI), while efficiently carrying out the task (*Biomechanics, Proximity rule*). The Number of Fixations and % Fixation Time for the Release phases of Movements 1 and 3 are the same because the future locations are close in proximity to the drop-off locations and a fixation to these locations is a necessity to ensure optimal task efficiency; as well, the Release phase of Movement 4 is close in proximity to the drop-off location, but due to Home area being low in the field of view of the first-person camera, this location was often missed (*Proximity rule*). The Release phase of Movement 2 has infinitesimally small values of Number of Fixations and % Fixation Time to Future AOIs because the Home area (Future) is quite far from the drop-off location, low in the field of view of the scene camera, and requires the participant to have their right arm outstretched to the left of their body (*Biomechanics, Proximity rule*). As discussed in the previous section, the high values of Numbers of Fixations and % Fixation Time to the Future AOIs during the Release phases of Movements 1 and 3 correlate with low % Fixation Time to Current AOIs during these phases. This shows evidence that participants stopped fixating to the drop-off location earlier in these Release phases than in those of Movements 2 and 4. As discussed above, we hypothesize this is because when an upcoming object interaction is going to occur sooner, the eyes will leave their current location, sacrificing a small amount of error in the Release for a more reliable next Reach and Grasp (*Proximity rule*).

See graphical representations of the Number of Fixations (d) and % Fixation Time (e) to Future in Figure 14.

3.3.2.2.4 – Hand in Flight

From looking at Figure 14 (f) and (g), it is obvious that participants fixated more on their hand when an object was in it (e.g. Hand in Flight fixations in Transport > Reach). But, within the Transport phases, it appears moving the cups from left to right using a side grasp (Movement 3) caused participants to fixate slightly more to their hand than when moving objects from right to left (Movements 1 and 2), as the Transport phase of Movement 3 had a significantly larger Number of Fixations and % Fixation Time to the Hand in Flight than those of Movements 1 and 2 (*Direction of hand movement*). As mentioned, a side grasp makes the grasping hand a larger target, which we believe contributes to these elevated values (*Grasp type*). Additionally, there were more Fixations to the Hand in Flight during Movement 3 than Movement 4, which could be attributed to the scene camera missing the Movement 4 looks to the hand as they were lower in the field of view, and more susceptible to data loss.

3.3.2.2.5 – Eye-Hand Latency

Again, the Eye-Hand Latency values were calculated for each ORA of the C task using the difference of the time of the first fixation to the Current AOI during the Reach phase of the ORA and the time of the start of the Grasp phase of this ORA. As is seen in Figure 15 (b), the Eye-Hand Latency of Movement 3 (0.714s) is significantly longer than those of Movements 1, 2, and 4 (0.462s, 0.508s, and 0.454s, respectively). The difference causing this result is the longer distance over which the hand moves during Movement 3 compared to the other Movements (*Movement distance*). This longer distance increases the duration of the Reach phase and thus allows the eyes to fixate longer before the hand arrives at the object. As well, some participants chose to keep their visual fixation on the Blue cup after its release at the end of Movement 2 as the hand moved and touched the Home area, eliminating any time the eyes would have to saccade from the Home area to the Blue cup at the start of Movement 3. The Eye-Hand Latency being much longer in Movement 3 is evidence for the proposal that the eyes will fixate on wherever the next target of action of the hand is as soon as they are able to dedicate their attention to it. As has been discussed earlier in the Current and Future sections of the C task analysis, the proximity of the Current and Future objects of action in portions of the C task sheds some light on the amount of time the brain needs to compute information about an object before the hand arrives at it. Specifically, it is interesting to consider what happens when participants are transitioning between the Near Target 2 and Far Target 1 locations (the two closest targets to one another) at the end of Movement 1 and start of Movement 2, and at the end of Movement 3 and start of Movement 4. The short distance in these transitions, coupled with the Eye-Hand Latencies of Movements 2 and 4, add more evidence to the idea that the brain requires visual information about an object approximately half a second in advance to the hand arriving at the object to interact with it (*Proximity rule*). "Presumably, this means that the motor system has half-a-second's worth of information available to it from a visual buffer" (Land & Hayhoe, 2001). Finally, a less notable difference is that Eye-Hand Latency of Movement 4 is slightly less than that of Movement 2. This could be due to the Near Target 2 being difficult for the scene camera to record, or that the grasp distance used to segment the data was slightly larger during Grasp 4 than for Grasp 2 (*Grasp type*).

Table 5: Cup Transfer Task Dependent Measure Means

Phase	Movement	Duration (s) Mean	Duration (s) $F_{(9,171)}$	Duration (s) Pair	Number of Fixations to Current (#) Mean	Number of Fixations to Current (#) $F_{(9,171)}$	Number of Fixations to Current (#) Pair	% Fixation Time to Current (%) Mean	% Fixation Time to Current (%) $F_{(9,171)}$	% Fixation Time to Current (%) Pair
Reach	1	0.678	**	1<<3; 2<<3; 4<<3; 2<<1; 4<<1; 4<2	0.995	ns		66.212	**	1<<2; 1<<4; 1<3; 3<<2
	2	0.525			1.003			96.147		
	3	0.876			1.025			80.801		
	4	0.488			1.003			89.934		
Grasp	1	0.178	**	1<<2; 1<<3; 4<<2; 4<<3; 4<<1	0.957	*		85.567	*	4<2
	2	0.221			0.980			91.148		
	3	0.222			0.980			93.369		
	4	0.145			0.875			76.502		
Transport	1	1.017	**	1<<2; 1<<3; 4<<2; 4<<3	1.008	ns		81.220	**	3<2; 3<1; 4<<2; 4<<1; 4<3
	2	1.162			1.013			82.408		
	3	1.164			1.015			72.240		
	4	1.038			1.000			66.903		
Release	1	0.303	ns		0.945	*	1<2; 1<4	62.980	**	3<<2; 3<<4; 1<<2; 1<<4; 1<3
	2	0.314			0.995			88.887		
	3	0.302			0.970			73.218		
	4	0.325			0.985			86.068		

Phase	Movement	Number of Fixations to Future (#) Mean	Number of Fixations to Future (#) $F_{(9,171)}$	Number of Fixations to Future (#) Pair	% Fixation Time to Future (%) Mean	% Fixation Time to Future (%) $F_{(9,171)}$	% Fixation Time to Future (%) Pair	Number of Fixations to Hand in Flight (#) Mean	Number of Fixations to Hand in Flight (#) $F_{(3,57)}$	Number of Fixations to Hand in Flight (#) Pair	% Fixation Time to Hand in Flight (%) Mean	% Fixation Time to Hand in Flight (%) $F_{(3,57)}$	% Fixation Time to Hand in Flight (%) Pair
Reach	1	0.013	*		0.030	ns		0.005	ns		0.037	ns	
	2	0.005			0.174			0.074			0.789		
	3	0.033			0.340			0.071			0.597		
	4	0.003			0.026			0.089			0.712		
Grasp	1	0.133	ns		4.473	ns							
	2	0.068			3.015								
	3	0.055			2.728								
	4	0.013			0.637								
Transport	1	0.043	*	2<1	0.283	ns		0.510	**	4<3; 2<3; 1<<3	5.352	**	2<3; 1<<3
	2	0.000			0.000			0.571			6.725		
	3	0.018			0.095			0.808			10.759		
	4	0.008			0.156			0.620			9.240		
Release	1	0.746	**	4<<1; 4<<3; 2<<1; 2<<3; 2<<4	27.806	**	4<<1; 4<<3; 2<<1; 2<<3; 2<4						
	2	0.000			0.000								
	3	0.632			19.963								
	4	0.190			5.084								

Table 5: The means of the dependent measures calculated for each Phase of the C task, with statistical effects from a Movement X Phase interaction indicated by: ns = not significant, * = $p < 0.05$, ** = $p < 0.005$. As well, any significant pairwise contrasts of Movements within a Phase, and their direction are indicated by: < = $p < 0.05$, << = $p < 0.005$.

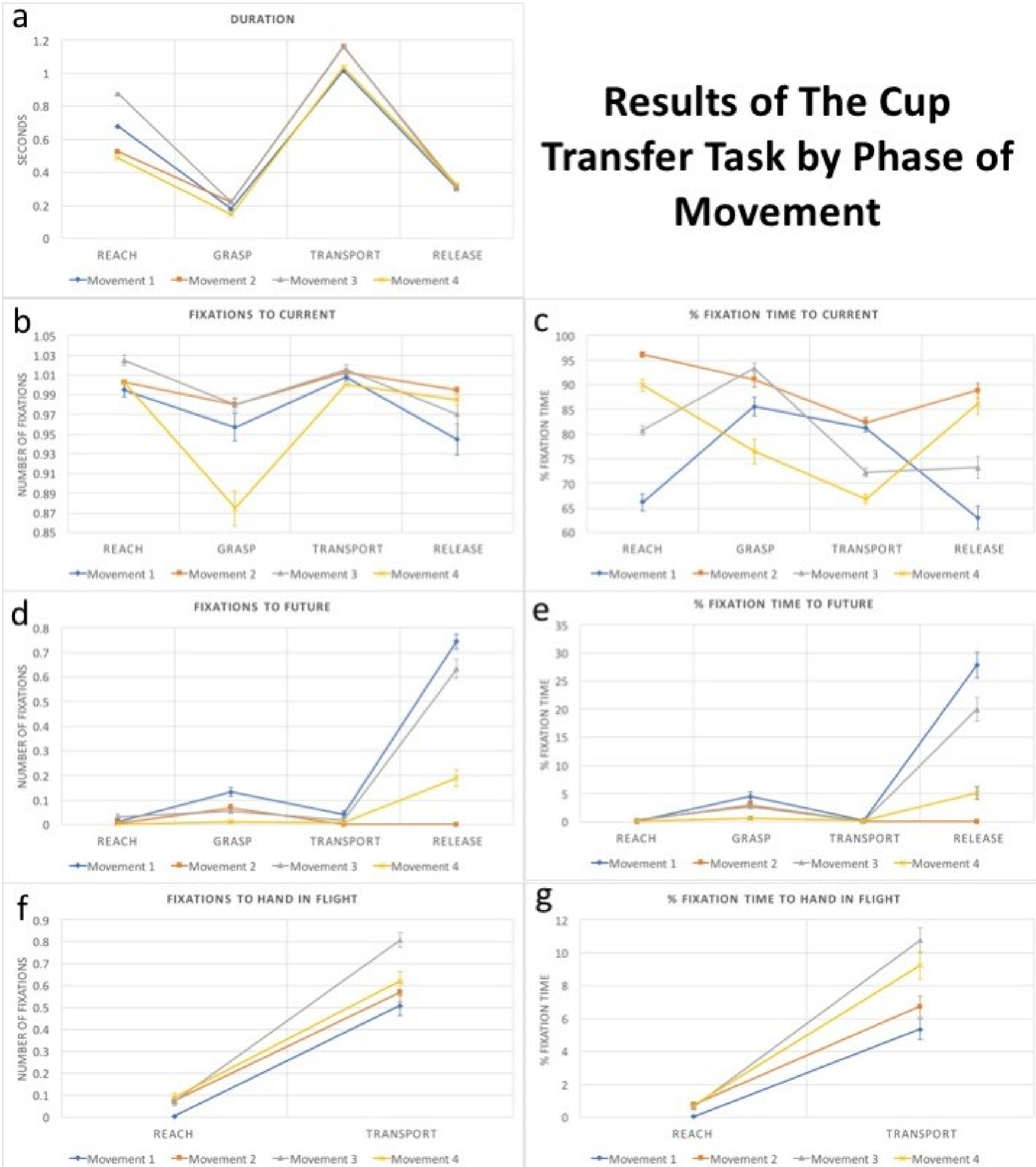

Figure 14: Data from the C task comparing each Phase of Movement by a) Duration (s), b) Number of Fixations to Current, c) % Fixation Time to Current, d) Number of Fixations to Future, e) % Fixation Time to Future, f) Number of Fixations to Hand in Flight, and g) % Fixation Time to Hand in Flight. The error bars shown here were calculated by first obtaining each participant's standard error, then calculating a mean value across all participants.

3.3.3 – Pasta Box versus Cup Transfer Tasks

In the final analysis, the PB task was compared to the C task for each dependent measure divisible by phase. Because the PB task has 3 Movements with 4 Phases in each, and the C task has 4 Movements with 4 Phases in each, a direct comparison of the Phases within each Movement was not possible, and, even if it were, the dimensions and tasks differed significantly, so analyzing by Movement is not meaningful for these measures in this task comparison. Instead, in each task, an average of each of the 4 Phases was obtained. Then, two-factor (Task, Phase) repeated measures ANOVA tests with Greenhouse-Geisser sphericity corrections were run for each dependent measure (excluding Eye-Hand Latency) with significant results followed up with pairwise comparisons with Bonferroni corrections. For Eye-Hand Latency, an average of the values of each Movement within each Task was obtained, with these values being compared in a pairwise comparison with a Bonferroni correction. Given the detailed results reported above from within each task, only those results involving either a main effect or an interaction effect involving Task are reported here. The means of the dependent measures and results of the statistical tests outlined above can be found in Table 7 for the Phase dependent measures and Table 8 for the Eye-Hand Latency.

3.3.3.1 – Duration

The C task Transport phases are significantly shorter in duration than those of the PB task. This can likely be attributed to the Transport phase of Movement 3 in the PB task (from the High Shelf Target, all the way down to the low Start/End Target) being a much further distance than any of the other Transport phases in either task, leading to the longest physical movement in the study (*Movement distance*). This would also explain the fact that no difference in the Reach phases of the two tasks was observed, as there is no comparably long Reach distance to that of the Transport phase of Movement 3 in the PB task.

3.3.3.2 – Current

There were no significant differences between the two tasks in the Number of Fixations to Current, with each Phase falling near 1.0 for both tasks. However, the PB task has a significantly lower % Fixation Time for both the Reach and Transport phases than those of the C task. This difference can be attributed to the C task having all potential fixation targets directly in front of the participant and in relatively close proximity to one another, while the PB task has the Start/End Target location out of the field of view of the participant, and all targets further away from one another (*Proximity rule*). In general, the out-of-sight location (Reach Movement 1 and Transport Movement 3 in PB) led to shorter % Fixation Times which can account for the specific differences between Reach and Transport seen here between the two tasks. Additionally, it is possible that the greater % Fixation Time to Current in the Reach and Transport phases of the C task could be due to the increased risk of failure when interacting with cups filled with beads as opposed to a larger, easier to grasp Pasta box.

3.3.3.3 – Future

The PB task also has fewer Numbers of Fixations, as well as less % Fixation Time to the Future targets in the Release phases than the C task. This can also be attributed to the C task having all potential fixation targets directly in front of the participant and in close proximity to one another (*Proximity rule*). More specifically, it is likely that the Release phases of Movements 1 and 3 in the C task drive the entire C task to be higher than the PB task, in turn because the drop-off location targets and the next objects to be picked up in these phases are the closest in distance of any combination in either task. Due to the close proximity of these targets, the participant shortens their % Fixation Time to the Current location of Release to the Future location of action to ensure they have enough time to analyze the visual information about the Future location before the hand arrives (*Proximity rule*). This is less of a concern in the PB task, as no objects are interacted with consecutively, instead after every Release the hand returns to the Home area before interacting with the next object.

3.3.3.4 – Hand in Flight

There are no significant differences in Number of Fixations or % Fixation Time to the participant's own Hand in Flight between the two tasks. In both tasks, there are increases in these values during the Transport phases. This is an indication that these differences are not due to differences in the two tasks, like height of the object movement, angle of the scene camera, type of grasps, or direction of movement, but rather due to the specific difference between Reaches and Transports, of which there is only one. The fact that participants fixate slightly more to their own hand when moving an object than when not moving an object is an important finding. Although we may feel as though we gain enough information about the stability of the object in our hand from other sensory modalities like proprioception, it is obvious that we still dedicate more visual attention when an object is in our hand. However, it is important to note that, although we do dedicate more visual attention to our hand when transporting an object, the actual increase in values is still quite small (% Fixation Time to Hand in Flight ≈ 8%), which supports the notion that we may just fixate slightly longer at a specific portion of the Transport phase (likely the initial pick up portion after the Grasp phase ends).

3.3.3.5 – Eye-Hand Latency

The significantly shorter Eye-Hand Latency in the PB task (0.367s) compared to the C task (0.534s) could be due to a few factors. First, the C task involves interacting with objects with a higher risk of failure (cups filled with beads) and with specific grasp demands (top and side grasp versus just side grasp). Perhaps participants were forced to fixate slightly longer to the riskier-to-move cups, with their painted grasp targets, than to the Pasta box to allow their brain time to compute more complex calculations and ensure a successful interaction. Next, the C task had one Eye-Hand Latency that was far greater than the other values (Eye-Hand Latency of Movement 3 = 0.714s) which is responsible for pulling the average Eye-Hand Latency of the Cups Task up to 0.534s. As mentioned above, at least some of this particularly elevated value is likely driven by participants not moving their eyes on the one return to home movement in the C task. As well, the PB task had one Eye-Hand Latency that was far smaller than the other values (Eye-Hand Latency of Movement 1 = 0.216s) which is responsible for pulling the average Eye-Hand Latency of the PB task down to 0.367s. Again, this lower value is due to a confluence of factors, all centred around needing to turn the body to an out of sight location. If these two outlier values are removed from each task, the average values are much closer to one another (PB task = 0.444s; C task = 0.475s). Coupled with this, the fact that objects being sequentially moved without an intermediary action (touching the Home area) result in shortening of fixations to current objects in favour of beginning fixations to future objects (C task: Movements 1 $\rightarrow$ 2, Movements 3 $\rightarrow$ 4) supports a minimum Eye-Hand Latency required to successfully interact with an object. It seems as though there is strong evidence that this value falls somewhere around 0.5s (Land & Hayhoe, 2001).

Table 6: Dependent Measure Means of Eyes and Both Conditions

Task	Condition	Duration (s) Mean	Duration (s) $F_{(1,15)}$	Duration (s) Pair	Number of Fixations to Current (#) Mean	Number of Fixations to Current (#) $F_{(1,15)}$	Number of Fixations to Current (#) Pair	% Fixation Time to Current (%) Mean	% Fixation Time to Current (%) $F_{(1,15)}$	% Fixation Time to Current (%) Pair	Number of Fixations to Future (#) Mean	Number of Fixations to Future (#) $F_{(1,15)}$	Number of Fixations to Future (#) Pair	% Fixation Time to Future (%) Mean	% Fixation Time to Future (%) $F_{(1,15)}$	% Fixation Time to Future (%) Pair	Number of Fixations to Hand in Flight (#) Mean	Number of Fixations to Hand in Flight (#) $F_{(1,15)}$	Number of Fixations to Hand in Flight (#) Pair	% Fixation Time to Hand in Flight (%) Mean	% Fixation Time to Hand in Flight (%) $F_{(1,15)}$	% Fixation Time to Hand in Flight (%) Pair
Pasta	Eyes	0.541	*	E < B	0.987	ns		71.688	ns		0.086	*	B < E	1.659	*	B < E	0.131	ns		1.021	ns	
	Both	0.565			0.961			70.073			0.069			1.313			0.128			1.039		
Cups	Eyes	0.538	**	E < B	1.035	ns		76.828	ns		0.161	ns		3.441	ns		0.251	ns		2.308	ns	
	Both	0.562			1.046			78.967			0.170			3.451			0.232			2.001		

Table 7: Dependent Measure Means of Pasta and Cups Tasks

Phase	Condition	Duration (s) Mean	Duration (s) $F_{(3,57)}$	Duration (s) Pair	Number of Fixations to Current (#) Mean	Number of Fixations to Current (#) $F_{(3,57)}$	Number of Fixations to Current (#) Pair	% Fixation Time to Current (%) Mean	% Fixation Time to Current (%) $F_{(3,57)}$	% Fixation Time to Current (%) Pair	Number of Fixations to Future (#) Mean	Number of Fixations to Future (#) $F_{(3,57)}$	Number of Fixations to Future (#) Pair	% Fixation Time to Future (%) Mean	% Fixation Time to Future (%) $F_{(3,57)}$	% Fixation Time to Future (%) Pair	Number of Fixations to Hand in Flight (#) Mean	Number of Fixations to Hand in Flight (#) $F_{(1,19)}$	Number of Fixations to Hand in Flight (#) Pair	% Fixation Time to Hand in Flight (%) Mean	% Fixation Time to Hand in Flight (%) $F_{(1,19)}$	% Fixation Time to Hand in Flight (%) Pair
Reach	Pasta	0.622	ns		0.976	ns		59.771	**	P << C	0.020	ns		0.296	ns		0.084	ns		1.140	ns	
	Cups	0.642			1.006			83.273			0.013			0.142			0.060			0.534		
Grasp	Pasta	0.205	ns		0.958	ns		86.685	ns		0.033	ns		1.539	ns							
	Cups	0.191			0.948			86.646			0.067			2.713								
Transport	Pasta	1.177	**	C << P	0.997	ns		66.269	**	P << C	0.018	ns		0.220	ns		0.630	ns		7.440	ns	
	Cups	1.095			1.009			75.693			0.017			0.134			0.627			8.019		
Release	Pasta	0.321	ns		0.960	ns		74.365	ns		0.158	**	P << C	3.323	**	P << C						
	Cups	0.311			0.974			77.788			0.392			13.213								

Table 6: The means of the dependent measures calculated for each Task under both Conditions, with statistical main or interaction effects of Condition being indicated by: ns = not significant, * = $p < 0.05$, ** = $p < 0.005$. As well, any significant pairwise contrast of Condition within a Task, and its direction is indicated by: < = $p < 0.05$, << = $p < 0.005$.

Table 7: The means of the dependent measures calculated for each Phase for both Tasks, with statistical main or interaction effects of Phase being indicated by: ns = not significant, * = $p < 0.05$, ** = $p < 0.005$. As well, any significant pairwise contrast of Task within a Phase, and its direction is indicated by: < = $p < 0.05$, << = $p < 0.005$.

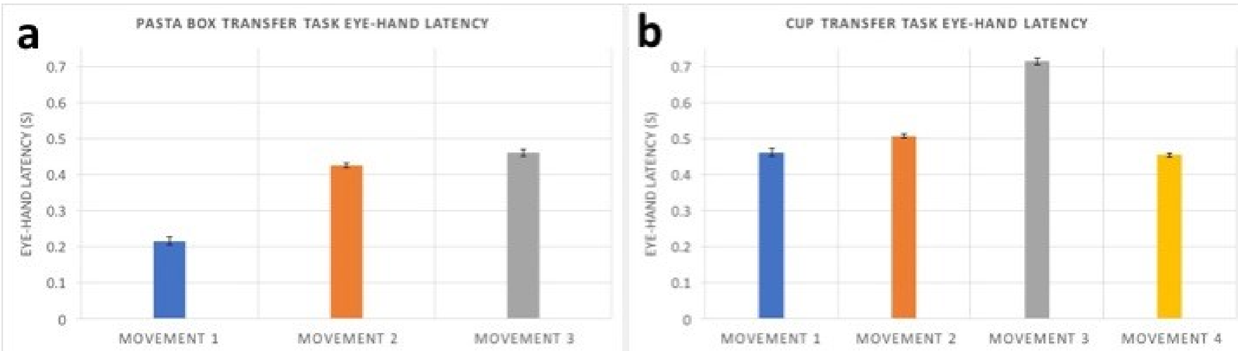

Figure 15: Eye-Hand Latency (s) values of: a) the PB task, and b) the C task, calculated for each ORA by using the difference in total time between the eyes fixating on the object to be picked up in the Reach or Grasp phase, and the beginning of the Grasp phase (the hand arriving at the object). The error bars shown here were calculated by first obtaining each participant's standard error, then calculating a mean value across all participants.

Eye-Hand Latency (s)					
Pasta			Cups		
Movement	Mean	Pair	Movement	Mean	Pair
1	0.216	1 << 2; 1 << 3	1	0.462	1 << 3; 2 <<3; 4 << 3; 4 << 2
2	0.427		2	0.508	
3	0.460		3	0.714	
			4	0.454	
Pasta Average 0.367		Pasta << Cups		Cups Average 0.534	

Table 8: Eye-Hand Latency (s) values for each ORA from both the PB and C tasks, a total average for each task, and statistical comparisons within and between the tasks. Significance and direction of pairwise contrasts are shown with: ns = not significant, $<$ = $p < 0.05$, $<<$ = $p < 0.005$.

4 - Discussion

4.1 – Results Summary

In two experiments under two separate conditions, I examined the eye movement behaviour of participants as they moved everyday objects in simulated real-world environments. The tasks had participants make either 3 movements of a pasta box onto shelves, or 4 movements of cups over a partition. Both tasks outlined a specified order and grasp pattern to be employed during the experiment. I explored which objects and locations were fixated during specific phases of each movement, how much time each object or location was fixated relative to the length of each phase of movement, how often and for how long fixations occurred to a participant's own hand or an object in their hand while the hand was moving, and the amount of time the eyes fixated an object before the hand arrived at the object. To do this, tasks were segmented into Reach, Grasp, Transport, and Release phases based on the location and behaviour of the hand in relation to the objects being interacted with. As well, objects and drop-off locations analyzed within each phase were simplified to the following Areas of Interest: the Current location being acted upon by the hand (Current), the Future location that the hand will act upon next (Future), and the Hand in Flight. A visual summary of these results can be seen in Figure 16.

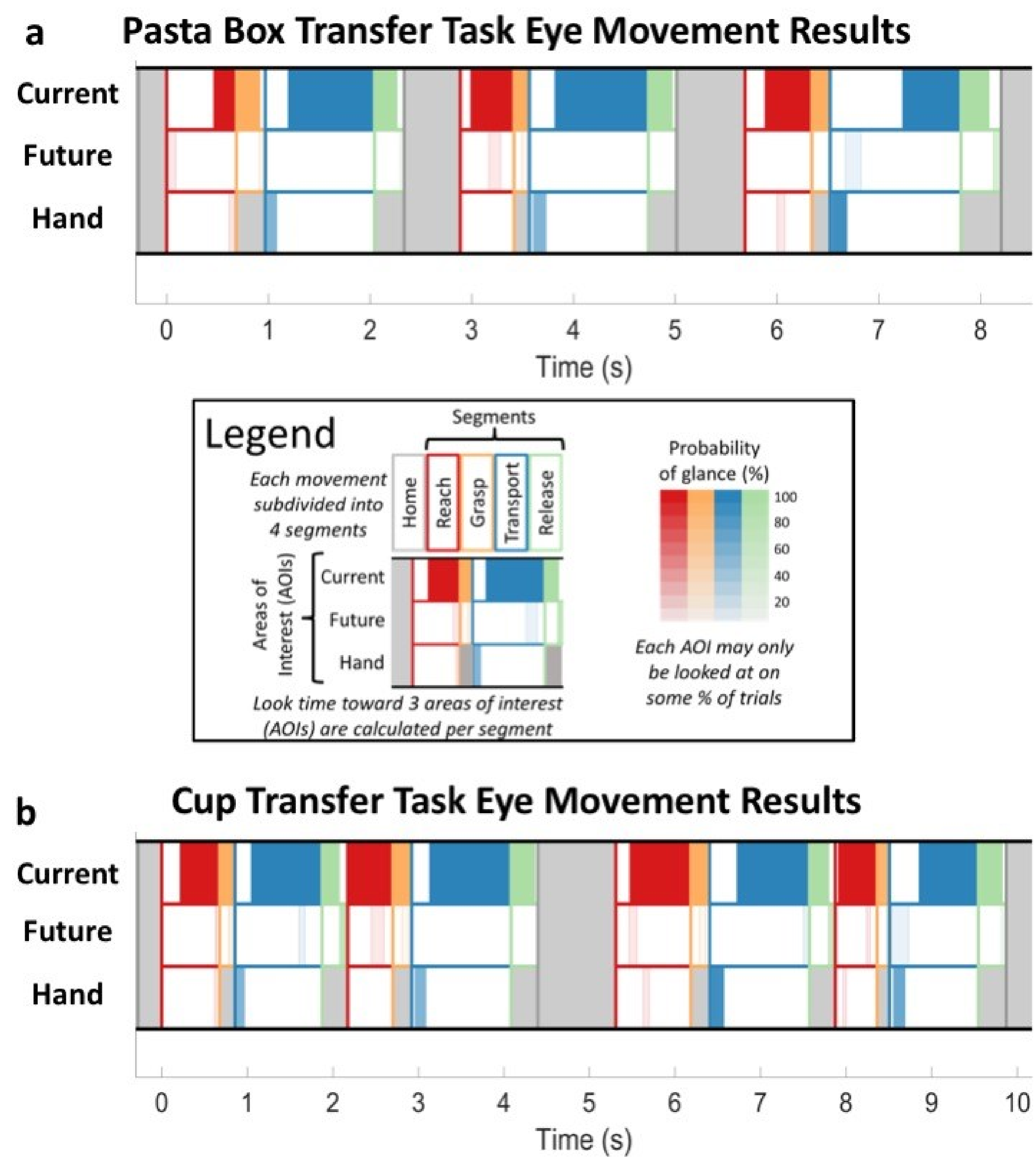

Figure 16: A timeline of the fixations to each of the Current, Future, and Hand in Flight AOIs during: a) the Pasta Box Transfer Task, and b) the Cup Transfer Task. The Legend explains that the more opaque a colour is on a timeline, the greater the probability that a fixation occurred on an individual trial.

First, based on previous work, I hypothesized that participants would spend little time fixating on objects irrelevant to the task, and the majority of the task time fixating on objects or locations relevant to the task (Buswell, 1935; Belardinelli et al., 2015; DeAngelus & Pelz, 2009;

Hayhoe, 2000; Land & Hayhoe, 2001; Land et al., 1999; Yarbus, 1967). Based on the findings of Land and Hayhoe, I expected that the eyes would arrive at objects being picked up well in advance of the hand, and, that the eyes would arrive at the drop-off location of an object being transported by the hand well in advance of the hand arriving with the object (2001). Additionally, I expected that objects or drop-off locations positioned outside of the field of view (Start/End Target location of the PB task) would elicit shorter fixation times than those positioned within the field of view, that objects with greater risk of failure (cups with beads) would elicit greater fixation time than objects with less risk (pasta box), and that different grasp types may require different durations of fixation than others. Lastly, I hypothesized that participants would make few fixations to, and spend little time fixating on their own hand.

Most of my predictions were shown to be true. When looking at the % Fixation Time to Current, Future, and Hand in Flight over the course of each of the two tasks, it can be seen that participants spend the majority of their time fixating on objects that are relevant to what their hand is currently interacting with, or what it will be interacting with in the near future. As well, time that is unaccounted for (approximately 25% does not fall under the Current, Future, Hand in Flight categories) can be explained as the time a participant was making a saccade, blinking, turning their head or making another movement that resulted in data loss. Not only did participants spend nearly all their detected fixation time during the task on action-relevant locations, but, of the objects that were fixated, these were nearly always made up of where the participant was currently acting upon, or would be acting upon in the immediate future. This 25% is much larger than the 5% value stated by Land and Hayhoe (2001) describing the amount of fixation time to objects irrelevant to the task. So, our 25% value includes any fixations that were made to irrelevant areas, along with the other reasons stated above. Moreover, the eyes lead the hand through both tasks. This can be seen in the high % Fixation Time to Current in the Reach and Transport phases of both tasks, and the approximately 0.5s delay between the eyes arriving at an object and the hand arriving at the same object to pick it up (Eye-Hand Latency). Interestingly, our prediction that at some point during the Grasp, the eyes saccade to the next target of action is found to be true, as in both tasks, the % Fixation Time to Future, while small, is on average non-zero. Similarly, we found that the % Fixation Time to Future during the Release phases in both tasks was non-zero, showing that participants also shifted their fixation to the next target of action during the Release phase. Interestingly, we see a substantial increase in % Fixation Time to Future during the Release phases of Movement 1 and 3 in the Cups Task. These two movements are the only occurrences in the whole experiment where a participant finishes interacting with one object and moves their hand directly to another object. This provides evidence for a minimum buffer time needed by the brain to analyze visual information in preparation for a manual interaction with an object. This buffer time is approximately 0.5s. As well, from the fairly consistent value of 1 for the Number of Fixations to Current across both tasks and phases, it is fair to say that once a participant fixated on an object during the end of the previous Release phase, they would stay fixated on that object until some point during the Grasp phase, at which point they would saccade to the drop-off location of that object, and stay fixated on that drop-off location through the Transport phase and part of the Release phase, at which point they would saccade to the next target of action. Figure 17 shows one ORA from the Cups task, illustrating many of the above points.

Next, as predicted, objects and target locations outside the field of view (the Start/End Target of the PB task) were shown to have lower % Fixation Time. Operating against one of our predictions, objects with a higher risk of failure when being transported (cups filled with beads) did not elicit greater Numbers of Fixations, nor greater % Fixation Time to the Hand in Flight. However, even though we did not see the expected variation due to risk, we did see, in both tasks, greater Numbers of Fixations and larger % Fixation Time to Hand in Flight during the Transport phases than in the Reach phases. This indicates that there is more attention being paid when the hand is moving with an object in it, than when the hand is moving by itself. This relates to our final hypothesis - relative to other objects and relevant locations in both tasks, participants spent very little time fixating on their own hand. And, even where we do see fixations to the Hand in Flight it could be incidental rather than functional, such as the hand and object passing into the line of sight as the participant is fixated on the drop-off location, or due to the hand and eye movement paths being similar enough that the eyes' saccade follows the hand for a few frames before separating.

Taken as a whole, these results provide a robust depiction of how humans use their eyes during object movement. Previous studies have looked at how humans coordinate reaching and grasping in far more limited lab tasks with artificial structure and many repetitions, and in tasks where participants can move around freely and carry out objectives in any order of their choosing, but with the added limitation of making analysis and decomposition of the results challenging. In the current experiments, we strike a balance between these two approaches by using the structure of repeatable movements while leveraging the freedom of new wireless eye and motion tracking systems. This enabled us to design functional movement tasks where participants could stand and move freely. Humans visually fixate on an object they are going to grasp in advance of their hand arriving at it. They stay fixated on the object long enough to ensure that their hand will be able to reliably grasp the object, at which point they shift their visual fixation directly to the location that they will place the object. As the hand transports the object, their eyes stay fixated on the location that the object will be placed, until, at some point during the release of the object, their eyes make a saccade to the next object to be picked up. A summary of the major results of the study can be found in Table 8. In subsequent sections I discuss the limitations of this study, propose a theoretical Attentional Landscape Model of Object Movement when objects are both inside and outside a person's field of view, and connect these ideas to Land and Hayhoe's Schema Theory of Object Interaction.

Table 9: Major Results Summary
• Participants spend the majority of task time fixating on objects/areas relevant to the task – specifically, where their hand is currently acting or where it will be acting in the immediate future • Participants make one fixation to an object they will pick-up, maintain this fixation until their hand nears the object to grasp it, at which point a saccade is made to the object's drop-off location where fixation is maintained until the object is being released from the hand • Participants' eyes lead their hand by approximately 0.5s when reaching for an object – this proposed buffer time is used to compute grasp specifications • Objects/areas outside participants' field of view were fixated for less time than those within their field of view • Participants spend little time fixating on their own hand – but, participants do spend more time fixating on their own hand when they are transporting an object than when they are reaching without an object in their hand

Table 9: A summary of the major results of eye movement behaviour during sequential object movement.

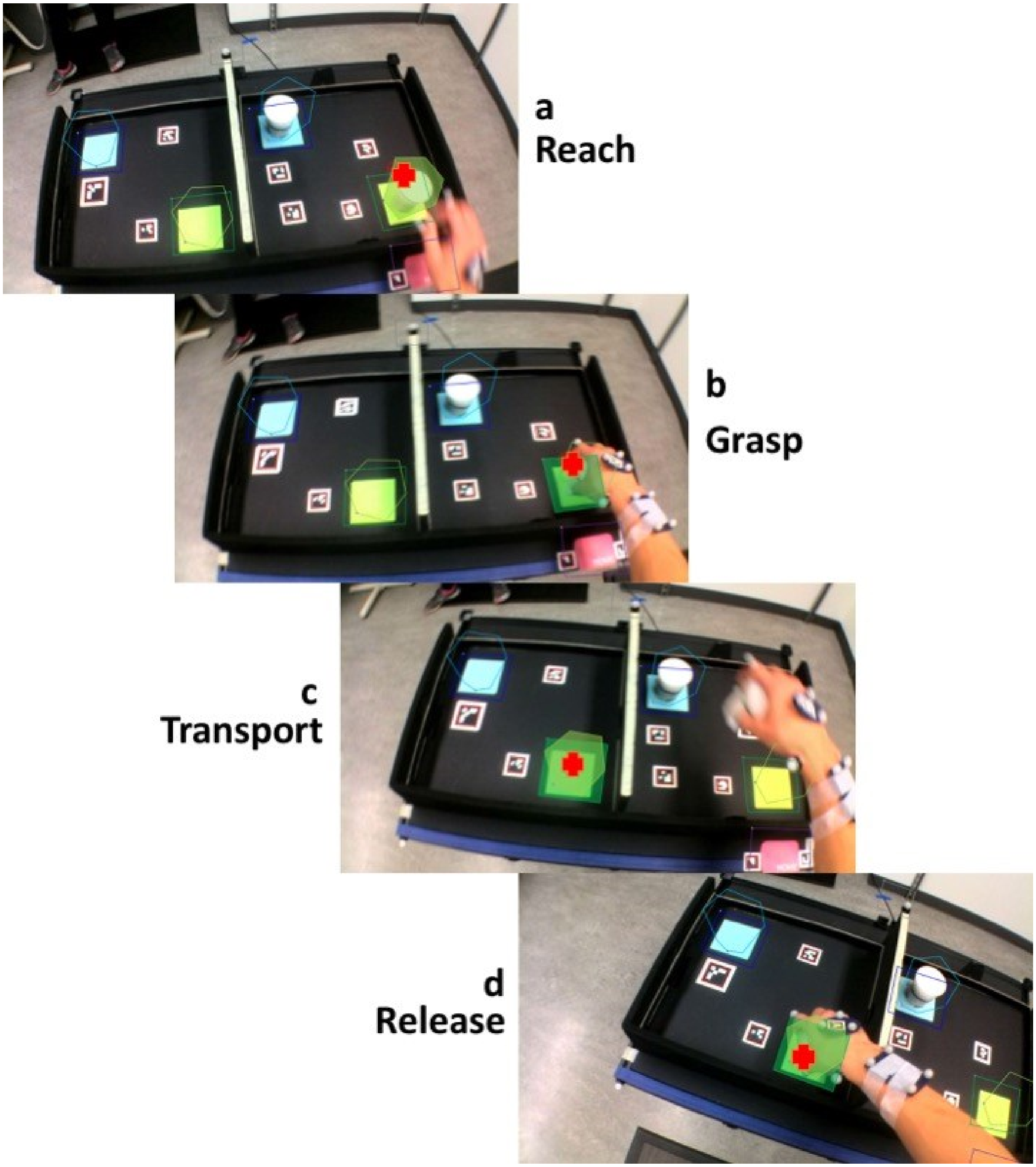

Figure 17: One ORA from the Cup Transfer Task showing the visual fixations (highlighted with a red cross) during a) Reach, b) Grasp, c) Transport, and d) Release phases.

4.2 – Limitations

Although certain limitations have been touched on already, here is a full description of all known shortcomings in this study.

4.2.1 – Eye-tracker Hardware and Software

The Dikablis 2.0 Professional head-mounted eye-tracker with D-Lab recording and analysis software was, overall, a good combination to use for our study. However, a few issues created limitations.

The headset itself allowed researchers to change the angle and placement of the pupil cameras, which, although beneficial for optimal pupil detection, took researchers varied amounts of time for each participant. And, due to differences in the height of the objects between the tasks, the pupil camera placement often had to be altered between the two. The scene camera could also be placed at a range of positions between 90° and 40° to the forehead. Again, this provided experimenters with choice, but also, due to the field of view provided by this camera, often lead to portions of the relevant environment being cut out of the frame. In addition, like most head-mounted eye-trackers, the Dikablis was uncomfortable to wear for long periods of time, and experimenters removed the device between Conditions to alleviate participants' discomfort. This added to the total amount of time each participant was involved in testing, as any removal and redonning, or alteration of any of the cameras required a recalibration of the pupil cameras to the scene camera. Moreover, the Dikablis 2.0 Professional eye-tracker was advertised as having a wireless option, where participants could wear a small backpack carrying a laptop to store recorded data, and a power source to provide energy to it. However, early pilot studies showed that in its wireless setup, the Dikablis eye-tracker and D-Lab software failed to yield data recordings with a consistent framerate of 60 Hz. So, for our experiment, a wired setup was utilized, which potentially inhibited participants' free movement.

Although the D-lab software is quite reliable at detecting the location of the pupil, many instances occurred where an eyelash or a dark portion of the iris was detected as the pupil. This required researchers to scan through each trial of every participant, frame by frame, to detect and correct misdetected pupils. Moreover, pupils were often lost at locations requiring participants' eyes to fixate at a sharp angle from its socket (eg. the Near Target 2 of the Cups Task). As well, the Calibration Adjustment tool described in the Methods section was unable to account for changes in depth. For example, a participant's fixation crosshairs, as detected by the D-Lab software, may be directly on the Neutral eye position at the beginning of the Pasta task, but when the participant turns to fixate the pasta box at the Start/End Target, the fixation crosshairs detected in D-Lab could be significantly shifted from their obvious actual fixation location, the pasta box. This was corrected as best as possible by experimenters, but is still reflected in the results, particularly at locations where depth was significantly different than the distance over which the pupil and scene cameras were calibrated at.

When participants turned their heads quickly, QR codes were blurred, often causing the AOIs for automatic fixation detection to disappear. This happened most consistently in the PB task when participants turned towards the Start/End Target, and back to the Mid Shelf Target. Again, experimenters scanned through each trial of every participant, after AOIs had been created, to correct instances where participants were fixated on relevant objects in the experiment, but blurred QR codes had led to missed object fixation time. The moving scene camera even made the manual corrective coding of fixations difficult, as researchers were still viewing a blurred, and therefore uncertain target. Additionally, target locations that were in less central locations in the environment, like the Near Targets and Home area in the Cups task, and the Start/End Target and Home area in the Pasta task, sometimes fell outside the field of view of the scene camera, restricting the logging of a fixation to these AOIs.

4.2.2 – Synchronization of Eye and Motion Tracking

Originally, researchers planned on using the full recordings collected from each of the eye and motion tracking software programs that had been signaled to start and end recording at the same time by a third piece of software. However, once the data had been collected, it was realized that the eye-tracking software regularly had more recording time than the motion tracking software. It was assumed that the recordings ended at the same time, and so, to bridge the synchronization of these two data streams, up to 0.400s was removed from the start of the eye-tracking data. On average, 0.124s of eye-tracking data was removed at the beginning of each trial to resolve this issue.

4.2.3 – Segmentation of Eye-tracking Data

The segmentation of the eye-tracking data based on the motion tracking data is a defining aspect of our study. The decision to refrain from using motion capture markers on relevant objects and participants' forefinger and thumb in the Eyes Only Condition therefore required a compromise. The purpose of the Eyes Only Condition was to quantify the eye behaviour changes, if any, that the motion capture markers caused. Except, removing most of these markers (participants still wore a headband with 4 markers, and 1 marker on the back of their hand in the Eyes Only Condition) from the environment inhibited our data analysis approach to the point that a completely different strategy had to be used. Using consistently-sized Regions of Interest around relevant locations to segment the data in the Eyes Only vs Both Analysis, although not as accurate as using grip aperture, object velocity, and distance, allowed us compare the two conditions successfully.

The Both Condition Analyses, including the individual task analyses and the comparison of the two tasks, also had some issues. Referring to Table 1, it is obvious that the distances used to define Grasp and Release phases (defined via peak grip aperture in each ORA) are consistent neither between nor within tasks. Therefore, differing distances could be the cause of variations in the Duration and % Fixation Time measures of different phases. There is solid ground to stand on in an argument for a comparison of the ROI segmentation strategy to the grip aperture, object velocity segmentation strategy, which could find a middle-ground between the two.

4.2.4 – Participant Differences

As mentioned in the Methods section, data was collected successfully from 20 participants. 16 of which were right-handed with normal or corrected to normal vision, 2 of which were right-handed without corrected vision, and 2 of which were left-handed with corrected to normal vision. The obvious standout issues here are the 2 participants who were not natural right-handers, and 2 participants who did not have normal to corrected vision. It is not hard to assume that using one's non-dominant hand could influence the eye movement behaviour of a person in a sequential object movement task. As well, although both participants assured the experimenters that they could perform the tasks normally and efficiently without their glasses, it is safe to assume that their eye movement behaviour could have been affected by this.

In addition, differences in the functioning of the D-Lab software based on participants' height were noted anecdotally by experimenters during the data post-processing stage. In the PB task, and to a lesser degree in the C task, D-Lab had greater difficulty accurately detecting fixations of shorter participants to Mid and High Shelf Targets. Because these participants stand with the scene camera at an angle closer to 90° to these targets and accompanying QR codes, the overall area on the scene camera recording that they cover is smaller than someone who stands at a taller height with the camera angled down. And, the QR codes are lost by the software more often as they are more likely to be in a position that has their dimensions altered with respect to the scene camera angle. The result is a greater number of missed fixations to higher targets in shorter participants. Furthermore, because the angle on the scene camera was closer to perpendicular in shorter participants, lower targets, like the Home area and Start/End Target, were also not viewed in the scene camera recording as much, thus resulting in instances where full fixations to these AOIs were not detected.

4.3 – Attentional Landscape Model in Object Movement

As has been discussed earlier, although humans have a central area of the retina with extremely high resolution, we are still able to gather information from the peripheral regions of our field of view. Numerous studies have been conducted using a 2-dimensional screen, to understand the ways in which we use the periphery of our field of view to plan movements (Baldauf et al., 2008; Baldauf & Deubel, 2008c, 2008b, 2008a, 2010; Baldauf et al., 2006). With this Attentional Landscape Model in mind, do humans dedicate varied levels of covert attention to different regions when interacting with objects in the real-world?

Recall from previous work on the Attentional Landscape Model of visual attention, when participants were required to fix their gaze on a central cross before shifting their fixation to touch a target with their finger, they were able to discriminate letters that flash on the screen before they started their movement more accurately at targets that they would reach towards than at those they would not (Baldauf & Deubel, 2010). And, Baldauf and Deubel showed that it is not only the first target that a participant will reach to that has heightened levels of covert attention given to it, but at least three future targets, with targets that will be interacted with sooner having greater levels of covert attention given to them than targets that will be interacted with later. Additionally, it was found that targets with more difficult requirements for success (smaller targets, or targets removed during reach), are given higher levels of covert attention than they would otherwise be given. Refer to Figure 4 for a visual representation of this.

If we extend the Attentional Landscape Model to the movement of sequential objects, like in the PB and C tasks, covert attention likely changes fluidly as a person moves within a phase of movement and shifts from transferring one object to the next. As well, it is highly probable that variables like the type of objects and their location, the type of grasps and their distance from the body, and any obstacles in the environment play a crucial role in the level of covert attention being dedicated to objects and their drop-off locations. As well, and possibly the most important variable to account for when discussing covert attention in object movement, is whether or not an area of interest (an object or its drop-off location) is within a person's field of view or not.

4.3.1 – Objects Visible in Field of View

The proposed Attentional Landscape Model of Object Movement when objects are visible in a person's field of view is most easily characterized in our study by the Cups task. Using the C task as an example, I propose a model for the shifting of covert and overt attention during sequential object movement. At the beginning of the task, the participant is overtly fixated on the Neutral eye position (shown by the yellow line in Figure 18 a). As they await the go cue, relevant objects and areas within their environment receive differing levels of covert attention, shown by the spectrum of red to white peaks. Based on Baldauf and Deubel's work showing that covert attention is differentially divisible with greater levels being dedicated to areas that will be acted upon sooner by the hand (2010), the Green cup has the highest level of covert attention given to it, with the Green cup's drop-off location having the next highest level of covert attention, with the Blue cup and its drop-off location following in descending order. This is an example of participants covertly preparing for the actions of their eyes and their hand in the near future.

Recall from previous literature that to make an overt shift of attention to an area, a person first must immediately shift their covert attention to this area. This supports our above proposed preparation model before the task begins, as every participant shifted their overt attention to the Green cup, as shown in Figure 18 (b), after the go cue was administered. With overt attention fixated on the Green cup as the hand approaches it, there is a shift in the levels of covert attention dedicated to the subsequent areas of action in the task. The drop-off location of the Green cup, on the other side of the partition, now has greater covert attention dedicated to it. This is facilitated by the fact that the brain no longer needs to dedicate covert attention to the Green cup as it has both overt visual attention and proprioceptive information from the hand about the Green cup. This excess covert attention can now be spread across the future areas of action in the remaining phases of the task, the Green cup drop-off location, the Blue cup, and the Blue cup drop-off location, with the lion's share going to the area that the hand will act at soonest, the Green cup drop-off location.

Once the confidence threshold has been met through the integration of overt visual attention and hand proprioceptive information about the grasp of the Green cup, the eyes saccade directly to the next highest level of covert attention, the Green cup drop-off location. This occurs during the grasp. And, the covert attention that had been dedicated to this drop-off location is now spread over to the next areas of hand action, the Blue cup and its drop-off location. Figure 18 (c) shows the participant transporting the Green cup over the partition with overt attention dedicated to its drop-off location, and heightened levels of covert attention now devoted to the Blue cup and its drop-off location, with the greatest amount at the soonest target of hand action, the Blue cup. Similar patterns of changing levels of covert attention play out over the remainder of the task.

Recent findings that show there are bimodal neurons in the brain that respond to both visual and haptic information. And, due to the greater density of mechanoreceptors in the palm of the hand and greater representation of the palm of the hand in the somatosensory cortex, the palm of the hand has greater density of these bimodal neurons dedicated to it than other body parts (Brown et al., 2009). Therefore, when the eyes are fixated on or near the palm of the hand, people enjoy greater levels of visual acuity. Therefore, the hand provides both haptic and proprioceptive information to its owner and bimodal neurons allow the owner to have greater accuracy of hand placement when the eyes are fixated on it. So I propose that the participant's hand plays a role in their ability to dedicate their covert, and therefore overt, visual attention throughout a sequential object movement task. As the hand nears its area of action, there is a subconscious increase of confidence in the part of the brain that decides where to fixate overt attention. For example, in Figure 18 (b), as the hand nears the Green cup, and covert attention is highest at its drop off location, a confidence level rises in the brain as the hand nears the Green cup to grasp it. At the point where the confidence level reaches a minimum threshold, which happens at some point during the grasp, a saccade is initiated to the highest area of covert attention, which happens to be the Green cup drop-off location (Figure 18 c). I propose that in people affected by compromised haptic and proprioceptive information, and therefore lower density of bimodal neurons dedicated to their limb, this minimum confidence level is greater, which results in overt attention shifting to the drop-off location later.

Obstacles must also play a role in the dispensing of covert attention around an environment during object movement. In the C task, the partition over which cups are transported is a prime example. As participants perform the task, there is very little, if any, overt visual attention being paid to it. Unlike the cup, which receives information from the hand, the partition must be signaling the participant to stay away from it somehow. Building from the obstacle avoidance literature that shows people change their reaching trajectory and grasping patterns based on the position and size of obstacles in the environment (Chapman et al., 2011; Chapman & Goodale, 2008), I propose that the participant dedicates covert attention to obstacles in the environment, but with a 'negative' value. For example, obstacles that are easily avoided will have covert attention dedicated to them, but only a small amount, and with a negative value (to signal the hand should stay away from them). And, obstacles that are more difficult to avoid will also have covert attention dedicated to them with a negative value, but will have a much larger amount. The larger the negative value of covert attention assigned to an object, the greater the change in hand movement path would be seen.

Now, it is important to stress that this Attentional Landscape changes fluidly as a person carries out a sequential object movement task. Before the task begins, these levels may be very constant, but once the task has begun, I envision an Attentional Landscape which rises and falls at different areas, like an ocean rises and falls during a storm. Variables that interact with one another and contribute to how these waves move include: the timing of overt fixations, the size of obstacles positioned in the environment, the hand's location, the type of grasp, the distance of the object from the body, previously interacted objects becoming obstacles, and the amount of time the brain needs to overtly fixate an object to reliably compute grasp information from it.

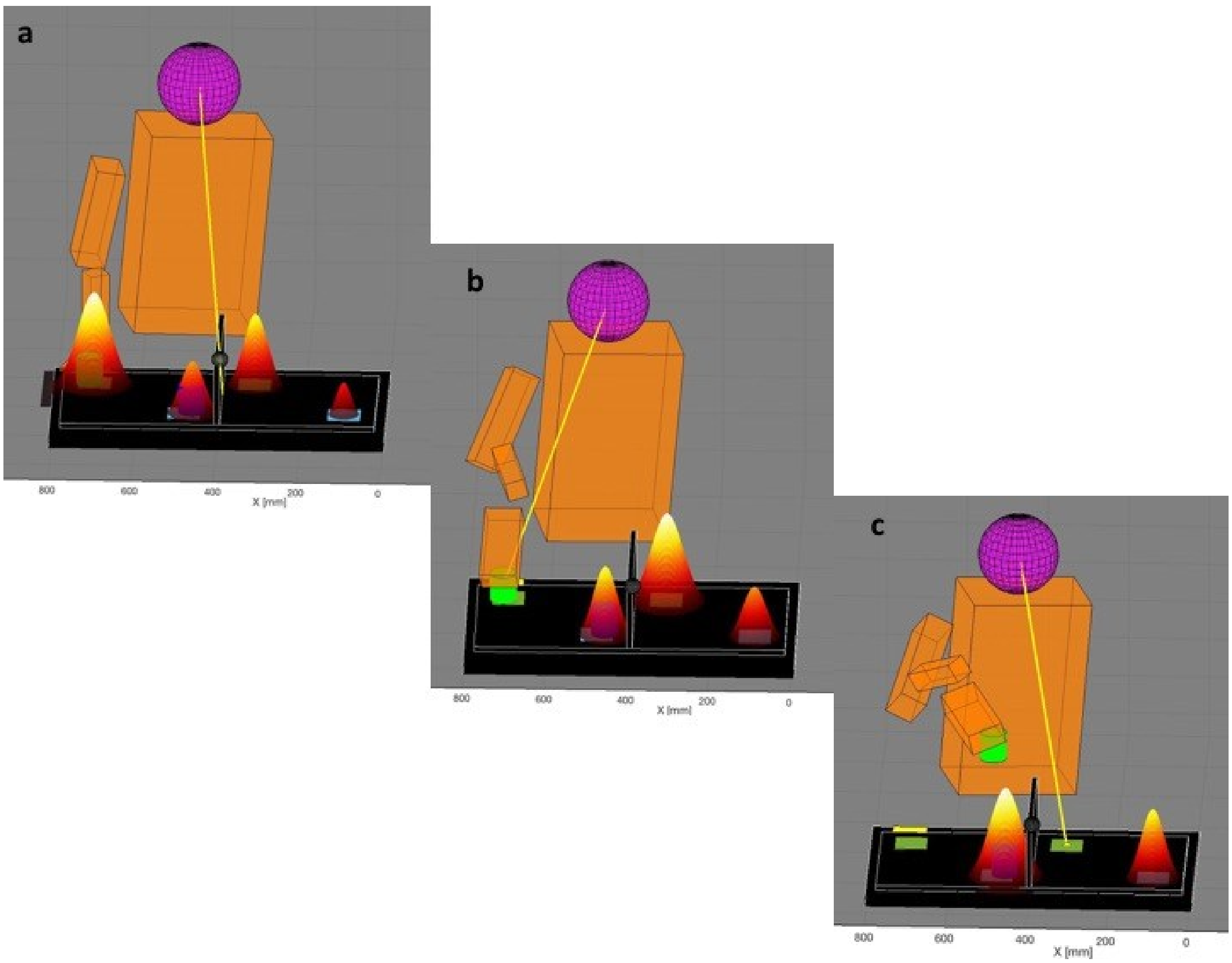

Figure 18: A virtual representation of the Attentional Landscape during sequential object movement. a) The participant fixates on the Neutral eye position, with covert attention on the Green cup, the Green cup drop-off location, the Blue cup, and the Blue cup drop-off location, with attention levels decreasing sequentially. b) The participant fixates on the Green cup, with covert attention now increased at the Green cup drop-off location and at the Blue cup and its drop-off location; c) The participant fixates on the Green cup drop-off location, with increased covert attention on the Blue cup, and the Blue cup drop-off location.

4.3.2 – Objects Not Visible in Field of View

How does this proposed Attentional Landscape Model of Object Movement differ when objects are not easily visible in a person's field of view? The biggest effect of this change is that a person is not be able to rely on visual information in planning an object interaction, but most likely relies on a form of working memory, at least in the initial stages, to complete an object interaction. Using the Two Visual Stream Hypothesis (for a review, see Goodale, 2011), which elucidates the differences between vision-for-action (Dorsal Stream) and vision-for-perception (Ventral Stream), we can further build on the Attentional Landscape Model of Object Movement when objects are outside the field of view. In brief, it has been shown in humans, that the Dorsal Stream receives visual input originally from the retina via the primary visual cortex (V1), and through a series of neural connections, transmits signals dorsally in the brain to the Posterior Parietal Cortex. Quite separate from the Dorsal Stream, the Ventral Stream also takes information from V1 but transmits signals ventrally in the brain to the Inferotemporal Cortex. For a simple visual of this, see Figure 19 which has been duplicated from Goodale's 2011 review paper, *Transforming vision into action*. The Dorsal Stream "is designed to operate in real time and is not normally engaged unless the target object is visible during the programming phase, when (bottom-up) visual information can be immediately converted into the appropriate motor commands" (Goodale, 2011). The Ventral Stream, on the other hand, "plays a fundamental role in constructing our perceptual representations of the world – but is not essential for the programming and online control of visually guided actions [to] goal objects that [it] has helped identify" (Goodale, 2011).

In our experiment, only the PB task has interactions to objects and areas outside the participant's field of view. Without the Dorsal Stream and its immediate visual information, the

Ventral Stream must be used with its memory of a representation of the location, size, and shape of objects in the environment. In each trial, the participant had to make 4 actions to areas that they did not have immediate visual information from. At the beginning of the task, they turn to the right to pick up the pasta box, and then must turn back towards the cart to place the box on the shelf. Near the end of the task, they must pick up the pasta box from the furthest shelf, turn to the right to place it on its original location, and then turn back and touch their hand on the Home area. Although separate, there is literature showing that connections from the Ventral to Dorsal Stream do exist, which allow transmission of information from the Ventral Stream to the Dorsal Stream to help plan actions to objects outside the current field of view. In all of these instances, it is expected that the Ventral Stream is more active, as information about the location and sometimes an object at the location has to be utilized in order to plan the hand's action. Consistent with literature showing that grip apertures are larger when subjects are initially blind to an object (Hu et al., 1999), the distance that a grasp was defined at the Start/End Target of the PB task was much greater than the other two targets (See Release of Movement 1 of the Pasta Task in Table 1). However, the Release distance at the Start/End Target (Release of Movement 3) was unaffected by this, perhaps because by the time a release was occurring, the Dorsal Stream was able to supply information about the object to the hand.

I propose that during sequential object movement involving areas outside the field of view, a similar Attentional Landscape is at work to areas within the field of view. However, attention, although not traditional covert visual attention, is also dedicated to areas outside the field of view. And, having to pay attention to an area outside the field of view, in preparation for a grasp, for example, will decrease the available covert attention to be used at areas within the visual field. For example, in the PB task, as participants are looking at the Neutral eye position, they are dedicating attention to their right side where their Ventral Stream approximates the pasta box to be sitting. As well, they are covertly giving visual attention to the drop-off location of the pasta box at the Mid Shelf Target. When the go cue occurs, they turn their bodies, and their Ventral Stream supplies information about the location of the pasta box, and as their visual field brings the pasta box into view, their covert attention locates it in order to make a saccade to overtly fixate the eyes on it. This enables the hand to accurately reach and grasp the box. Once proprioceptive information from the hand has reached a confidence threshold in the brain, the overt fixation on the box ends and the participant turns their visual field back towards the cart. Through this turn, and until the box is fully in view, their Ventral Stream supplies information about the general location of the Mid Shelf Target drop-off location, and as their visual field brings this target into the field of view, covert attention is paid to it, triggering an overt fixation, and enabling safe transport of the pasta box onto the shelf.

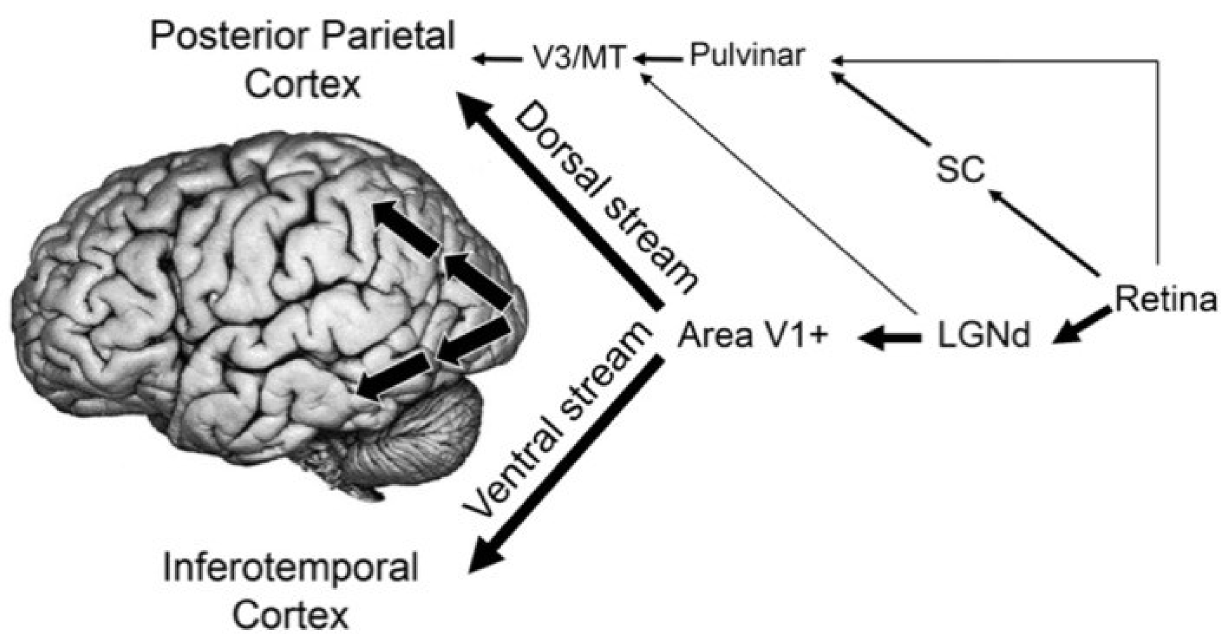

Figure 19: A visualization of the Two Visual Stream Hypothesis: the Dorsal Stream (vision-for-action), and the Ventral Stream (vision-for-perception). Duplicated from *Transforming vision into action* (Goodale, 2011).

4.4 – Flow of Sensory Modalities in Object Movement

The previous proposal on the Attentional Landscape Model with objects outside the field of view seems to converge close to the Schema Theory proposed by Land and Hayhoe (2001). Referring to Figure 5 (b), Land and Hayhoe describe how the Schema employs the Gaze system to search for the next object, or remember its previous location. This is similar to what the Ventral Stream would be doing in the previous section. The body then moves to direct the gaze in the general region of the object, and then the Visual system fixates the object, which enables the Motor system to begin its action toward the object. In the hypothesis described in the previous section, this aligns with the Dorsal Stream fixating the object and enabling the person to initiate a reaching and grasping movement towards the object.

Continuing with reference to the work of Land and Hayhoe, we noticed many similarities in our findings to theirs. But first, we will discuss a few differences. The breakdown of fixations into Locating, Directing, Guiding, and Checking were not easily noticeable in our tasks. Our tasks were rather simple when compared to "making a sandwich" or "making a cup of tea", and so we were not given the opportunity to record many Checking fixations. However, in both the PB and C tasks, Locating, Directing, and Guiding fixations were noted, but were not necessarily different from one another. For example, a fixation to the Green cup at the beginning of a trial could be called a Locating fixation, but as the hand reaches and begins to grasp the cup, there is no change in fixation to constitute a Directing fixation of the hand. Similarly, as the Green cup is transported over the partition, the eyes stay fixated on the drop-off location, which could constitute a Guiding fixation. But, this is not entirely different from the previous Locating and Directing fixations. Perhaps if our tasks were more complex, we would be able to pick out more noticeable differences between these types of fixations.

Recall work from Pelz and colleagues that the eyes often were the limiting factor when picking up and moving small blocks, in that the hand and body stalled their movements to allow the eyes to lead them (2001). And, Land and Hayhoe found a roughly half-second buffer of visual information required by the brain to prepare the hand to interact with an object. We observed similar findings and believe these are the result of the same phenomenon. First, we saw a slow movement of the hand at the beginning of both of our tasks, likely because the eyes had to travel from the Neutral eye position to the first object. Because the eyes had to travel a relatively far distance compared to other ORAs, they would not be fixated on the object with enough lead time to compute the necessary pick-up schematics. So, the hand moves at a slightly slower speed to allow the eyes time to fixate the object for just under a half-second to compute the grasp schematics. If the hand had moved at full speed, it would not have the necessary information computed in time to initiate an effective grasp, and would likely cost the participant by unsuccessfully grasping the object, forcing the participant to stop moving entirely to compute the necessary information, or dropping the object at some point during the transport. In essence, by slowing down the hand movement at the start of the task, the user is ensuring the remainder of the task can go faster. But, what does this 0.5s buffer allow for? There are many types of information that must be computed to ensure successful grasping of an object. Location, size, shape, weight, texture, type of grasp, and probable future grasping and lifting forces are just a few of the shopping list of items that the brain needs to have information about to grasp an item successfully. From our findings, the minimum amount of time needed to grasp household items like a pasta box or a cup is half a second. But when an item is out of sight, participants are less likely to get the full 0.5s of grasp computation time, and therefore have greater uncertainty about their probability of success, and consequently institute fail safes like increasing their grip aperture to increase their probability of success.

I agree with Hayhoe's description of vision being a "scarce and valuable resource" that can be "disengaged from a particular aspect of an action as soon as another sense is available to take over (Land & Hayhoe, 2001). It seems that in sequential object movement, the minimum amount of time is about half a second of visual information from an object to be interacted with for a successful grasp to take place, and the eyes will make a saccade to the next object of interaction to ensure a fluid flow of the hand and body. The other sensory modalities that are obviously at work in sequential object movement are proprioception and haptic feedback. Since visual information is powerful, but also scarce, we allocate it to areas that need it to complete a task efficiently, and employ other sensory modalities, like proprioception and haptic feedback, to duties that they can accomplish effectively. For example, during a reach and grasp, a person will fixate on an object they are going to interact with, and then at some point during the grasp, they will saccade to the drop-off location. Perhaps, they saccade because the eyes need to fixate on the drop-off location far enough in advance to ensure that the object is moved efficiently to it? And so, proprioception and haptic feedback from the hand allow the eyes to make this shift of overt attention.

The findings from the Two Visual Stream Hypothesis also have a place in this. It can be argued that when interacting with objects, people are hard-wired to optimally interact with objects by employing their Dorsal Stream toward objects that are to be picked up in their field of view, rather than having to recall information from their memory about the object while using their Ventral Stream. The Dorsal Stream does real-time calculation which does not utilize any information storage or recall. This makes sense from a time and energy efficiency perspective.

Why waste time and energy having to recall information from the Ventral Stream about an object and relay that information to the Dorsal Stream, if the eyes can fixate the object immediately using the Dorsal Stream? If a person can interact with an object by using their Dorsal Stream without having to recall information about it from their Ventral Stream, they will do it. So, one could argue that using the Ventral Stream is likely more of a last resort in sequential object movement if the Dorsal Stream is available to be used immediately. But, if the eyes are occupied and the next area of interaction is outside a person's field of view, the Ventral Stream will be engaged to prepare the head and body to turn towards this next area of interaction.

I propose a Flow of Sensory Modalities in Object Movement in which the Schema, the Ventral Stream, the Dorsal Stream including covert and overt visual attention, and proprioceptive and haptic information work together to allow people to interact with objects reliably and efficiently. For example before the start of the PB task, the schema identifies the relevant objects and areas in the task. Once the task starts, to grasp the pasta box at the Start/End Target at the beginning of the task, there is heavy use of the Ventral Stream to direct the field of view to the object, in order to allow the participant to initiate covert attention, and then overt attention to the pasta box. Without the memory representation and mental modeling functions of the Ventral Stream, this would not be possible. Next, overt attention to the pasta box, the Dorsal Stream functions, allows the participant to compute, subconsciously, the necessary information to enable the hand to successfully grasp the box. Without overt visual attention and real-time computation functions of the Dorsal Stream, this would not be possible. This real-time computation takes approximately 0.5s. While the eyes are overtly fixated on the pasta box and initiating a reach and grasp, the Ventral Stream is likely preparing the body to turn the field of view towards the drop-off location. And, once the hand is in the process of initiating a grasp, a confidence level is reached through a combination of proprioceptive and Dorsal Stream feedback, triggering a shift of the gaze away from the current object being grasped towards the next target of action. The body turns and directs the gaze to the drop-off location in order to dedicate covert attention and then overt attention there. The hand's proprioceptive and haptic abilities are relied upon to transport the object towards its drop-off location, while the Ventral Stream plans the next direction of body movement, and the eyes overtly fixate to enable the Dorsal Stream to plan the hand releasing movements on the drop-off location. Figure 20 depicts this idea as a flow of information.

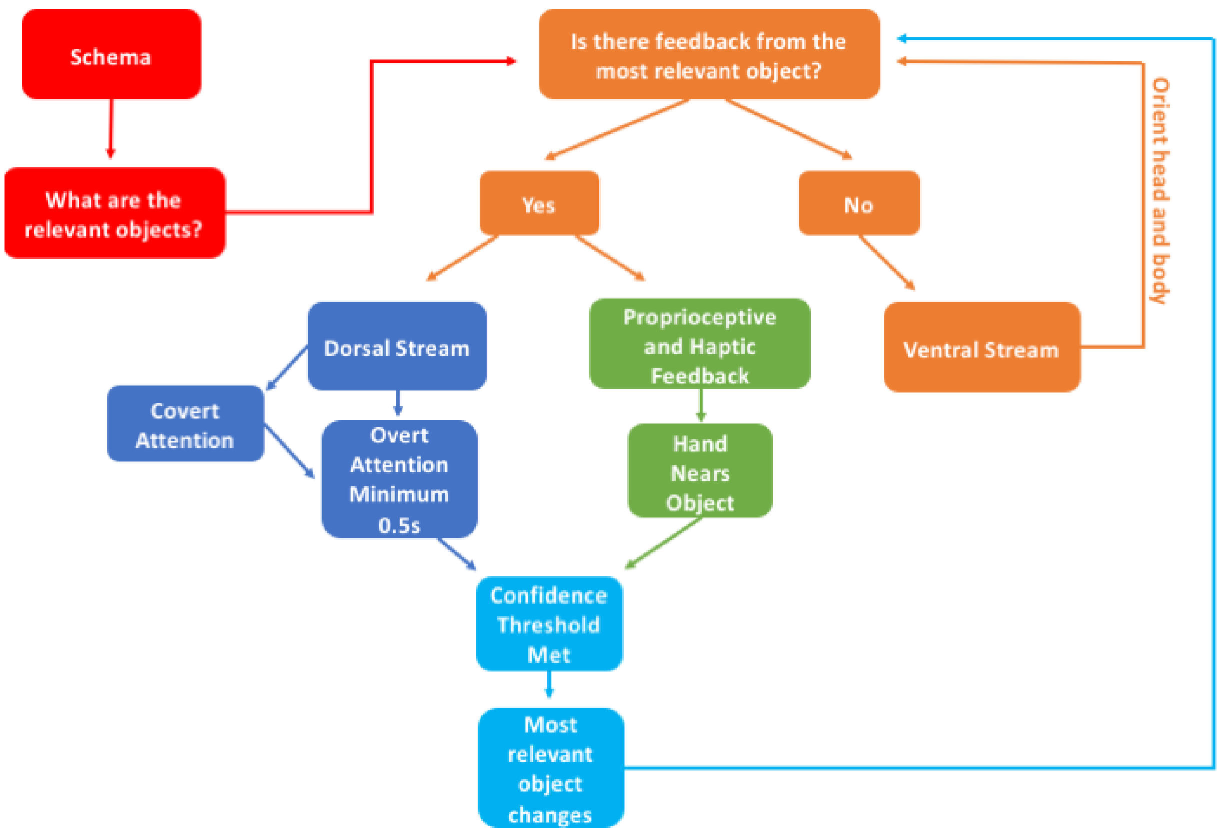

Figure 20: A visual representation of the Flow of Sensory Modalities in Object Movement.

4.5 – Future Directions and Final Remarks

This previous section has left several questions to be answered. The idea of an accumulation of confidence as the hand grasps an object that signals the eyes to make a saccade to the object's drop-off location is intriguing. As well, the proposed Flow of Sensory Modalities in Object Movement requires refining. Both of these can be investigated with the use of brain recordings like electroencephalography. Can we pinpoint an accumulation of confidence in the Dorsal Stream that initiates a signal to make a saccade? By studying certain pathological demographics, like upper-limb prosthetic users, we can explore some of the assumptions made about this Flow of Sensory Modalities to see if greater levels of overt visual attention are dedicated to certain areas during object interaction. As well, prosthetic users could require greater Ventral Stream usage to successfully interact with objects, as they are unable to rely on proprioceptive or haptic feedback to ensure accurate movement of their limb. This, in turn, may require, not only greater expenditure of overt attention, but also greater computation of information by prosthetic users as they move in, and interact with, the world. As well, prosthetic users may treat their limb as a separate object in the real-world, as opposed to a portion of their body, which would contribute to greater burden on both the Ventral and Dorsal Streams of Visual Information.

In closing, this thesis has analyzed information collected using a novel approach to object movement tasks. By enabling participants to move freely in the world while interacting with everyday objects, and allowing experimenters to hold the reins of experimental design, we have collected, processed, and defined eye movement behaviour in sequential object movement tasks with closer accuracy than ever before. Additionally, this experimental technique will be employed in clinical and research settings to assess and improve upper-limb prosthetic user functioning in the real-world, with further datasets having already been collected.

References

Anderson, B., & Ulvaeus, B. (1979). Angleyes [Recorded by ABBA]. On *Voulez-vous*. Polar Music Studio.

Baldauf, D., Cui, H., & Andersen, R. A. (2008). The Posterior Parietal Cortex Encodes in Parallel Both Goals for Double-Reach Sequences. *Journal of Neuroscience*, *28*(40), 10081–10089. https://doi.org/10.1523/JNEUROSCI.3423-08.2008

Baldauf, D., & Deubel, H. (2008a). Attentional Selection of Multiple Goal Positions Before Rapid Hand Movement Sequences: An Event-related Potential Study. *Journal of Cognitive Neuroscience*, *21*(1), 18–29. https://doi.org/10.1162/jocn.2008.21021

Baldauf, D., & Deubel, H. (2008b). Properties of attentional selection during the preparation of sequential saccades. *Experimental Brain Research*, *184*(3), 411–425. https://doi.org/10.1007/s00221-007-1114-x

Baldauf, D., & Deubel, H. (2008c). Visual attention during the preparation of bimanual movements. *Vision Research*, *48*(4), 549–563. https://doi.org/10.1016/j.visres.2007.11.023

Baldauf, D., & Deubel, H. (2010). Attentional landscapes in reaching and grasping. *Vision Research*, *50*(11), 999–1013. https://doi.org/10.1016/j.visres.2010.02.008

Baldauf, D., Wolf, M., & Deubel, H. (2006). Deployment of visual attention before sequences of goal-directed hand movements. *Vision Research*, *46*(26), 4355–4374. https://doi.org/10.1016/j.visres.2006.08.021

Ballard, D. H., Hayhoe, M. M., Li, F., Whitehead, S. D., Frisby, J. P., Taylor, J. G., & Fisher, R. B. (1992). Hand-Eye Coordination during Sequential Tasks [and Discussion].

Philosophical Transactions of the Royal Society of London B: Biological Sciences, *337*(1281), 331–339. https://doi.org/10.1098/rstb.1992.0111

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory Representations in Natural Tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80. https://doi.org/10.1162/jocn.1995.7.1.66

Bekkering, H., Adam, J. J., Aarssen, A. van den, Kingma, H., & Whiting, H. T. A. (John). (1995). Interference between saccadic eye and goal-directed hand movements. *Experimental Brain Research*, *106*(3), 475–484. https://doi.org/10.1007/BF00231070

Belardinelli, A., Herbort, O., & Butz, M. V. (2015). Goal-oriented gaze strategies afforded by object interaction. *Vision Research*, *106*, 47–57. https://doi.org/10.1016/j.visres.2014.11.003

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research*, *49*(24), 2992–3000. https://doi.org/10.1016/j.visres.2009.09.014

Brown, L. E., Morrissey, B. F., & Goodale, M. A. (2009). Vision in the palm of your hand. *Neuropsychologia*, *47*(6), 1621–1626. https://doi.org/10.1016/j.neuropsychologia.2008.11.021

Buswell, G.T. (1920). An experimental study of the eye-voice span in reading. *Supplementary Educational Monographs* No. 17. Chicago: Chicago University Press.

Buswell, G.T. (1935). *How People Look at Pictures: A Study of the Psychology of Perception in Art.* Chicago: Chicago University Press.

Butsch, R.L.C. (1932). Eye movements and the eye-hand span in typewriting. *J. Educational Psychology* 23, 104-121.

Chapman, C. S., Gallivan, J. P., Culham, J. C., & Goodale, M. A. (2011). Mental blocks: fMRI reveals top-down modulation of early visual cortex when obstacles interfere with grasp planning. *Neuropsychologia*, *49*(7), 1703–1717. https://doi.org/10.1016/j.neuropsychologia.2011.02.048

Chapman, C. S., & Goodale, M. A. (2008). Missing in action: the effect of obstacle position and size on avoidance while reaching. *Experimental Brain Research*, *191*(1), 83–97. https://doi.org/10.1007/s00221-008-1499-1

Clapton, E. (1998). My father's eyes. On *Pilgrim*. Reprise.

Collewijn, H., van der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*, *15*(3), 447-IN5. https://doi.org/10.1016/0042-6989(75)90098-X

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., … Shulman, G. L. (1998). A Common Network of Functional Areas for Attention and Eye Movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/S0896-6273(00)80593-0

de Haan, B., Morgan, P. S., & Rorden, C. (2008). Covert orienting of attention and overt eye movements activate identical brain regions. *Brain Research*, *1204*, 102–111. https://doi.org/10.1016/j.brainres.2008.01.105

D-Lab 3.0 Manual, 2017. http://www.ergoneers.com/faq/index.php?action=artikel&cat=4&id=52&artlang=en

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6–7), 790–811. https://doi.org/10.1080/13506280902793843

Deco, G., & Zihl, J. (2001). A Neurodynamical Model of Visual Attention: Feedback Enhancement of Spatial Resolution in a Hierarchical System. *Journal of Computational Neuroscience*, *10*(3), 231–253. https://doi.org/10.1023/A:1011233530729

Descartes, R. (1637/1902). *La dioptrique*. In C. Adam and P. Tannery (Eds.), *Oevres de Descartes*. Vol. 6. (pp. 81-228) Paris: Cerf.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4

Deubel, H., Schneider, W. X., & Paprotta, I. (1998). Selective Dorsal and Ventral Processing: Evidence for a Common Attentional Mechanism in Reaching and Perception. *Visual Cognition*, *5*(1–2), 81–107. https://doi.org/10.1080/713756776

Dodge, R., & Cline, T.S. (1901). The angle velocity of eye movements. *Psychol. Rev.* 8, 145-157.

Dohlman, G. (1925). Physikalische und physiologische Studien zur Theorie des kalorischen Nystagmus. *Acta Oto-Laryngologica Supplement*, 5, 1-196.

Du Laurens, A. (1599). A discourse of the preservation of the sight: of melancholic diseases; of rheumes, and of old age. *Trans R. Surphlet.* London: The Shakespeare Association.

Eimer, M., Van Velzen, J., Gherri, E., & Press, C. (2006). Manual response preparation and saccade programming are linked to attention shifts: ERP evidence for covert attentional orienting and spatially specific modulations of visual processing. *Brain Research*, *1105*(1), 7–19. https://doi.org/10.1016/j.brainres.2005.10.060

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & psychophysics*, *40*(4), 225-240.

Farid, M., Murtagh, F., & Starck, J. L. (2002). Computer display control and interaction using eye-gaze. *Journal of the Society for Information Display*, *10*(3), 289-293.

Findlay, J. M. (1997). Saccade Target Selection During Visual Search. *Vision Research*, *37*(5), 617–631. https://doi.org/10.1016/S0042-6989(96)00218-0

Frey, G., & Henley, D. (1975). Lyin' eyes [Recorded by The Eagles]. On *One of These Nights*. Hollywood, CA: Asylum.

Geisler, W. S., & Cormack, L. K. (2011). Models of overt attention. *The Oxford Handbook of Eye Movements*, 439-454.

Gherri, E., & Forster, B. (2012). The orienting of attention during eye and hand movements: ERP evidence for similar frame of reference but different spatially specific modulations of haptic processing. *Biological Psychology*, *91*(2), 172–184. https://doi.org/10.1016/j.biopsycho.2012.06.007

Goodale, M. A. (2011). Transforming vision into action. *Vision Research*, *51*(13), 1567–1587. https://doi.org/10.1016/j.visres.2010.07.027

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*, 43–64.

Helmholtz, H. (1867). *Handbuch der physiologischen Optik.* In G Karsten (Ed.), *Allgemeine Encyklopadie der Physik.* Vol. 9. Leipzig:Voss.

Henderson, J. M. (1992). Visual Attention and Eye Movement Control During Reading and Picture Viewing. In K. Rayner (Ed.), *Eye Movements and Visual Cognition* (pp. 260–283). Springer New York. https://doi.org/10.1007/978-1-4612-2852-3_15

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. Retrieved from https://core.ac.uk/display/22875642

Henriques, D. Y. P., Klier, E. M., Smith, M. A., Lowy, D., & Crawford, J. D. (1998). Gaze-Centered Remapping of Remembered Visual Space in an Open-Loop Pointing Task. *Journal of Neuroscience*, *18*(4), 1583–1594.

Henry, F. M., & Rogers, D. E. (1960). Increased Response Latency for Complicated Movements and A “Memory Drum” Theory of Neuromotor Reaction. *Research Quarterly. American Association for Health, Physical Education and Recreation*, *31*(3), 448–458. https://doi.org/10.1080/10671188.1960.10762052

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787–795. https://doi.org/10.3758/BF03206794

Hu, Y., Eagleson, R., & Goodale, M. A. (1999). The effects of delay on the kinematics of grasping. *Experimental Brain Research*, *126*(1), 109–116. https://doi.org/10.1007/s002210050720

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: linked or independent? *Cognitive Brain Research*, *18*(1), 102–105. https://doi.org/10.1016/j.cogbrainres.2003.08.006

Hunter, J. (1786). *Observations on Certain Parts of the Animal Oeconomy*. London.

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. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational Modelling of Visual Attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–Hand Coordination in Object Manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932.

Kepler, J. (1611). *Dioptrice*. Augsburg: Franci.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive Ethology: a new approach for studying human cognition. *British Journal of Psychology (London, England: 1953)*, *99*(Pt 3), 317–340. https://doi.org/10.1348/000712607X251243

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. S. Nickerson (Ed.),*Attention and performance* (Vol. 8, pp. 259–276). Hillsdale, NJ: Erlbaum.

Klein, R. M., & Hilchey, M. D. (2011). Oculomotor inhibition of return. Retrieved from http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199539789.001.0001/oxfordhb-9780199539789-e-026

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916. https://doi.org/10.1016/0042-6989(94)00279-U

Krieger, G., Rentschler, I., Hauske, G., Schill, K., & Zetzsche, C. (2000). Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics. *Spatial Vision*, *13*(2), 201–214.

Kulke, L. V., Atkinson, J., & Braddick, O. (2016). Neural Differences between Covert and Overt Attention Studied using EEG with Simultaneous Remote Eye Tracking. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00592

Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2012). A new look at social attention: Orienting to the eyes is not (entirely) under volitional control. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1132–1143. https://doi.org/10.1037/a0027075

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565. https://doi.org/10.1016/S0042-6989(01)00102-X

Land, M., Mennie, N., & Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. *Perception*, *28*(11), 1311–1328. https://doi.org/10.1068/p2935

Land, M. F., & Tatler, B. W. (2009). Looking and acting. *Vision and eye movements in natural behaviour*. Oxford Scholorship.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (Eds.). (2011). *The Oxford handbook of eye movements*. Oxford ; New York: Oxford University Press.

Martin, C. (2002). Green eyes [Recorded by Coldplay]. On *A rush of blood to the head*. Liverpool, UK: Capitol.

Merchant, J., Morrissette, R., & Porterfield, J. L. (1974). Remote Measurement of Eye Direction Allowing Subject Motion Over One Cubic Foot of Space. *IEEE Transactions on Biomedical Engineering*, *BME-21*(4), 309–317. https://doi.org/10.1109/TBME.1974.324318

Orschansky, J. (1899). Eine Methode die Augenbewegungen direct zu untersuchen. *Centralblatt fur Physiologie,* 12, 785-790.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123. https://doi.org/10.1016/S0042-6989(01)00250-4

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *ResearchGate*, *139*(3), 266–77. https://doi.org/10.1007/s002210100745

Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics*, *66*(3), 398–405. https://doi.org/10.3758/BF03194888

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(4), 726-744.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. *Attention and performance XV*, *15*, 231-265.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*(1), 31-40.

Rose, F. (1975). Blue eyes crying in the rain [Recorded byWillie Nelson]. On *Red Headed Stranger*. Columbia.

Ruete, C.G.T. (1857). *Lehrbuch der Opthalmologie fur Aerxte und Studirende.* (2nd ed.) Braunschweig: Vieweg.

Rybak, I. A., Gusakova, V. I., Golovan, A. V., Podladchikova, L. N., & Shevtsova, N. A. (1998). A model of attention-guided visual perception and recognition. *Vision Research*, *38*(15–16), 2387–2400. https://doi.org/10.1016/S0042-6989(98)00020-0

Scheiner, C. (1630). *Rosa ursina*. Bracciani: Phaeum.

Schill, K., Umkehrer, E., Beinlich, S., Krieger, G., & Zetzsche, C. (2001). Scene analysis with saccadic eye movements: Top-down and bottom-up modeling. *JOURNAL OF ELECTRONIC IMAGING*, *10*, 152–160.

Stark, L. W., & Choi, Y. S. (1996). Experimental metaphysics: The scanpath as an epistemological mechanism. In H. S. S. and C. F. W.H. Zangemeister (Ed.), *Advances in Psychology* (Vol. 116, pp. 3–69). North-Holland. https://doi.org/10.1016/S0166-4115(96)80069-0

Steinman, R. M., & Collewijn, H. (1980). Binocular retinal image motion during active head rotation. *Vision Research*, *20*(5), 415–429. https://doi.org/10.1016/0042-6989(80)90032-2

Underwood, G., & Everatt, J. (1996). Automatic and controlled information procession: the role of attention in the processing of novelty. In O. Neumann, & F. Sanders, *Handbook of perception and action*, vol. 2 9pp. 185-227). London: Academic Press.

Velzen, J. van, Gherri, E., & Eimer, M. (2006). ERP effects of movement preparation on visual processing: attention shifts to the hand, not the goal. *Cognitive Processing*, *7*(1), 100–101. https://doi.org/10.1007/s10339-006-0089-z

Weaver, H. E. (1943). A survey of visual processes in reading differently constructed musical selections. *Psychological Monographs* 55, 1-30.

Wells, W.C. (1792). *An essay upon single vision with two eyes: together with experiments and observation on several other subjects in optics.* London: Cadell.

Wilkie, R. M., & Wann, J. P. (2003). Eye-movements aid the control of locomotion. *Journal of vision*, *3*(11), 3-3.

Wundt, W. (1862). *Beitrace zur Theorie der Sinneswahrnehmung.* Leipzig: Winter.

Yarbus, A. (1967). *Eye Movements and Vision*. New York: Plenum Press.