Analysis of Visual Scanning Behaviours for the Objective Assessment of Psychiatric Disorder

by

Kai-Ho Fok

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Graduate Departments of Edward S. Rogers Sr. Department of Electrical and Computer Engineering

and Institute of Biomaterials and Biomedical Engineering
University of Toronto

Analysis of Visual Scanning Behaviours for the Objective Assessment of Psychiatric Disorder

Kai-Ho Fok

Master of Applied Science,
Edward S. Rogers Sr. Department of Electrical and Computer Engineering
Institute of Biomaterials and Biomedical Engineering
University of Toronto
2012

Abstract

Biases in selective information-processing or attention biases are features of most psychiatric disorders. Attention biases can be measured by monitoring visual scanning behaviour (VSB) which is directly linked to attention allocation processes. This thesis presents a general framework for studies of VSB when multiple images are presented simultaneously to the viewer. Within this general framework, a novel set of VSB parameters that characterize the different stages of the visual scanning process was developed. Using this set of parameters, biases towards thin and fat body shape images were detected in Anorexia Nervosa patients. A log-likelihood ratio detector utilizing VSB parameters had both high sensitivity (92%) and high specificity (90%). Preliminary results in VSB studies also show biases in adults with Major Depressive Disorder and elderly apathetic Alzheimer's patients. The development of sensitive physiological markers in individuals with mental illness is crucial to the advance of research, diagnosis, and treatment in psychiatry.

Acknowledgements

I would like to thank my supervisor Dr. M. Eizenman as well as Drs. L. Pinhas, P. Giacobbe, and L. Grupp for their guidance and mentorship throughout the course of this work. I would also like to thank Dmitri Model, my fellow lab member, and Eileen Lam, Reva Schachter, Sarah Chau, Chris Tang, William Sanh, Gabilasan Gengatharan, Amanjit Dhillon, Shayne Lin, and Alex Ji not only for their work on the projects but for their support and friendship. You have made my graduate school experience a fun and memorable one.

I would like to acknowledge the Vision Science Research Program for their financial support. I am grateful towards the staff and students of Edward S. Rogers Sr. Department of Electrical and Computer Engineering and the Institute of Biomaterials and Biomedical Engineering for providing a stimulating learning environment.

Finally, as always, I am deeply indebted to my parents, May Lam and William Fok, my brother, Kai-Lon Fok, and Leanne Grace Buzon for their endless patience, support, and encouragement.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

Mental illness, the second leading cause of human disability and premature death, affects thinking, mood or behaviour and can be associated with distress and/or impairment of functioning, with symptoms that vary from mild to severe (Dewa, Chau, and Dermer, 2010). One in five Canadians will experience a mental illness in their lifetime and the remaining four will have a friend, family member or colleague who will experience it (Health Canada, 2002). An estimated $51 billion was the total cost of mental illness to Canadian economy in terms of health care and lost productivity in 2003 (Lim et al., 2008) and according to the World Health Organization, depression will be the single biggest medical burden on health by 2020. Therefore, there is an urgent need to undertake research in psychiatry because of the scale, challenges, and relative lack of understanding in comparison to other medical fields (Royal College of Psychiatrists, 2011).

1.1. The Need for Objective Indicators

One important need is for objective indicators of psychiatric disorder. In 1973, it was reported that 12 pseudopatients, who pretended to hear hallucinations, were hospitalized and received a diagnosis of schizophrenia even though they acted normal after admission (Rosenhan, 1973). As a resolution to this crisis, the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) was published in 1980 to address the unreliability of psychiatric diagnosis (Nesse and Stein, 2012). Revolutionizing psychiatry, clinical impressions were replaced with checklists of operationalized indicators (DSM-III, 1980). This change created reliably defined conditions and standardized interviews researchers could use to collaborate, search for pathology, and measure prevalence and outcomes of treatment (Spitzer et al., 1992; Wilson, 1993). Though this operationalization of diagnosis remained in later revisions of the DSM, psychiatric nosology is still highly criticized because of its focus on utility rather than validity (Kendell and Jablensky, 2002). Excluding neurological disorders, none of the main DSM mental disorders can be validated by laboratory or imaging biomarkers (Ness and Stein, 2012). While other medical disorders are defined by specific etiology or distinctive anatomical or molecular abnormalities, no comparable objective indicators have been found for any major mental disorder (Insel and Wang, 2010). Some neurobiological differences (such as brain imaging) may

characterize certain diagnostic groups but they are neither specific nor sensitive enough to validate any diagnosis. A need for such objective markers arises from several current challenges facing psychiatry: heterogeneity of patients within diagnostic groups, a need for early indicators of treatment efficacy, and comorbidity. Heterogeneity occurs when two patients qualify for the same diagnostic group but do not share any specific symptoms in common. Furthermore, there is little evidence that supports the existence of natural boundaries that separate currently recognized mental disorders and drugs do not respect the boundaries of the disorders. For example, antidepressants are found to be effective in treating anxiety disorders and obsessive compulsive disorders as well as depression (Kendell and Jablensky, 2003; Hyman, 2007). Since treatment is dependent on diagnosis, objective markers would help identify potential subgroup or between-group patients where a modified treatment plan would be more effective and can contribute greatly to the understanding of the progression of illness and response to treatment. Objective markers may identify early signs of response to treatment and highlight patients who are at risk of relapse.

Comorbidity, which occurs when patients qualify for multiple diagnoses, is highly prevalent (Kessler et al., 2005). A high rate of comorbidity may naturally occur if one illness is a risk factor for another illness but may also reflect the scenario where different symptoms arise as a result of shared risk factors (Kendler et al., 1992). Objective markers will provide understanding of comorbidities and guide treatment selection. For all the above reasons, there is a strong push towards finding objective biomarkers in psychiatry. Current approaches include finding genetic/epigenetic markers (Kendler et al., 1992) and integrating neurological brain circuit models (Hyman, 2007; Insel and Wang, 2010) but the long and exhaustive search has yet to be fruitful.

1.2. Visual Scanning Behaviour as an Objective Indicator

Genetic, biological and environmental factors contribute to the development of psychiatric disorders. However, previous efforts have proven that studying this "ballet choreographed interactively over time among genotype, environment, and epigenetic factors, which gives rise to a particular phenotype" (Gottesman, 2003) is challenging. Instead, it may be more important to identify measurable physiological components that reflect these interactions along the pathway between the observed phenotype and distal genotypes and biological factors. These measurable

components that cannot be seen by an unaided eye (endophenotypes) can be an objective indicator that is useful in the diagnosis, classification and evaluation of treatment of patients with mental illness (Gottesman, 2003).

This thesis describes a potential endophenotype: visual scanning behaviour (VSB). VSB is based on the observation that in humans, detailed information (high resolution, colour, etc) of objects-of-interest in the visual field (approximately 180°*135°) is obtained by moving the point-of-gaze (eye movements) so that the objects' images fall on the rod-free, capillary-free portion of the retina - the foveola- (0.3 mm, approximately 1°). The patterns of movement (visual scanning patterns) provide continuous records of regions in the visual field that are considered relevant by the subject. Visual scanning is controlled by both low-level perception processes (e.g., temporal and spatial characteristics of the visual stimuli) and cognitive processes, which are driven by the subject's memories, emotions, expectations and goals. VSB provides not only behavioural end products of cognitive processes but also clues to the process through which these products are achieved. Moreover, during natural viewing of visual stimuli, subjects are unaware of their eye movements and VSB can be monitored without requests for meta-cognitive reports or other overt responses.

VSBs have been shown to illuminate information processing at several levels which range from perception to imaginative processes. Object copying tasks reveal a cognitive process of "indexing" whereby the spatial location on an object is maintained in working memory and other properties of the object are "looked-up" while needed (Ballard et al., 1997). A study involving the recall of auditory information revealed that information is coupled with moving visual stimuli whenever possible. Further, the recall of the information produces significantly more eye movements towards the last seen position of the visual stimulus even when the visual stimulus is no longer present. This finding suggests that eye movements are sensitive to partially active representations in cognition that may not be detectable by most other experimental measures (Richardson and Spivey, 2000). When there are no spatial cues to couple information presented to subjects, eye movements have been demonstrated to reveal imaginative processes. In one particular study, subjects' eye movements were recorded while they listened to spoken descriptions of spatiotemporally dynamic scenes and faced a large white projection screen that

took up most of their visual field. When subjects were told, "Imagine that you are standing across the street from a 40-story apartment building. At the bottom there is a doorman in blue. *On the 10th floor, a woman is hanging her laundry out the window. On the 29th floor, two kids are sitting on the fire escape smoking cigarettes. On the very top floor, two people are screaming.*" they made reliably more upward saccades than any other direction while listening to the italicized sentences (Spivey et al., 2000). These studies demonstrate that eye movements provide a rich source of observations at many levels of information processing.

Evidence that VSBs may be useful in identifying psychiatric disorders comes from the presence of attention biases in psychiatric patients. Attention bias is a selective information-processing bias that occurs early in the information-processing sequence, often independently of awareness or intent (Mathews and MacLeod, 2005) and refers to the tendency to selectively attend to disorder-relevant stimuli (Williamson et al. 2004). Traditionally, attention bias has been studied indirectly using various forms of the Stroop and dot-probe tests. The Stroop colour-naming task (where words are printed in different colors and subjects are asked to name the color) operates on the premise that increased reaction time means a subject is distracted by the meaning of the word. This distraction is then inferred as attention or cognition bias (Stroop, 1935). The dot-probe test (where two stimuli are simultaneously shown for a short time and subjects are asked for the location of the probe) is also used to indirectly study attentional bias and operates on the premise that decreased reaction time when the probe is on the side of the stimulus represents attention bias. Pictures as well as words can be used in these tests and studies generally show similar results to the Stroop tests (MacLeod, Mathews, and Tata, 1986). Dot-probe and Stroop tests are indirect measures of attention bias and as such they cannot be used to rule out the hypothesis that non-attention processes are responsible for the observed response bias (e.g. button-press responses utilized in the studies can be affected by psychomotor changes). As well, the Dot-probe and Stroop tests are usually presented for brief time periods and therefore are more suitable for the measurement of attention bias in the initial response to the appearance of the stimulus rather than attention biases that are associated with continuous interrogation of an image. Finally, the Dot-probe and Stroop tests provide only indirect assessment of gaze position at the end of the presentation of the stimuli (words, pictures, etc.) and thus cannot provide a full picture of the manner in which the word/picture was observed/looked at/interrogated by the

subject. These limitations affect the ability of these two methodologies to detect small changes in attention bias, necessitate many repetitions, and prevent them from being used as objective indicators of psychiatric disorder.

In the last decade, multiple studies have utilized a new method of assessing attention: that of directly studying gaze (VSBs), which can provide a continuous measure of attention (Hermans et al. 1999; Toh, Rossell, and Castle, 2011). This method is unobtrusive and has no response bias, as the subject has to view images on the screen. In a new paradigm that uses direct analysis of eye movements to estimate attention biases (Eizenman et al., 2003), subjects' visual scanning patterns are analyzed when competing visual stimuli (multiple images) are presented simultaneously and subjects are allowed adequate time to scan and re-scan the images. Detailed analysis of these visual scanning patterns results in the estimation of a set of visual scanning parameters that describes details of the attention allocation process that cannot be assessed in Dot-probe or Stroop tests (Eizenman et. al, 2003, Yu and Eizenman 2004). Using the new paradigm in depression, Eizenman et al. (2003) found that not only do depressed subjects have longer total fixation time to dysphoric images, but that they also had difficulty shifting attention away from dysphoric images. Subsequent studies using similar paradigms were able to validate the findings (Caseras et al., 2007; Kellough et al., 2008) and apply it to measure biases in anxiety disorders (In-Albon et al. 2009, Rinck and Becker 2006) and eating disorders (Jansen et al., 2004; Smeets et al., 2011). However, none of these studies have yet to identify a measure sensitive enough to detect differences in behaviour between individuals. To reach the goal of finding objective indicators of mental illness, more sensitive parameters must be derived from the information-rich VSBs. To realize this objective, there is a need to test a wide array of possible mechanisms that explain the observed biases between psychiatric and non-psychiatric populations.

1.3. Goals and Organization of this Thesis

The main goals of this thesis are a) to develop a general framework for studies of VSB in psychiatric disorders and b) to demonstrate that parameters of VSB can be used as objective indicators of mental illness.

Chapter 2 describes a general system for studies of VSB in patients with psychiatric disorders. The system has three components: visual stimuli creation and presentation, data recording, and analysis. First, a method to select and present visual stimuli in a manner which evokes differences in VSB is discussed. Secondly, the process of recording VSBs and linking them to the visual stimuli is described. Lastly, visual scanning parameters are described to characterize the various aspects of VSB. These visual scanning parameters are used in Chapters 3, 4, and 5 to characterize behaviour in psychiatric patients and provide evidence that objective indicators can be obtained.

Chapter 3 describes how attention biases in the adolescent Anorexia Nervosa population may be characterized by visual scanning parameters. It is shown how a simple log-likelihood ratio detector can use visual scanning parameters to identify individual patients with high sensitivity and specificity. Further, it is shown how the detector may be modified to produce better performance by incorporating information from several mechanisms of bias.

In Chapter 4, visual scanning parameters of patients with Major Depressive Disorder are analyzed. Depressed participants and their VSBs are followed over 8 weeks of a drug trial in order to find out if VSB can be used as an early indicator of antidepressant efficacy. The variability that is found between individuals is discussed.

Chapter 5 describes how VSB may be utilized to distinguish Apathy from Depression in seniors with Alzheimer's Disease. The visual scanning parameters that include temporal information of VSBs improve specificity when identifying apathy in individuals compared to visual scanning parameters that assume visual scanning is a stationary process. Additionally, a potential method to identify depressed Alzheimer's patients is discussed.

Chapter 6 presents a summary of the contributions of this thesis and avenues for further research.

Chapter 2. Framework for Studying Visual Scanning Behaviours in Psychiatric Disorders

2.1. Overview

Visual scanning biases may be elicited by presenting patients with stimuli congruent with their psychiatric disorder. The paradigm developed by Eizenman et al.in 2003 utilizes multiple observations of visual scanning responses to gain statistical confidence in the behaviour measured for an individual. In this paradigm, visual stimuli are presented for relatively long time periods (e.g., 10 seconds) and the analysis of the subject's eye movements support the detection of biases to both early and late processing of the images. In the analysis, sequences of eye movements are summarized into visual scanning parameters to characterize the visual scanning behavior of an individual. As the control of visual scanning behaviour is affected by both high-level cognitive processes and low level perceptual processes the framework to study visual scanning behaviours must be flexible enough to test a wide array of hypotheses about the mechanisms that lead to biases in visual scanning processes. Such a system is described in the following chapter and has three main components: Presentation Creation, Data Recording, and Analysis.

2.2. Presentation Creation

Visual scanning bias is defined as the difference in visual scanning behaviour between a reference group (usually, age-matched control subjects with no psychiatric disorders) and patients with psychiatric disorders. The objective of stimuli presentation is to maximize the difference between the visual scanning behaviour of the two groups. When multiple images are presented simultaneously, they compete for the viewer's attention. The relative attraction of one image over another at specific time instances during the visual scanning processes provides a measurement of visual scanning bias. Visual scanning bias can be a function of processes such as rumination, avoidance, and stimuli-checking (Rawal et al., 2010) or perceptual processes such as the colours and the contrasts of the images. For a specific study, images are selected so as to maximize differences in the visual scanning behaviour between the tested groups. For example, the two graphs in Figure 2.1 show that the difference in glance durations when an image is visited for the first time (one of the VSB parameters), between controls and patients with

anorexia nervosa, is not constant for the set of images that was used in the study. For some images (i.e., image 6 and 28) the mean values of the two populations are the same and for other images (i.e., 10-15) the mean values are significantly different. Images that attract the same amount of attention from both populations are less likely to help in differentiating control subjects from patients.

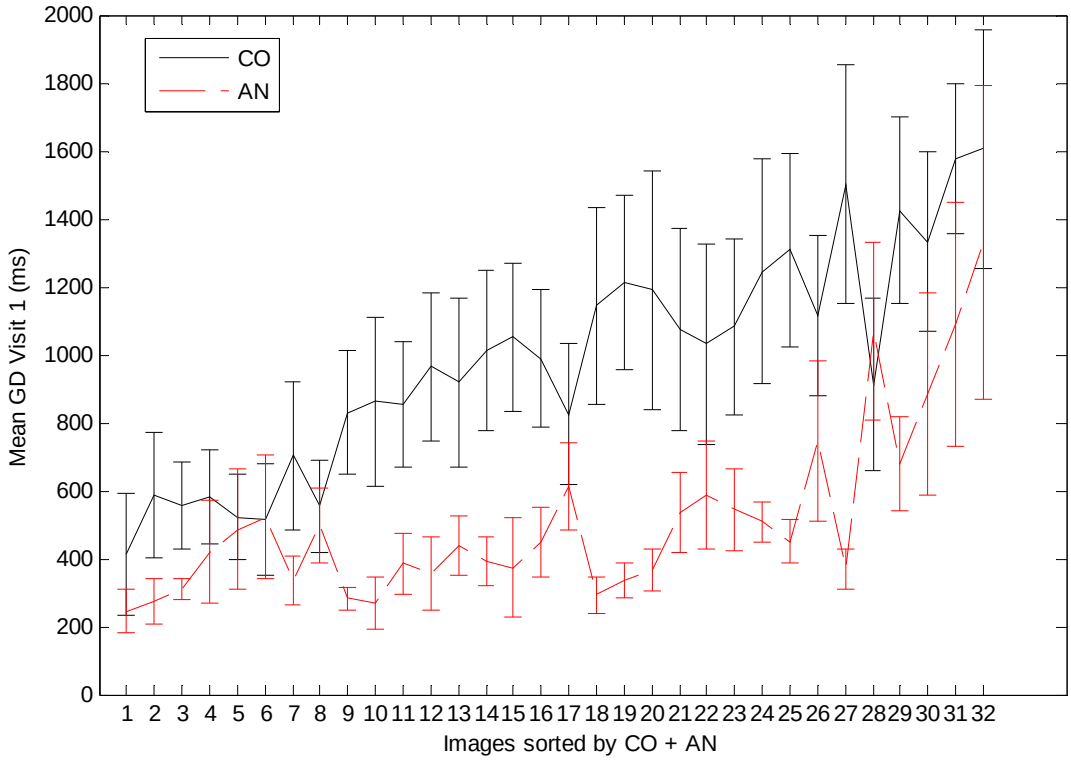


Figure 2.1: Mean Glance Duration on Visit 1 (a measure of attention, See Glance Duration By Visit Below) by images sorted by total mean time spent by both Anorexic and Control populations. Region of high attraction for both Anorexic and Control (28-32) and region of low attraction for both Anorexic and Control (1-8) provide little differentiation between the two populations. Images in the region (9-25) provide best differentiation between the two populations.

In the presentation creation process, images are collected and organized into slides for simultaneous presentation. In general, three types of images are collected: psychiatric congruent, control pleasant, and neutral. Psychiatric congruent stimuli are selected based on previous attention bias studies (i.e. dot-probe, Stroop, visual scanning). These stimuli are thought to evoke attention bias from patients with specific psychiatric disorder relative to control viewers. Control-pleasant stimuli are images of pleasant social interaction that have been shown to be

preferred by control populations (Eizenman et al., 2003). Neutral stimuli, objects with low emotional content, provide a baseline for the minimum attention an image receives on a slide. Assembling a slide requires a hypothesis on how psychiatric and control viewers will allocate their visual attention to the images on the slide. For example, body shape concern in Anorexia Nervosa manifests in a need to evaluate body shape stimuli in a manner that is different from non-anorexic viewers (Jansen et. al., 2005). Therefore, it is hypothesized that placing body shape stimuli on a slide together with control pleasant stimuli will cause anorexic viewers to allocate their visual attention to body shape stimuli while controls will pay more attention to control-pleasant images. In this example, control-pleasant images are selected such that they do not contain body shapes and the amount of details, contrast and color in the two image-types should be similar. Repeated measurements of visual scanning behaviour (each parameter of VSB is a random variable) are required to gain statistical confidence in identifying disorder-congruent behaviour. Repeated measurements are obtained by presenting multiple slides with the same configuration of image types. Multiple hypotheses can be tested by presenting multiple sets of slides with different configurations of image types. A more detailed description of image selection for specific studies is provided in Chapters 3, 4 and 5.

For each set of images for a specific study (presentation), a database of image properties is created to organize the information for analysis. This database consists of the following mandatory fields: Slide Number, Slide Type, Unique Image ID, Image X Coordinate, Image Y Coordinate, Image Width, Image Height, and Image Type. It may also contain any number of additional fields, for example: Image Description, Orientation, valence, and arousal (ratings of the image by control groups). A slide generator program takes this database of image properties along with the collected images and produces the slides and a slides property file necessary for presentation. The slide generator program has several options that allow the user to define image positions on the display monitor.

Figure 2.2: Example of a slide created for the Major Depressive Disorder study. On this slide, two dysphoric (top-right, bottom-left) and two control-pleasant (top-left, bottom-right) are shown to the viewer simultaneously.

2.3. Presentation and Data Recording

The visual stimuli (slides) are presented to participants on a 19" computer monitor that is part of the Visual Attention Scanning Technology (VAST, EL-MAR. Inc., Toronto, Ontario, Canada) binocular gaze estimation system (Guestrin & Eizenman, 2006). A second monitor displays the subject's fixation points on each image and provides the operator real-time estimation of gaze position. During presentation, participants can move their heads freely within 1 cubic foot, which supports natural viewing of the visual stimuli. Participants sit at a distance of approximately 65 centimeters from the monitor so that the visual angle subtended by each of the four images on each slide is approximately (15.2º x 11.4º). The horizontal and vertical separation between any two images is greater than 2.5º. Following a 9 points calibration procedure in which the participant is asked to follow a moving target on the computer screen, the participant views the presentation of slides which appear sequentially. The duration that each slide is presented and the time interval between slides can be set for each experiment. Figure 2.3 is a picture of the monitor that a participant sees and Figure 2.4 is a screen caption of the eye tracker operator's monitor.

VAST generates a single output file every time a set of slides are presented. This output file includes experiment, subject, and slide properties along with the gaze estimation. Experiment properties include the Study Name, Date and Time, and Operator Name and Comments. Subject properties include Unique ID, Type (Control, Patient, etc.), Age, Gender, and other optional attributes such as Weight, Height, and Condition. Slide properties are the information in the file that was generated during the creation of the presentation (See previous section). In the output file, each gaze estimation by the eye tracker is time stamped and provided as x and y locations on the screen in pixels for both eyes along with the pupil area in mm^2. Also, the visual stimulus that was presented at the time of each gaze-estimation is recorded.

The presentation, recording, and analysis of visual scanning parameters (see following section) are automated and organized into databases by Microsoft Windows Applications written in C++. These automated algorithms allowed the collection and analysis of the large amounts of data required to study visual scanning behaviours.

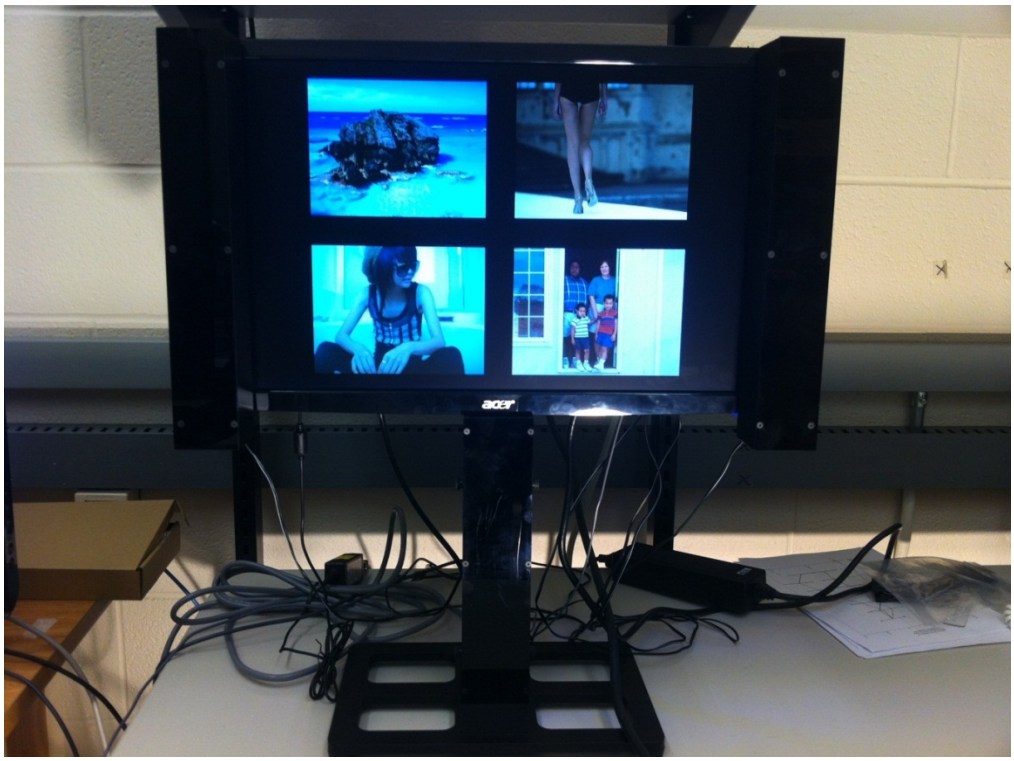


Figure 2.3: Picture of participant's monitor. Camera and infrared light sources are hidden behind infrared filters on the bottom and sides of the monitor.

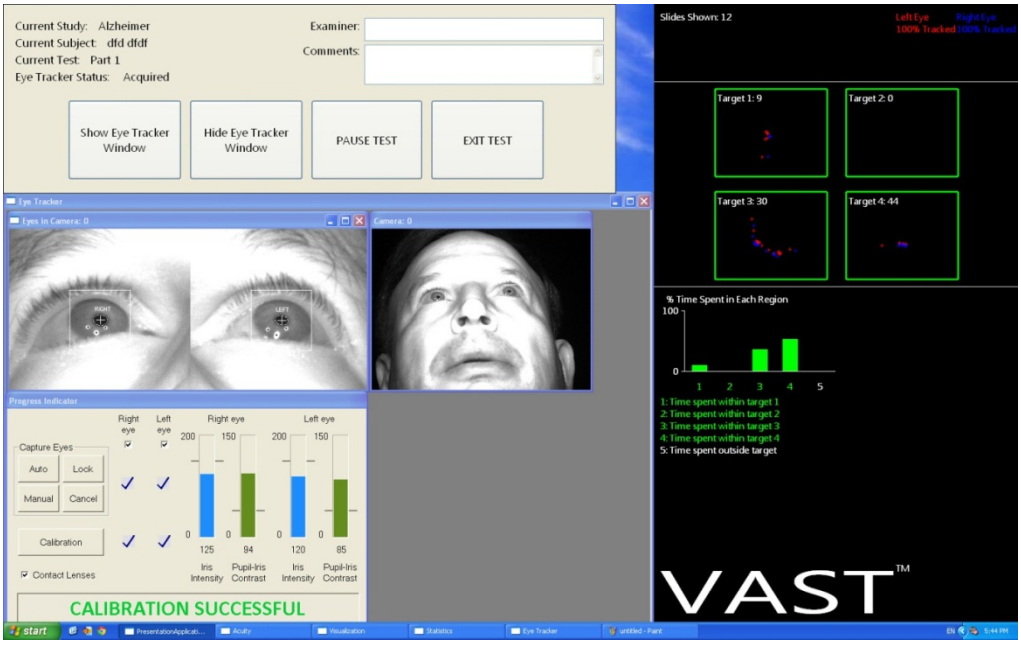


Figure 2.4: Screen caption of eye-tracker operator's monitor. Estimated gaze positions are shown real-time in the upper-right hand corner of the screen. Total time spent on each of the four images on the current slide is presented in the bottom-right hand corner of the screen. The image from the camera, metrics of eye-tracking quality, and eye-tracker controls are displayed in the bottom-left.

2.4. Analysis

When represented as a sequence of gaze positions, visual scanning behaviours are high dimensional data generated by a complex network of interactions. To analyze the many aspects of visual scanning behaviours that may pertain to psychiatric disorders a set of visual scanning parameters was developed. Visual scanning parameters vary in complexity and may be organized into three categories according to the manner that the raw gaze estimation data is preprocessed. Sections 2.4.1. and 2.4.2. describe parameters similar to those currently found in visual scan path research (Spivey et al. 2000; Eizenman et al., 2003). Section 2.4.3. describes a novel method of organizing the time-sequence nature of VSBs to assess its temporal characteristics.

2.4.1. Parameters derived directly from Gaze Positions (Raw Data)

The following parameters are derived directly from the raw sequence of gaze positions recorded by VAST.

Relative Time on Image (RTOI)

A simple method of estimating bias in scanning behaviour is to compare time spent on one image of a slide to all other images on the slide. Relative Time on Image (RTOI) is determined by dividing the total number of eye estimates on each image by the total number of valid estimates on all images (Equation 2.1)(To account for noise in the eye-tracker's gaze estimates the boundaries of each image are extended by 1 degree for these calculations). RTOI is dependent on all the images on the slide and the sum of RTOI for all images on a slide is 1. If all four images on a slide attract equal amount of attention, the expected value of RTOI on a slide is 0.25.

$$RTOI_i = \frac{TOI_i}{TTOI}$$

Equation 2.1: Calculation of RTOI for image *i* on a slide where TOI_i is the sum of the total number of valid estimates on image *i* and TTOI is the total number of valid estimates for the entire slide.

Average Pupil Area

Pupil area (size) is controlled, in part, by the sympathetic nervous system which is affected by emotions. Therefore, pupil size may provide a measure of the emotional state of the viewer. However, pupil size is also affected by luminance which is uncontrolled in the paradigm developed by Eizenman et al., 2003. Therefore, pupil size was not used in the analysis described in this thesis. The average pupil-size is the mean of all pupil-sizes for each eye when the same type of images is viewed by a participant. The typical average pupil width ranges from 21 to 34 mm^2.

2.4.2. Parameters derived from Fixations and Saccades

VAST records visual scanning behaviour as a sequence of eye positions at fixed sampling intervals. A more compact physiologically meaningful representation of this data can be obtained by describing the VSB as a sequence of fixations or saccades. Fixations are periods of time (>165 ms) where the eye is relatively still. Saccades are rapid ballistic movements of the eye that move the point of gaze from one location to another. During saccades visual processing is suppressed. A description of the visual scan path as a sequence of fixations and saccades allows for more accurate determination of the time periods in which visual information is processed (fixations). To determine fixations and saccades sequences, gaze positions are clustered temporally and spatially by clustering algorithms that were developed initially by D.

Cassel and M. Eizenman (2009). The algorithms were refined to support the estimation of the following visual scanning parameters: Number of Visits, Average Glance Duration, Relative Fixation Time, Fixation Frequency Within Image, Fixation Order and Time to First Fixation.

2.4.2.1. Fixation Estimation

The clustering algorithm considers each estimate of gaze position from the eye-tracking system and forms a fixation cluster if it finds a cluster of gaze estimates that are each within a "maximum distance" (e.g., 1 degree) of at least one other gaze estimate in the cluster and if the time interval over which the points in the cluster were recorded is longer than the minimum "fixation time" (i.e., longer than 165 ms). The parameters that define the "maximum distance" and the "fixation time" are based on the eye-tracker noise characteristics and the physiological characteristics of human fixation eye movements. When a new cluster is formed (i.e. after a saccade), the time interval from the beginning to the end of the last cluster is examined and gaze estimates that were not included in the last cluster (i.e. gaze estimates that are more than "maximum distance" from all the points within the fixation cluster) are added to the fixation cluster if the time interval associated with consecutive estimates that are not included in the cluster is less than 33 ms. This parameter is referred to as the "maximum time gap allowance". To process eye-blinks, the algorithm allows for a condition where there are no gaze position estimates from the eye-tracker for a time interval that is less than 500 msec (the "maximum no estimates period"). Following an eye-blink, subsequent gaze positions are compared to the current fixation cluster. If the period with no gaze-position estimates is longer than the "maximum no estimate period" subsequent gaze estimates are considered for a new fixation cluster.

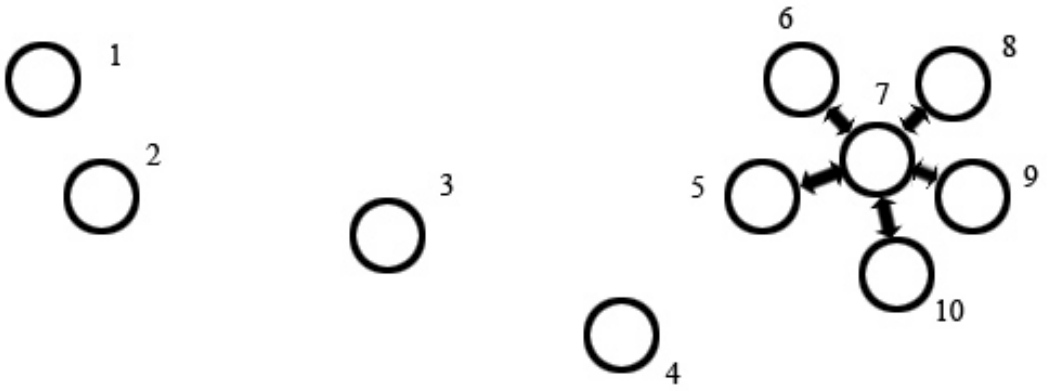


Figure 2.5 illustrates how the clustering algorithm processes data.

The numbers in Figure 2.5 indicate the temporal sequence of the gaze estimates. Estimate 1-4 are from a saccade, while 5 to 10 represent a fixation. The clustering algorithm classifies it as a fixation cluster based on the fact that estimate 7 is sufficiently close (maximum fixation distance) to enough other points (minimum time gap allowance) to constitute a fixation.

2.4.2.2. Parameters

Number of Visits

Number of Visits is the total number of times that each image was visited. A visit occurs when there is a fixation within the boundaries of the image and the previous fixation was outside the boundaries of that image. Figure 2.6 illustrates the definition of a visit. The number of visits to all images on a slide is subject-dependent. Figure 2.7 shows that the average number of visits to all images on a slide varies widely for both anorexic ($Range_{anorexic}$ = [6.2,11.3]) and non-anorexic control subjects ($Range_{Control}$ = [4.8, 11.5]) and it is clear that this visual scanning parameter is not likely to differentiate control subjects from patients with Anorexia Nervosa. The Number of Visits to each image can be normalized by dividing by the total number of visits to all images on the slide (Equation 2.2). This normalization preserves the differences in Number of Visits between competing images while making it possible to compare the subjects with different scanning rates. Number of Visits is used (Chapter 3) to characterize bias in visual scanning in Anorexia Nervosa.

$$Normalized\ Number\ of\ Visits_i = \frac{Number\ of\ Visits_i}{\sum_{j=1}^{n} Number\ of\ Visits_j}$$

Equation 2.2: Calculation of Normalized Number of Visits for image *i* where *n* is the total number of images on the slide.

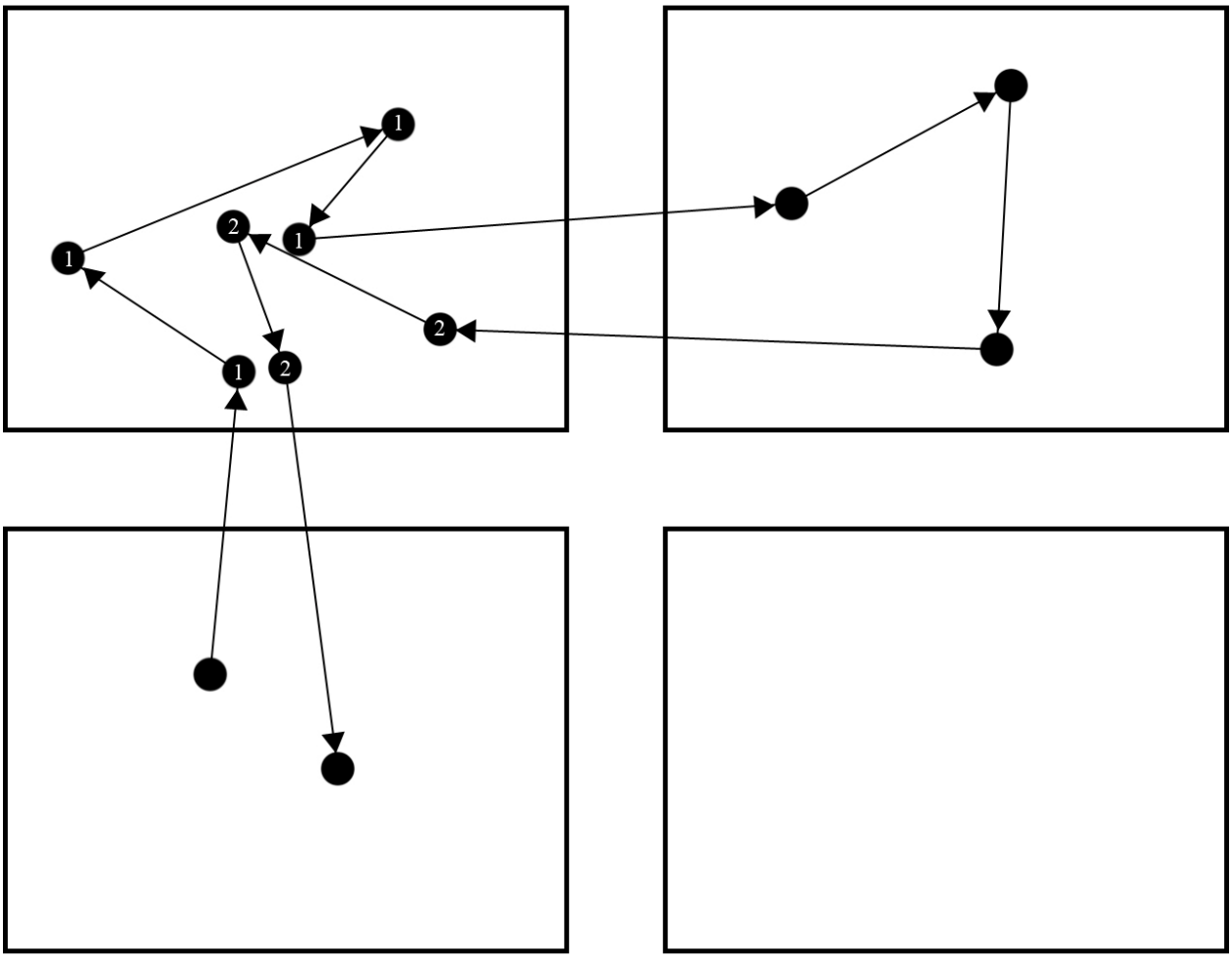


Figure 2.6: Diagram depicting visits to the top-left image. Black dots represent fixations and arrows between dots represent saccades. Visit 1 and Visit 2 marked for the top-left image.

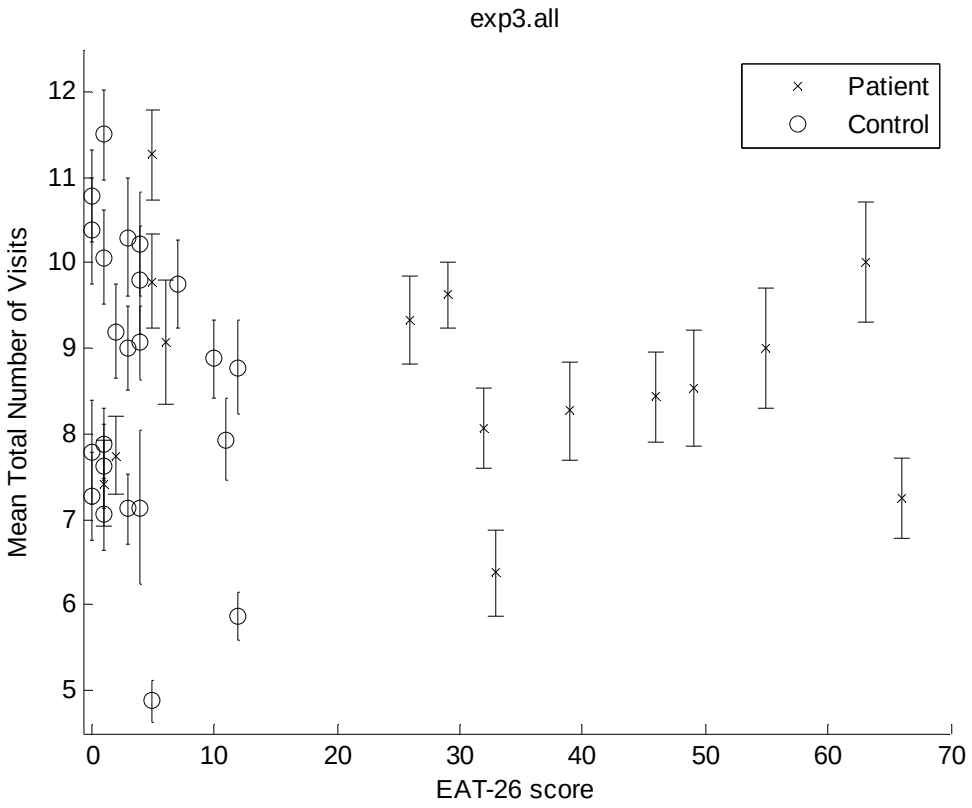


Figure 2.7: Mean Number of Visits to all images on a slide for anorexic and age-matched control group. EAT-26 score is a self-assessment questionnaire on eating attitudes and behaviour. A higher score means more disordered eating.

Average Glance Duration

Average Glance Duration is the average duration in milliseconds spent on each image on a slide. It is the total time of all fixations within the boundaries of an image divided by the Number of Visits (Equation 2.3). Difficulty disengaging from the image will result in a longer Average Glance Duration while avoidance (quick disengagement from the image) will result in a shorter Average Glance Duration. The study in Chapter 3 provides an example of bias in Average Glance Duration in patients with Anorexia Nervosa.

$$Average\ Glance\ Duration_i = \frac{TOI_i}{Number\ of\ Visits_i}$$

Equation 2.3: Calculation of Average Glance Duration on image *i* where TOI_i is the total time on image $_i$ and $Number\ of\ Visits_i$ is the total number of visits to image *i*.

Relative Fixation Time

The Relative Fixation Time of an image is equal to the total time of all fixations on the image divided by the total time of all fixations on the slide (Equation 2.4). Relative Fixation Time is

highly correlated with the Relative Time on Image. Figure 2.8 is a histogram of the differences in total time on images and the total fixation time on all the images on a slide (i.e., total time minus saccadic eye movements). The Figure shows that the mean duration of saccadic sequences is 1.22s (Std=0.71s) which is approximately 12% of the time that each slide is presented. However, Relative Fixation Time and Relative Time on Image are very similar. Figure 2.9 shows a histogram of the difference of Relative Fixation Time and Relative Time on Image for all images and subjects included in the Major Depressive Disorder study (Chapter 4) where slides are presented for 10.5 seconds. Ninety percent of images fall within ±0.038 (approximately 10.5s×0.038 = 0.4 s). In the small number of instances (1.2%) where large differences (>1s) between the two parameters occur, these differences are likely attributed to poor gaze-estimation. RFT is highly correlated with RTOI even though saccades remove approximately 1s of data suggesting that saccades are distributed evenly between images.

$$RFT_i = \frac{FT_i}{Total\ fixation\ time\ on\ Slide}$$

Equation 2.4: Calculation of Relative Fixation Time for image *i* where FT_i is the total fixation time on image *i*.

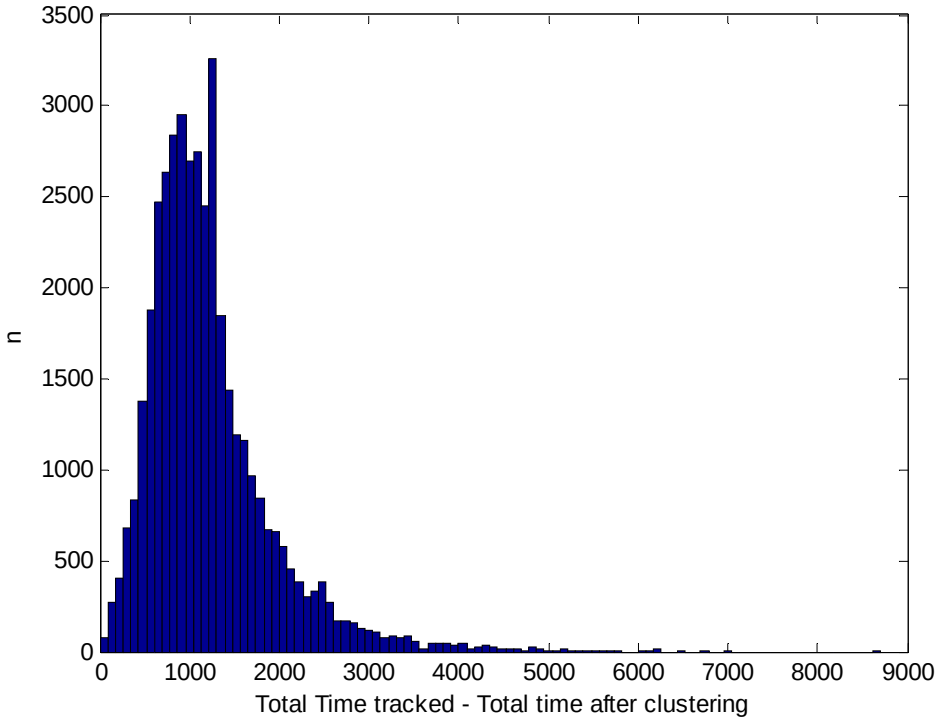


Figure 2.8: Histogram of the differences in total time and total time after fixation. Data taken from subjects over 8 visits of data recording. Slides where tracking was poor (<80% of time was not tracked by VAST) were removed and not included. All the values are positive because total frames tracked are always larger than total frames after clustering.

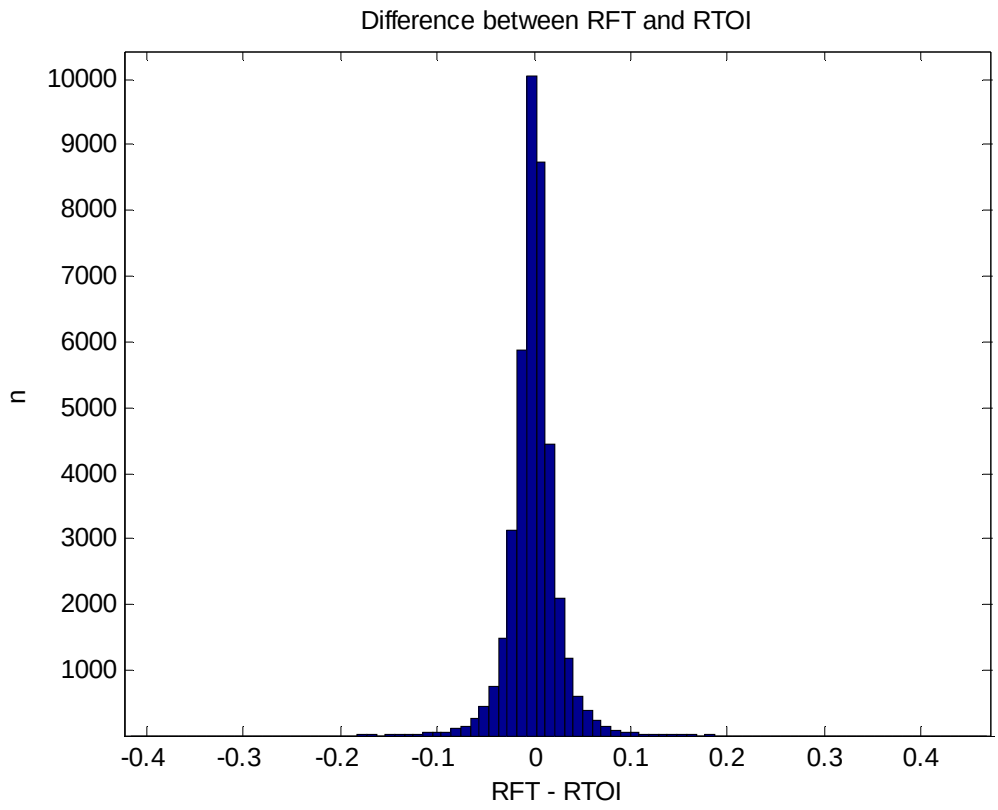


Figure 2.9: Histogram of difference in Relative Fixation Time and Relative Time on Image. Data taken from subjects over 8 visits of data recording. Slides where tracking was poor (<80% of time was not tracked by VAST) were removed and not included.

Number of Visits and Average Glance Duration are the components that make up Relative Fixation Time. This relationship is expressed in Equation 2.5. In some instances, biases in visual scanning behaviour are better understood by analyzing the components of the relative fixation time.

$$RFT_i = \frac{AGD_i \times NV_i}{\sum_{j=1}^{n} AGD_j \times NV_j}$$

Equation 2.5: Relationship between Relative Fixation Time, Average Glance Duration and Number of fixations for image i where n is the total number of images in the slide, AGD_i and NV_i are the Average Glance Duration and Number of Visits of image i respectively.

Number of Fixations on Image

The Number of Fixations on an Image indicates the subject's interest in exploring image details (Equation 2.6). Figure 2.10 illustrates that both patients with Anorexia Nervosa and age-matched controls exhibit large variations in the total number of fixations on a slide. Once again,

this effect can be normalized by dividing the Number of Fixations on Image by the total number of fixations on a slide (Equation 2.7). This normalization preserves the differences in the number of fixations between competing stimuli while making the parameter comparable between slides and subjects. The Normalized Number of Fixations is one component of the Relative Fixation Time which includes both the duration of each fixation and the number of fixations. In this way, Normalized Number of Fixations emphasizes the intent of the subject to gather information from an image. In the Anorexia Nervosa Study with 33 subjects, 82.8% of the variability between the two parameters (R =0.9109) can be explained by a linear relationship. Figure 2.11 plots A) Relative Fixation Time vs. Normalized Number of Fixations and B) the residuals of the linear regression between the two parameters. The linear relationship is clearly visible and the residuals are distributed evenly for the range of observations.

$$Number\ of\ Fixations_i = \sum_{j=1}^{n_i} 1$$

Equation 2.6: Number of Fixations on Image i where $j = 1 \ldots n_i$ are all fixations that land within the boundaries of image i.

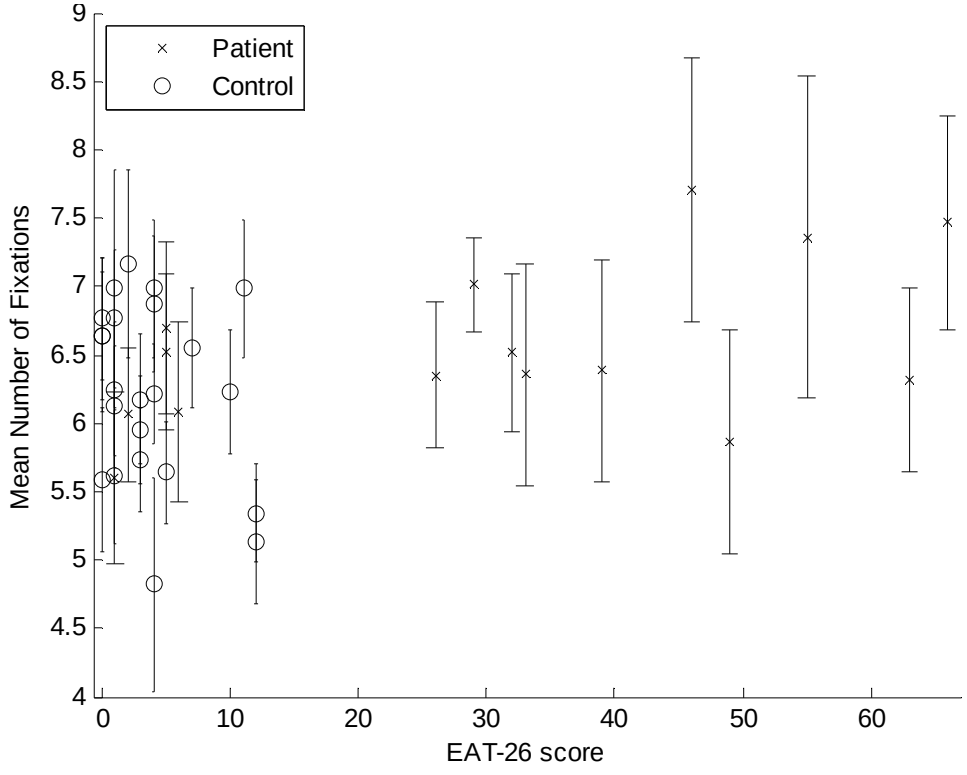


Figure 2.10: Mean Number of Fixations on Image varies between anorexic and non-anorexic subjects. EAT-26 score is a self-assessment questionnaire on eating attitudes and behaviour. A higher the score means more disordered eating.

$$Normalized\ Number\ of\ Fixations_i = \frac{\sum_{j=1}^{n_i} 1}{\sum_{k=1}^{n} 1}$$

Equation 2.7: Number of Fixations on Image i where $j = 1 \ldots n_i$ are all fixations that land within the boundaries of image i and $k = 1 \ldots n$ are all fixations for the entire slide.

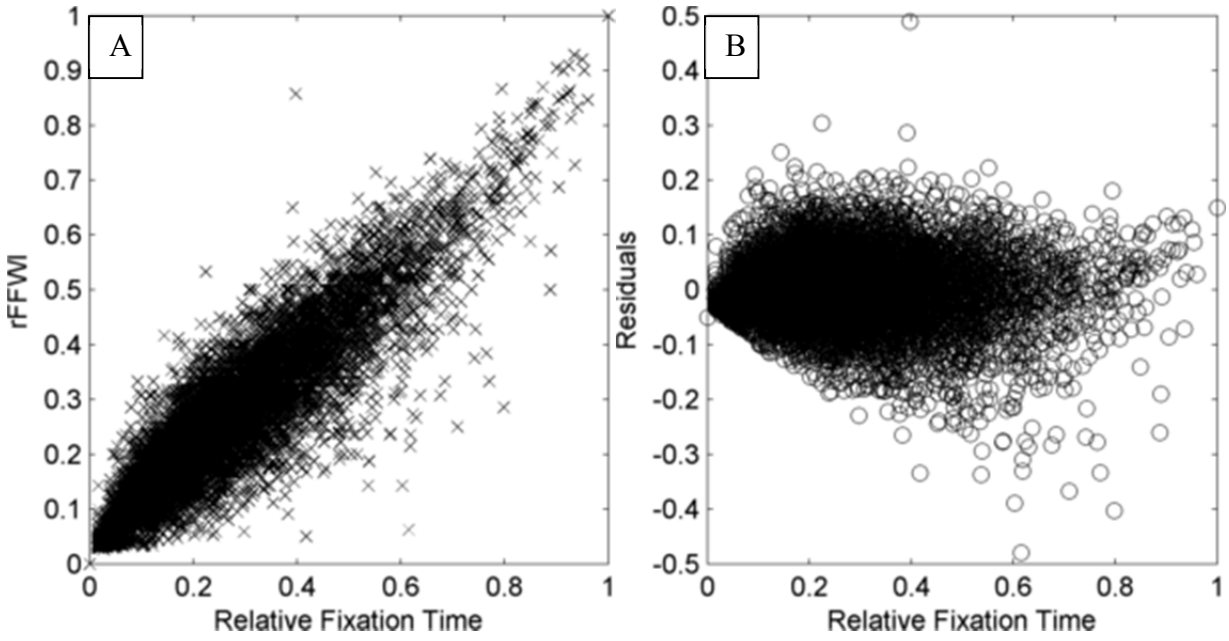


Figure 2.11: A) Relative Fixation Time vs. Normalized Number of Fixations. B) Residuals of linear regression between Normalized Number of Fixations on Image and Relative Fixation Time.

Order of First Visit (Fixation Order) and Time to First Fixation

When a slide is first presented the content of the four images on the screen is unknown to the viewers. The order in which subjects view the images is largely a function of where subjects were looking on the previous slides and the preferred scanning pattern of each subject. Figure 2.12A illustrates that participants in the Anorexia Nervosa study viewed the top-left image position first, slightly more (10%) than the other image positions (top-right, bottom-left, and bottom-right). In the studies this bias is mitigated by equally distributing the image types in each image position. For example, Figure 2.12B shows that body-shape images and control-pleasant images in the eating disorder study are visited first on a slide an equal number of times. Figure 2.13 shows the first visit transition probabilities matrix by image position for all participants of the Anorexia Nervosa study. Probabilities along the principal diagonal are 0 because it is impossible to transition to the image that is being viewed. Transition probabilities on the minor diagonal are low (<0.08) suggesting that diagonal saccades (i.e. saccades from top-left to bottom-

right, top-right to bottom-left, bottom-left to top-right, and bottom-right to top-left) are infrequent. Transitions along horizontal or vertical directions are equally as likely, with the exception that it is more likely to move from bottom-left to top-left (0.54) than it is to move from bottom-left to bottom-right (0.37). The order in which the images on the slide are visited can be correlated with subjects' interest in these images. For example, patients with Anorexia Nervosa have a significantly lower average order of first visit on thin body shape images (i.e. they tend to visit these images earlier) than control-pleasant images ($U = 182574$, $P<0.01$) while this bias also occurs in controls ($U = 417401$, $P<0.01$) in the Mann-Whitney U test.

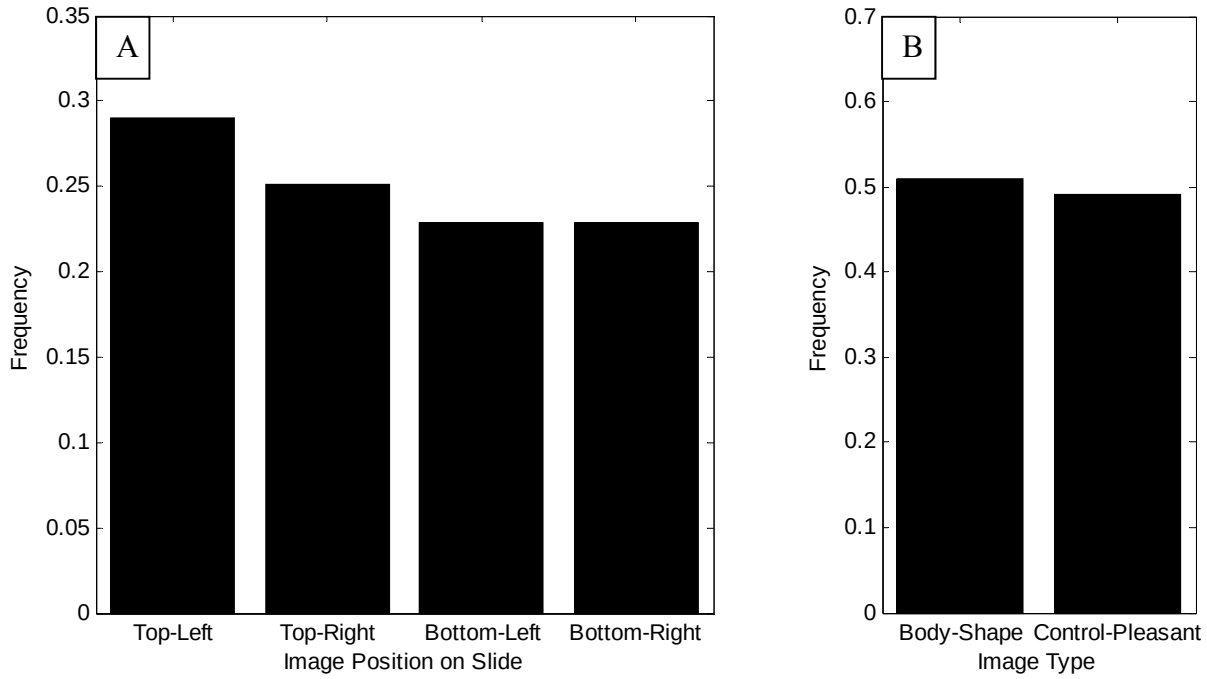


Figure 2.12: **A)** Number of images participants viewed first on slide by image position (n=2653). Images in the top-left position are visited first approximately 10% more than the image positions. **B)** Body-shape images and control-pleasant images are visited first an equal number of times on hypothesis-testing slides (n=555).

		End Image Position			
		Top-left	Top-right	Bottom-left	Bottom-right
Start Image Position	Top-left	0	0.5122	0.439	0.0488
	Top-right	0.450	0	0.050	0.500
	Bottom-left	0.5405	0.0811	0	0.3784
	Bottom-right	0.0313	0.500	0.4688	0

Figure 2.13 Transition probability matrix between image positions on a slide.

The Time for First Fixation is measured from the time that a new slide appears on the screen. Anxiety is often associated with lower Time to first fixation as subjects try to look at all the images on the slide as soon as they can to identify threats (Becker and Detweiler-Bedell, 2009). Figure 2.14 shows the Time for First Fixation for all subjects in the Anorexia Nervosa study.

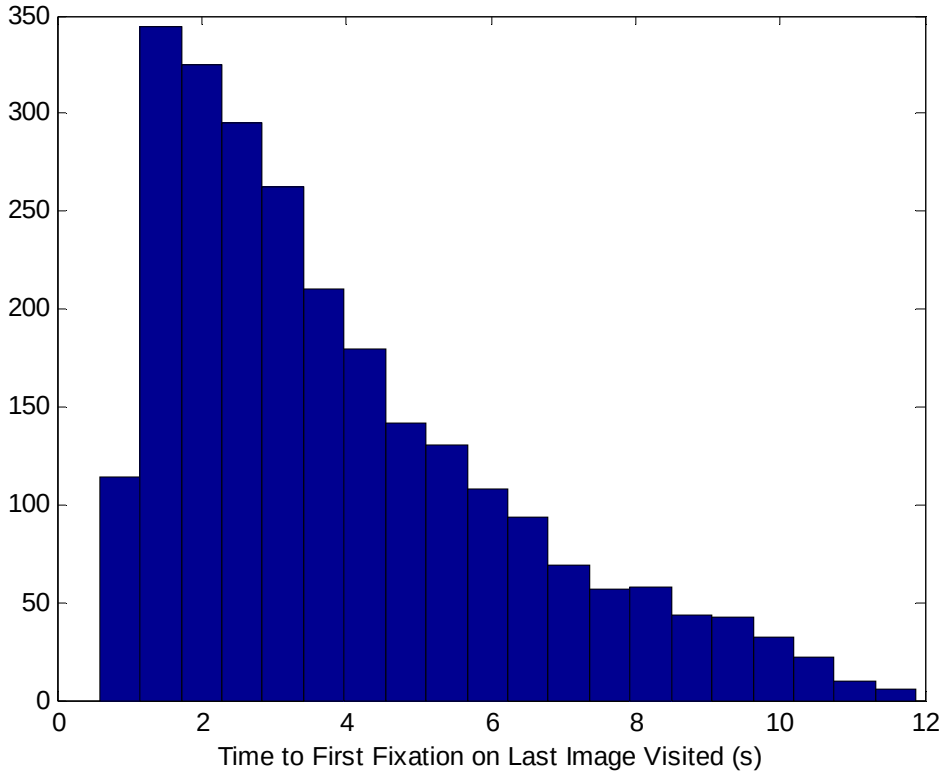


Figure 2.14: Histogram of Time to First Fixation of the last image visited for each slide. Data is taken from 33 subjects in the Anorexia Nervosa Project (see Chapter 3).

2.4.3. Parameters derived from Sequence of Visits

Many of the visual scanning parameters described previously, such as Average Glance Duration and Relative Fixation Time, characterize visual scanning behaviour by averaging the visual scanning behaviour over the presentation of an entire slide into a single parameter per image. However, it is evident from the Order and Time to First Visit parameters that scanning behaviour is dependent not only on how long images are viewed but also 'when' they are viewed. Though it is possible to analyze the visual scanning process by considering the time-sequence of fixations (i.e., not only the length of each fixation but also when it occurs) this is challenging because the large variability associated with these sequences (even for a single person). One method of retaining time-sequence information without increasing the dimensionality of the problem

significantly is through an intermediate representation of the data as a sequence of visits. Illustrated in Figure 2.15, sequences of fixations can be transformed into a sequence of visits for each image on a slide. Each visit to an image is described by four parameters: the visit order (I) total glance duration (II), number of fixations (III) and fixation durations (IV).

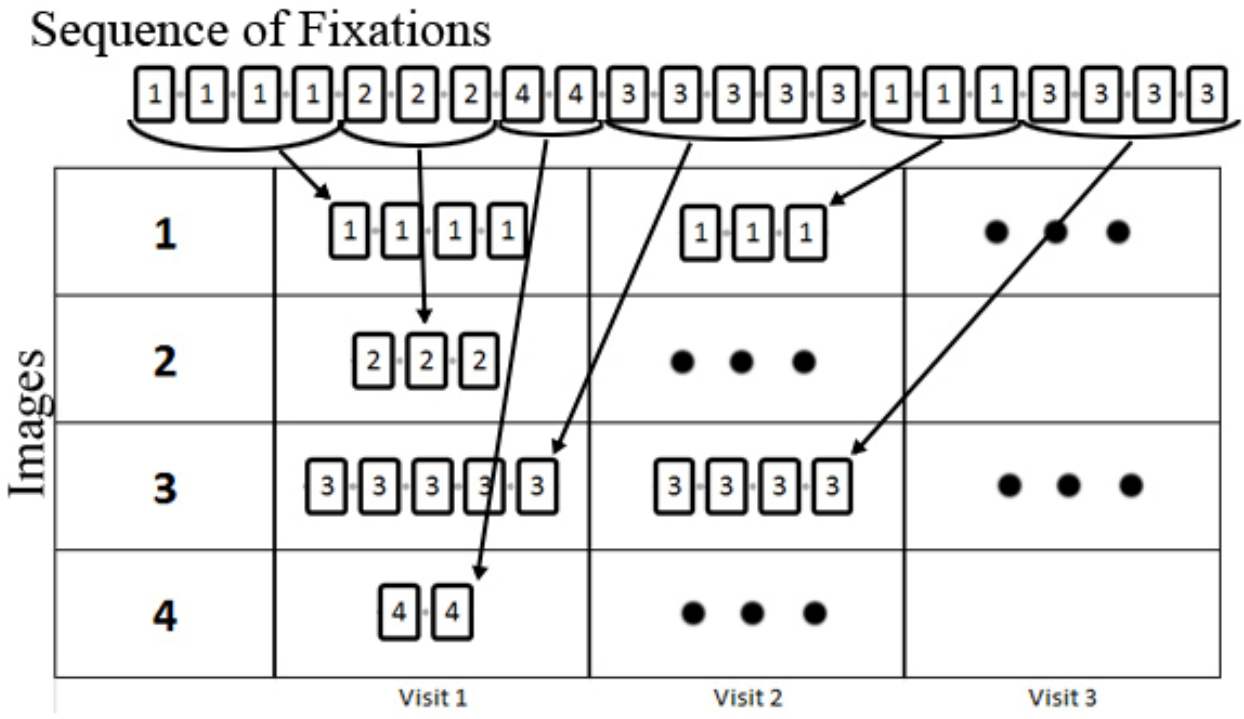


Figure 2.15: Sequence of fixations to sequence of visits per image.

Glance Duration by Visit

A Glance Duration by Visit is calculated as the time interval between the first gaze estimate of the first fixation on the image, during the visit, to the last gaze estimate of the last fixation before leaving the image (i.e. saccades are included in the calculation of this glance duration). Separating the durations by visit allows for a refined view of visual scanning behaviour. In Figure 2.16 and Figure 2.17, cumulative distribution functions (CDF) for Glance Durations by Visit are plotted for an anorexic and control subject respectively. These glance durations are recorded when the subjects view slides where all four images are control-pleasant ("filler slides) and slides with 2 thin body shapes images and 2 control pleasant images ("Thin body shapes" and "Control-Pleasant"). In Figure 2.16, there is a significant difference between the CDF of Glance duration visit 1 for images with "thin body shape" and the CDFs of the "control-

pleasant" and "filler" images. The differences between the CDFs in Glance durations visit 2 are much smaller. In Figure 2.17, the three CDFs overlap for the two visits. These results show that the anorexic viewer exhibited bias in visual scanning behaviour towards images with thin body shapes while the control subject did not. Further, this bias in scanning only manifests during the first time the patient visits the images. Thus, any visual scanning parameter which averages behaviour over the entire duration of the slide reduces the ability to detect transient differences in visual scanning behaviour. The CDF plots in Figures 2.16 and 2.17 can also be used to describe differences in the approach to scanning the images between these viewers. For example, in the anorexic patient (Figure 2.16) 90% of the observations on "filler" images are less than 2000ms in the first glance duration while only 40% are below 2000ms for the control subject (Figure 2.17) (i.e, the control subject views the images for longer durations on the first visit). In Chapter 5 Glance Duration by visit is used to develop a measure for apathy in the Alzheimer study.

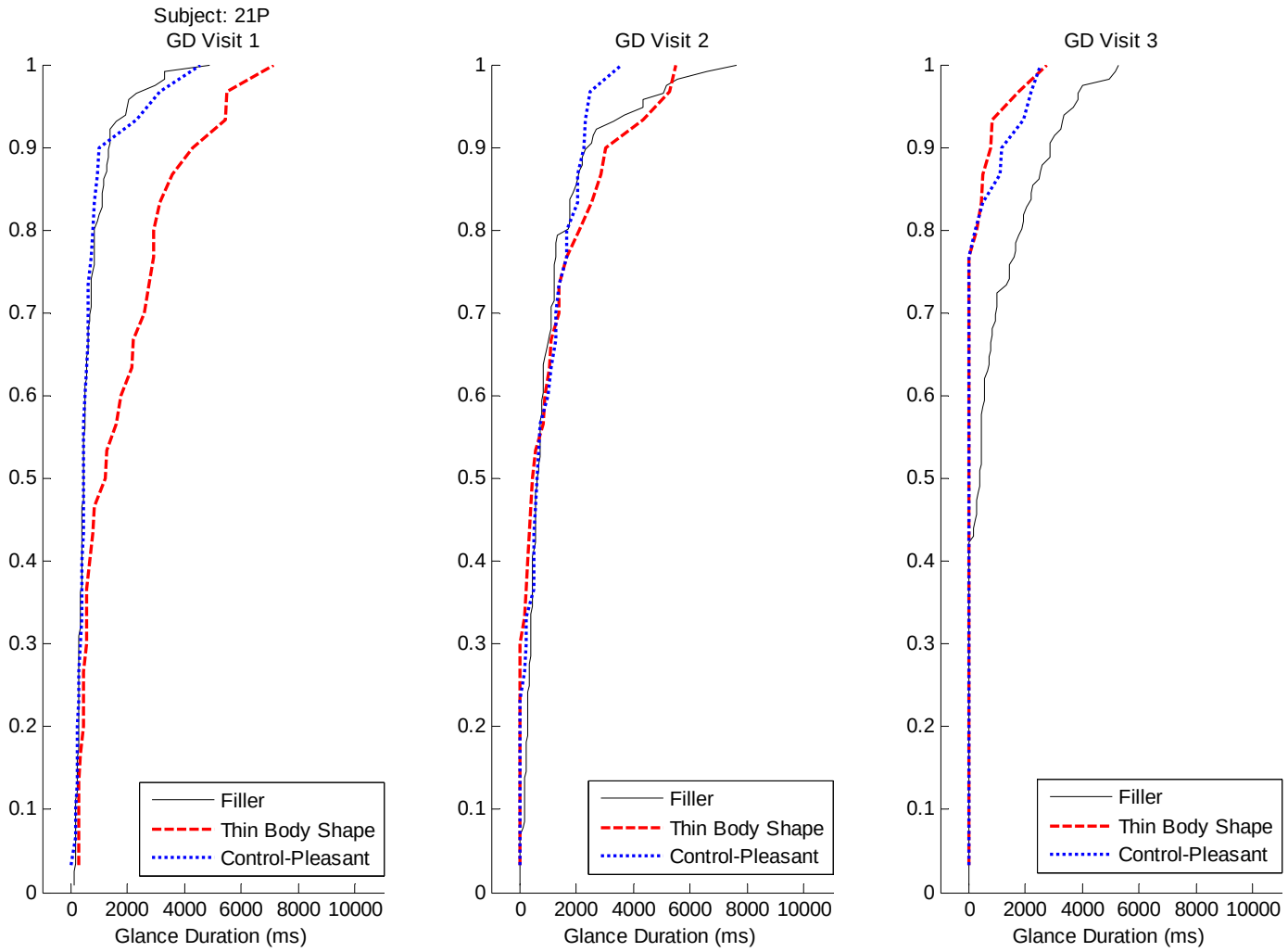


Figure 2.16: Cumulative distribution function of Glance Duration by Visit of an anorexic patient viewing a set of 16 slides with 2 thin body shape images (dashed line) and 2 control-pleasant images (dotted line) placed on the same slide along with a set of 20 filler slides (solid line) where all four images are control-pleasant. Differentiation between scanning thin body shape images and control-pleasant images may be observed in visit 1 while not present in GD visit 2 or 3.

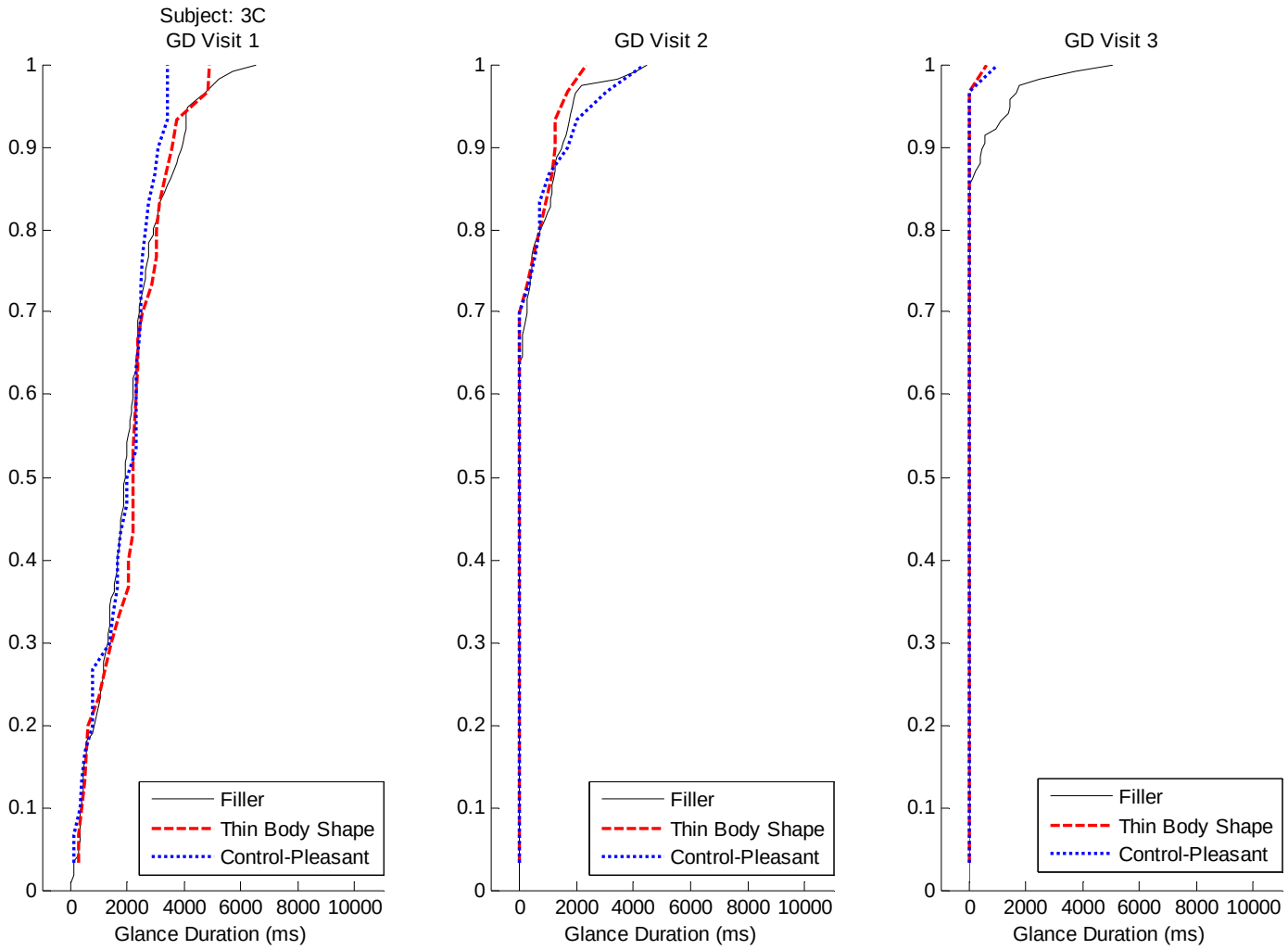


Figure 2.17: Cumulative distribution function of Glance Duration by Visit of a control subject viewing a set of 16 slides with 2 thin body shape images (dashed line) and 2 control-pleasant images (dotted line) placed on the same slide along with a set of 20 filler slides (solid line) where all four images are control-pleasant. Differentiation between scanning thin body shape images and control-pleasant images is not observed in any visit.

Fixation Durations by Visit and Number of Fixations by Visit

In the same way that Average Glance Duration and the Number of Visits are the components of the Relative Fixation Time, Fixation Durations and Number of Fixations by Visit are the components of the Glance Duration by Visit. In the dataset of the Major Depressive Disorder study (22 participants) 79.3% of the variability in Glance Duration by Visit can be explained as a linear transformation of the Number of Fixations by Visit (R=0.8903). As expected, these two parameters are highly correlated with slight differences caused by the length of fixation durations. These parameters of the visual scanning behaviour help to explain how biases manifest within a visit to an image.

2.5. Conclusions

A framework for studying VSB in psychiatric disorder is presented in this chapter. As the underlying cause of mental illness is different between psychiatric disorders, the manifestations of those underlying causes as attention biases are also likely to be different. Therefore, the presentation creation component of the framework is flexible enough to control which, where, and how visual stimuli are presented to viewers. The presentation, recording, and analysis of visual scanning parameters are automated to handle large amounts of data required to study VSB. Visual scanning parameters found in the literature that characterize VSBs are summarized and assessed. In addition, a novel set of visual scanning parameters that describe behaviour by visits is described. Chapters 3, 4, and 5, demonstrate how these visual scanning parameters may be analyzed to produce objective indicators for the detection of psychiatric disorders.

Chapter 3. Visual Scanning Behaviours in Anorexia Nervosa

3.1. Introduction

Anorexia nervosa (AN) is a severe and chronic disorder with one of the highest mortality risks of any psychiatric illness (Hannerz, Borga, & Borritz, 2001; Harris, Barraclough, 1998) and adolescents with AN have 18 times the risk of death (Steinhausen, 1998). The reduction in life expectancy ranges from 20-25 years (Steinhausen,1998; Harbottle et al., 2008). The disorder has an onset typically in adolescence and prevalence ranges from 0.3 to 1% (Hoek and vanHoeken, 2003). AN is characterized by food restriction leading to weight loss, an extreme fear of fat or weight gain despite being underweight, body image distortion, amenorrhea and a denial of the severity of the illness (DSM-IV-TR, 2000).

Cognitive behavioral theory for AN suggests that the core psychopathological process accounting for the persistence and severity of AN is a cognitive schema whereby patients judge themselves (evaluate self-worth) largely in terms of their eating habits, shape and weight and their ability to control them. Weight-control behavior, over evaluation and preoccupation with thoughts about eating, shape and weight stem directly from this core psychopathology. In some patients, the core psychopathological process is supported by mechanisms such as clinical perfectionism, low self-esteem and the "starvation syndrome" that help to maintain the disorder (Fairburn et. al., 2003). In clinical perfectionism, the fear of failure leads to frequent and selective attention to performance through repeated calorie-counting and frequent shape and weight checking (Shafran et. al., 2002). Self-criticism arising from negatively biased appraisal of these frequent checks, in turn, encourages even more determined effort to control eating shape and weight. Low self-esteem and reduced capacity to experience reward or pleasure (anhedonia) (Davies and Woodside, 2002) lead to negative cognition processing biases that tend to reaffirm the patients' overall negative view of themselves and to increase their determination to achieve in their valued domains (control over eating, shape and weight). The "starvation syndrome" appears to maintain AN through a process of pronounced social withdrawal that encourages self-absorption and isolation from external influences that might diminish their over-evaluation of eating, shape and weight and their control (Garner et al., 1997). The above maintaining

mechanisms in AN are associated with selective information-processing biases (attentional biases) towards food, body weight and body shapes (Long et al. 1994, Green et al. 1994, Overduin et al, 1995, Sackville et. al. 1998, Dobson et. Al., 2004, Johansson et al., 2005, Jansen et al., 2005, Shafran et.al., 2008, Shafran et.al. 2007, Lee and Shafran, 2008, Giel et al., 2011a, Giel et al., 2011b, Smeets et. al., 2011, Wietersheim et al., 2012).

In AN, selective attention (attention bias) to shape and weight images has largely been inferred through the use of the Stroop (12-14) and the Dot-probe test (Shafran et al., 2007). Meta analysis of the Stroop interference test for body weight and shape related words (Dobson and Dozios, 2004) suggested that when compared with controls, patients with AN demonstrate modest interference for weight and shape related words. Similar interference patterns were found with control words or food words in patients with AN, suggesting that the differences between AN patients and controls might not be simply or only associated with attentional biases to body weight and shape words (Dobson and Dozios, 2004). A subsequent meta-analysis of the Stroop interference test (Johansson et al, 2005) found modest interference for body weight and shape words with negative overtones (large body physique and weight) in patients with AN. Studies with the dot-probe paradigm (Shafran et al., 2007) showed that patients with eating disorders were significantly faster to respond to the probe when it appeared in the same location of a negative or a natural body shape stimuli but not for positive shape stimuli. In summary, data from the Stroop and Dot-probe studies suggest that patients with AN have modest attentional bias to negative body weight and shape related words and pictures.

A more direct method to estimate attentional biases is to measure visual scanning behavior (VSB), which provides continuous record of the attention allocation processes (Jansen et. al., 2005, George et. al., 2011, Giel et. al., 2011a, Wietersheim et al., 2012). VSB can be monitored without requests for meta-cognitive reports or other overt responses and it provides not only behavioural end products of cognitive processes but a continuous measure of attention (Hermans et al., 1999). Using visual scanning measures (i.e., gaze duration) Giel at al., (2011a) showed that patients with AN viewed food and control-pictures approximately the same amount of time, while healthy controls who had not eaten 1-8 hours before the tests looked longer at food pictures. George et al., (2011) showed that when estimating body size and attractiveness, control

subjects fixate mainly on the subject's stomach while patients with AN have wider fixation patterns that encompass other body features such as hip and collar bones. Jansen et al, (2005) showed that participants who were extremely dissatisfied with their weight or shape focused on the ugly parts of themselves and the most attractive parts of others while participants who were satisfied with their bodies concentrated on the self-identified ugly parts of others. In a subsequent study of patients with AN, Wietersheim et al. (2012) showed that these attentional biases are often small and vary greatly. Even though group attentional biases in AN were identified in all of the above studies, the large variability between patients and the overlap between the visual scanning behavior of patients and controls did not allow for robust detection (high sensitivity and specificity) of attentional biases in individual subjects with AN.

In this chapter attentional bias to positive and negative body shapes/weight images are studied in adolescent AN patients through the analysis of VSB (Chapter 2). The current literature has only involved subjects in early adulthood. As eating disorders most commonly have their onset in adolescence, subjects in their 20s would potentially already have a chronic form of the illness (Steinhausen, 2009; DSM-IV-TR). It is therefore important to assess an adolescent population to determine attention bias in the beginning stages of the disorder.

Based on the core psychopathology of AN and the fact that images of both thin and fat people are found on pro-anorexia webpages, the following three hypotheses were explored in three experiments:

1) Attentional biases for images with thin body forms (positive) will a) be present and b) be stronger in adolescent AN patients than in age-matched controls.
2) Attentional biases for images with fat body forms (negative) will a) be present and b) be stronger in adolescent AN patients than in age-matched controls.
3) In adolescent patients with AN, attentional biases for images with thin body forms will be stronger than attentional biases for images with fat body forms. i.e., in AN "thinspiration" is stronger than "reverse thinspiration".

Self-reported clinical assessment instruments such as the Eating Attitudes Test (EAT-26) (Garner et. al., 1982) are often used as screening tools but are limited in utility when participants

minimize or misrepresent their behavior. It is estimated that only 30% of the population of people with AN actually receive treatment (Preti, et al., 2009) and the lack of objective indicators for anorexia nervosa impact identification, diagnosis and the course of treatment of the disorder (Pinhas and Bondy, 2010).

To increase the sensitivity and specificity of detecting attentional biases, the novel method that we present analyses biases along more than one axis of the multi-dimensional attentional bias space. In this study, attentional biases are analyzed along two main axes. The first is attentional bias to images of thin body forms, which is associated with the core psychopathology of AN and with clinical perfectionism. The second is attentional bias to images with social interactions and pleasant scenery, which is associated with social withdrawal and anhedonia that help to maintain AN.

3.2. Study Design

3.2.1. Participants

Female participants between the ages of 12-18 were recruited from the Eating Disorder Program at The Hospital for Sick Children (HSC) and from an online advertisement on the HSC website. All participants completed the Eating Attitudes Test (EAT-26) (Garner, et al., 1982) on the day their visual scanning behaviors were recorded. The patient participant group had all been previously diagnosed with AN. Patient characteristics, including treatment duration, secondary diagnoses, and body weight were also collected. The never-AN control participant group consisted of individuals who never had any known eating disorder or any other mental illness. Control participants who scored above the clinical cut point of 20 on the EAT-26 were excluded.

A total of 13 patient participants and 20 control participants were recruited. The mean age of patient and control participants were 14.54 ± 1.6 and 14.40 ± 1.82, respectively, $t(31) = 0.22315$, $P = 0.41244$. The mean EAT-26 score for the patient and control groups were 38.0 ± 22.3 and 6.5 ± 5.6 respectively, $t(31) = 6.07$, $P = 5.01 \times 10\text{-}7$. The patient characteristics are summarized in Table 1. Four out of the thirteen AN participants had a sub-clinical EAT-26 score (<30). These participants have either intentionally or unintentionally self-reported behavior that is inconsistent with the other clinical data available and were labeled as “deniers” of pathological

behavior. The proportion of deniers within the participants (30.8%) is similar to the 27.6% found by Pryor, Johnson, Wiederman, and Boswell (1995) when denial was defined.

Participant	Total EAT Score	Age	Percent of Healthy Weight	Diagnosis
7A	**54**	16	83.6%	AN-BP
8A	**11**	16	99.1%	AN
9A	**47**	16	100%	AN
11A	**59**	14	100%	AN
15A	**14**	15	89.6%	AN
17A	**63**	15	91.9%	AN-BP
18A	**68**	14	90.4%	AN
21A	**6**	15	92.0%	AN
22A	**5**	16	83.3%	AN
25A	**45**	12	73.2%	AN
26A	**32**	16	96.0%	AN-BP
39A	**51**	12	75.6%	AN
47A	**39**	12	96.6%	AN

Table 3.1: Patient characteristics for the 13 anorexic patients recruited.

3.2.2. Visual Stimuli

Visual stimuli were organized into slides and were presented on a computer monitor. Each slide had four images from four different categories. To test the first hypothesis (Experiment 1), slides had images of thin body parts (eg. rib cage or hip bone visible); b) thin subjects with faces; c) images with social interactions; and d) high valence images with pleasant scenery. Two different categories for thin human body forms were used, as thin body parts without heads are processed differently from thin body parts with heads (Bauser et.al., 2011) and there is evidence that distinct cortical regions in humans respond selectively to human body and human faces (Downing, 2001). Images for the first two categories were selected from non-copyrighted pictures available on the internet, including those found on "pro-ana" websites (those that promote anorexia as a lifestyle rather than a mental illness) that contain images meant to inspire readers to continue in their restrictive eating patterns. These images have thus already been "self-selected" by people with anorexia symptoms as meaningful. Images with social interactions and

pleasant scenery (valence >6) were selected from the International Affective Picturing System (IAPS) database of images (Lang, Bradley, and Cuthbert, 2008). Images of celebrities, morbid thinness, and people inappropriately dressed were excluded. To test the second hypothesis (Experiment 2) slides have images of a) fat body parts (eg. "love handles, cellulite or obviously rounded abdomens), b) fat subjects c) social interactions and d) high valence images with pleasant scenery. To test the third hypothesis slides have images of a) thin body forms, b) fat body forms, c) social interactions and d) neutral objects. Images for the second and third experiments were selected in a similar manner to that described for the first experiment.

Each subject was presented with 78 slides. Each slide was presented for 12 seconds for a total presentation time of 15.6 minutes. The slides included 48 test slides (16 slides in each experiment) and 30 filler slides. Filler slides had images from the same theme and were used to mask the purpose of the experiment (only test slides were analyzed). The first five filler slides were used to familiarize the subjects with the testing procedure. The 48 test slides were randomly distributed among the filler slides. The four images on each slide were arranged in a 2 x 2 configuration with the specific spatial position of images from each category inter-mixed (i.e., each category of stimuli appeared in each quadrant of the slide the same number of times).

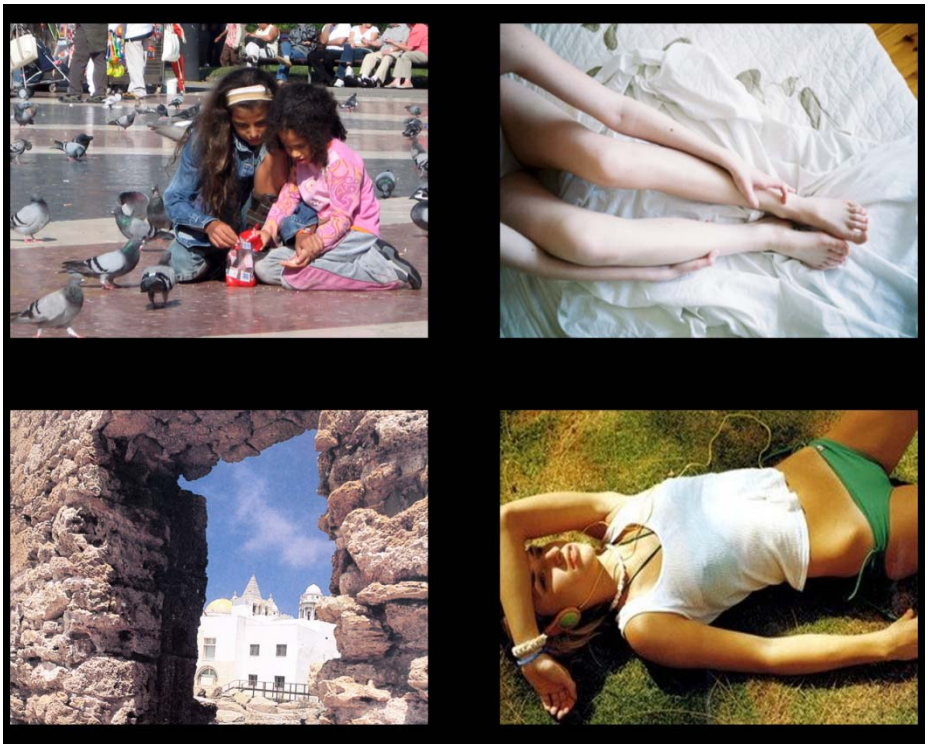


Figure 3.1: Example Experiment 1 test slide. Two images of thin body subjects, one with the face visible and one without the face visible, one image of social interaction, and one image of pleasant scenery are presented simultaneously.

Figure 3.2: Example Experiment 2 test slide. Two images of fat body subjects, one with the face visible and one without the face visible, one image of social interaction, and one image of pleasant scenery are presented simultaneously.

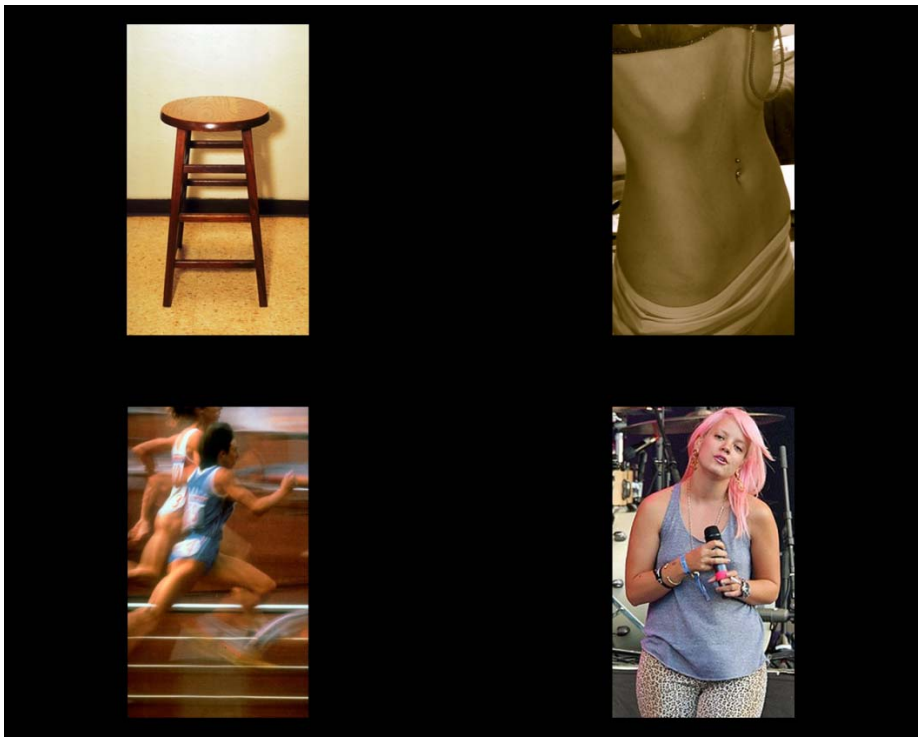


Figure 3.3: Example Experiment 3 test slide. One images of a thin body subject, one image of a fat body subject, one image of social interaction, and one image of pleasant scenery are presented simultaneously. Throughout the 16 test slides, thin and fat body images have equal number and distribution of ones with faces visible and faces hidden.

3.2.3. Procedure

On the day of the test, a research assistant explained the testing procedure to each participant and a consent form was signed (Hospital for Sick Children Toronto, REB 1000018863). All participants were told that the study is to measure the pupillary reaction of adolescent girls when viewing pictures. Visual scanning patterns and pupil-sizes were recorded and analyzed by EL-MAR's Visual Attention Scanning Technology (VAST, EL-MAR Inc., Toronto, Ontario, Canada) as described in Chapter 2. The horizontal and vertical separation between any two images was greater than 2.5º. Following a 9 points calibration procedure in which the subject was asked to follow a moving target on the computer screen the subject looked at the 78 slides. At the end of the visual scanning procedure participants completed the EAT-26 questionnaire (Garner et al., 1982).

The Relative Fixation Time, Relative Fixation Frequency, and Average Glance Duration were analyzed using the automated algorithms described in Chapter 2. For each subject, these visual scanning parameters for each theme of images (social, thin/fat body shapes), in all the test slides

of the same experiment, are averaged to provide the average relative fixation time and relative fixation frequency for each theme of images for each subject.

Factors such as the subject looking away from the screen, moving their head outside of the estimation range, or disruption by eyelashes or hair inhibited proper gaze estimation on occasion. Slides were removed and excluded from analysis where the number of valid estimates was less than 80% of the presentation time (9.6s). Seventeen slides were removed from the AN participants (8.2%) with a maximum of 8 slides removed for a single participant. From the control participants, 26 slides were removed (8.1%) with a maximum of 8 slides removed for a single participant.

3.3. Attention Biases in the Adolescent Anorexia Nervosa Population

The mean Relative Fixation Time (RFT) on all image types across the three experiments of a representative AN patient and a control subject are shown in Figure 3.4. On average, the AN subject spends more time on thin body shape images than the social interaction and high valence (scenery) images while the control subject spends more time on social images than thin images in Experiment 1. These results provide evidence for hypothesis 1; that attention bias towards thin images is present in AN patients and are greater than in control subjects. Similarly, this AN subject spends more time on fat body shape images than the competing social and high valence images in Experiment 2 while the control subject spent more time on the social images (Hypothesis 2) The biases observed in Experiment 1 and Experiment 2 are also found in Experiment 3 where the configuration of the images are different. Though thin and fat images are competing for visual attention on the same slide, this AN patient spends more time on both thin and fat images relative to the social image. In contrast, the control subject spends more time on social images than both thin and fat images. Little time is spent on neutral images in both subjects. Neutral images have less than other images. Therefore, it may be seen as a baseline for the amount of time a person will spend on any image presented on a slide. In Experiment 1, the difference in RFT between the body shape images and non-body-shape images is slightly larger than in Experiment 2. In Experiment 3, the AN patient's bias for thin body shape images over fat body shape images is clearly evident in the large difference in mean RFT between the two image types. Interestingly, the AN patient appears to view body shape images without a face

slightly more than images with a face while the opposite is true for the control subject. This result suggests a different mechanism of scanning between the two subjects.

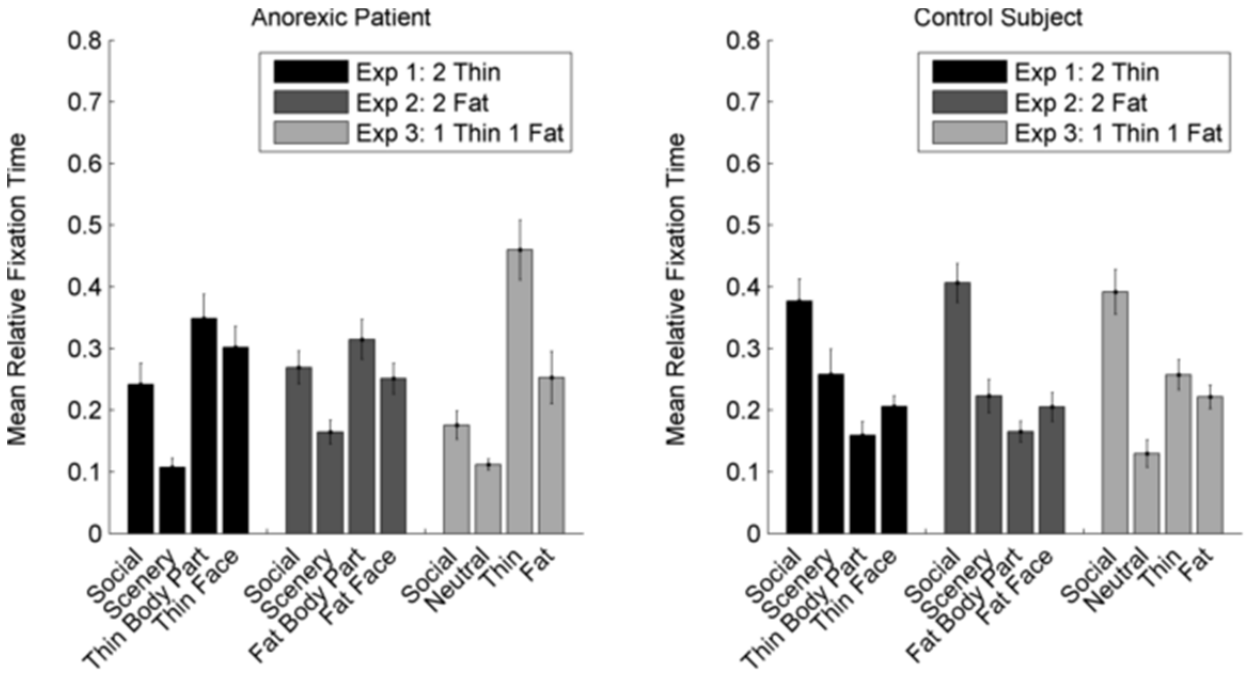


Figure 3.4: Mean Relative Fixation Time for a single Anorexic (left) and Control (right) subject on all image types of all three experiments. See 3.2.2 for a detailed description of images in Experiment 1, Experiment 2, and Experiment 3 test slides. Error bars show standard error of the mean.

A series of 2 (Group: control, AN) X 2 (Theme: social + pleasant scenery, sum of body shape images) repeated measures analysis of variance (ANOVA) were performed for Experiments 1 and 2 (thin and fat body shape images. For Experiment 3, a 2 (Group: control, AN) X 3 (Theme: social, thin body images, fat body images) repeated measures ANOVA was performed. Significant group interactions were found between Group X Image Type interactions for all three experiments ($P<0.01$). The results of the ANOVAs are shown in Tables 3.2, 3.3, and 3.4. Levene's test of equality of variances was used to verify that the variances of the means of the two populations (Anorexic, Control) and all image types (social, body shapes, high valence) were homoscedastic ($F=0.9467$, $df1=23$, $df2=372$, $P = 0.5352$).

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Group	0	1	0	0	1
Image Type	1.14302	1	1.14302	87.28	0
Group*Image Type	0.95348	1	0.95348	72.81	0
Error	0.81196	62	0.0131		
Total	2.54348	65			

Table 3.2: 2 (Group: control, AN) X 2 (Theme: social + pleasant scenery, sum of both thin body images). ANOVA table for Experiment 1.

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Group	0	1	0	0	1
Image Type	0.27981	1	0.27981	22.39	0
Group*Image Type	0.22767	1	0.22767	18.22	0.0001
Error	0.77491	62	0.0125		
Total	1.19417	65			

Table3.3: 2 (Group: control, AN) X 2 (Theme: social + pleasant scenery, sum of both fat body images). ANOVA table for Experiment 2.

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Group	0.00986	1	0.00986	1.52	0.221
Image Type	0.49494	2	0.24747	38.13	0
Group*Image Type	0.34708	2	0.17354	26.74	0
Error	0.60359	93	0.00649		
Total	1.31202	98			

Table3.4: 2 (Group: control, AN) X 3 (Theme: social, thin body images, fat body images). ANOVA table for Experiment 3.

The observations of the two representative subjects in Figure 3.4 are similar to the biases observed in the populations of 13 AN patients and 20 controls. Figure 3.5 illustrates the mean RFT for all AN patients (left) and all controls (right). The mean sum RFT of the two thin body shape images in Experiment 1 is significantly higher than the mean sum of the social and high valence images ($t(24) = 13.45$; $P < 0.01$) for the AN population but not for the control population ($t(38) = 0.5959$; $P = 0.5548$). In Experiment 2, the mean sum RFT on fat body shape images is significantly higher than the mean sum of non-body shape images in the AN population ($t(24) = 4.9178$; $P < 0.01$) but not in the control population ($t(38) = 0.4238$; $P = 0.6741$). Similar observations are found in Experiment 3. The mean RFT on thin and fat images are higher than the mean RFT on social images (Thin: $t(24) = 9.5025$; $P < 0.01$; Fat: $t(24) = 4.8403$; $P < 0.01$) for the AN population but not for the control population (Thin: $t(38) = 1.1792$; $P = 0.2456$; Fat: $t(38) = 0.9947$; $P = 0.3262$). The mean RFT of thin images for AN patients were significantly

higher than the mean RFT of fat images ($t(24) = 5.8414$; $P < 0.01$). This bias does not exists in controls ($t(38) = 0.1702$; $P = 0.8657$). Taken together, these observations of RFT in Experiments 1, 2, and 3 provide evidence for the three hypotheses stated at the beginning of the chapter. AN patients have a bias towards thin and fat body shape images that is stronger than in controls (hypothesis 1 and 2). Further, when thin and fat images directly compete for attention, the attention bias towards thin images is stronger than the attention bias towards fat images (hypothesis 3). Other interesting observations are that controls tend to view social interaction images for longer durations than scenery images (Exp1: $t(38) = 2.3246$, $P = 0.0255$; Exp2: $t(38) = 3.9814$, $P < 0.01$). Body shape images with faces have significantly higher mean RFT's than body shape images without faces for controls (Thin: $t(38) = 3.0750$, $P < 0.01$; Fat: $t(38) = 4.6344$, $P < 0.01$). These results may be because there are qualitatively more details in social interaction and body shape images with faces than pleasant scenery images and body shape images without faces which require subjects extra time to process. AN participants also have larger mean RFT's on social interaction images than pleasant scenery images (Exp1: $t(24) = 2.4130$, $P = 0.0238$; Exp2: $t(24) = 3.3409$, $P < 0.01$) but, in contrast, have similar mean RFT's between body shape images with and without faces (Thin: $t(24) = 1.3234$, $P = 0.1982$; Fat: $t(24) = 0.1187$, $P = 0.9065$). This difference suggests that AN participants view the two body shape stimuli similarly and this behaviour is different from the way controls view those images. In Experiment 3, social interaction and body shape images have similar mean RFT's in controls while AN patients have larger mean RFT's on body shape images at the expense of social interaction and neutral images.

Figures 3.6 and 3.7 show the Average Glance Duration (AGD) and Normalized Number of Visits (NNV) for all patients and all controls. In Figure 3.5, the relative magnitudes of AGD are quite similar to the RFT graph for all images types across all three experiments for both populations of subjects. This result suggests that the biases observed in the mean RFT parameter are a result of longer glance durations more so than additional visits to body shape images. Figure 3.6 shows that there are also some biases associated with the number of visits to competing image types. Controls appear to visit all images a similar number of times for all image types and in all three experiments with the exception of the neutral images in Experiment 3. A decrease in the number of visits to social and high valence images relative to thin body shape images is observed in

Experiment 1 for the AN population. In experiment 2, a decrease in the number of visits is only observed in high valence images and not social images relative to fat body shape images. This result is another indication that there is a stronger bias towards thin body shape images than fat body shape images in the AN population.

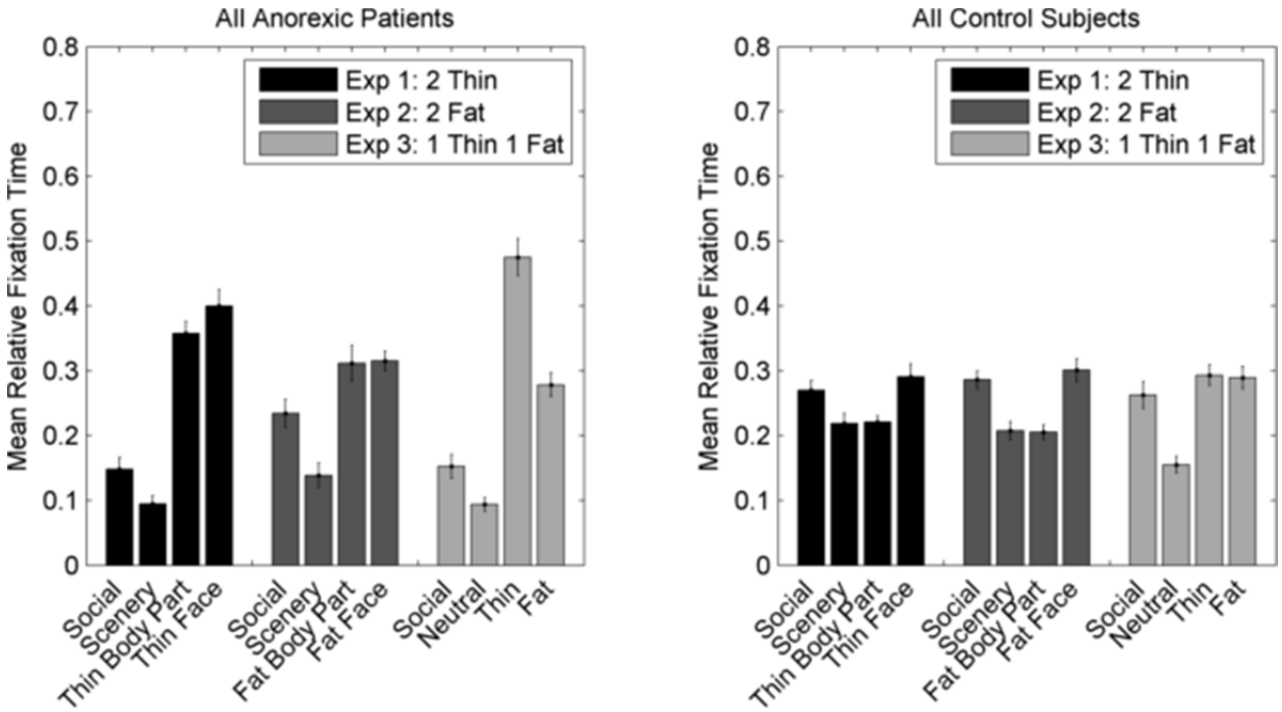


Figure 3.5: Mean Relative Fixation Time for all Anorexic (left) and Control (right) subject on all image types of all three experiments. See 3.2.2 for a detailed description of images in Experiment 1, Experiment 2, and Experiment 3 test slides. Error bars show standard error of the mean.

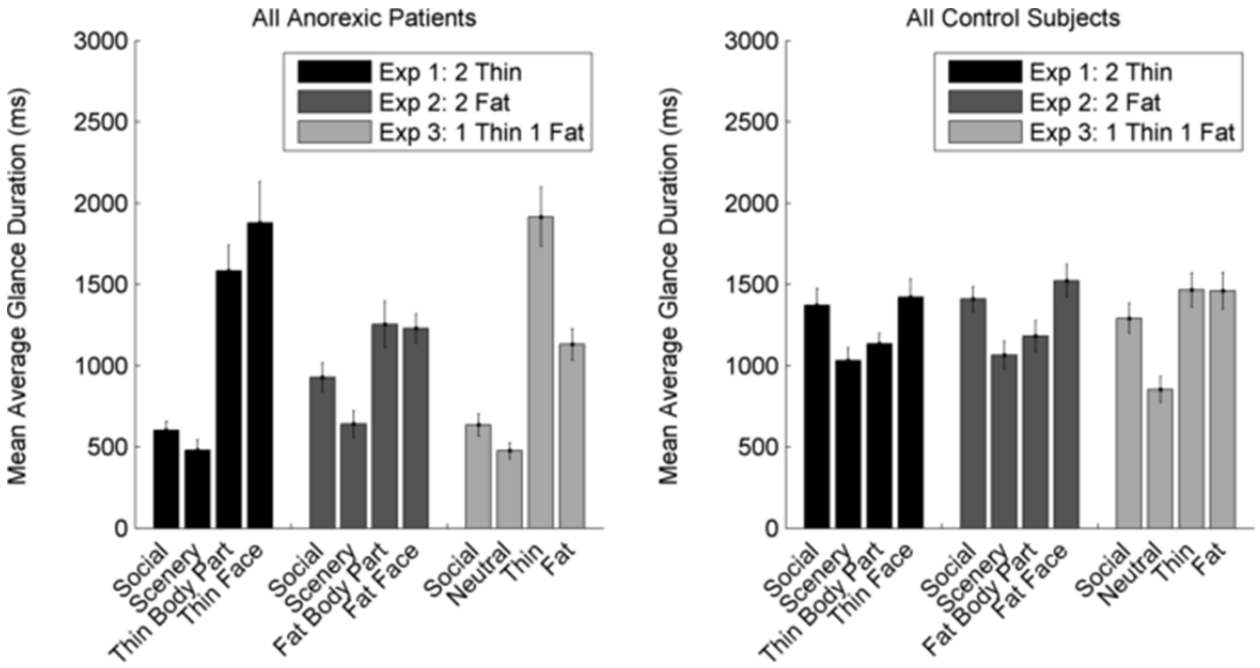


Figure 3.6: Mean Average Glance Duration for all Anorexic (left) and Control (right) subject on all image types of all three experiments. See 3.2.2 for a detailed description of images in Experiment 1, Experiment 2, and Experiment 3 test slides. Error bars show standard error of the mean.

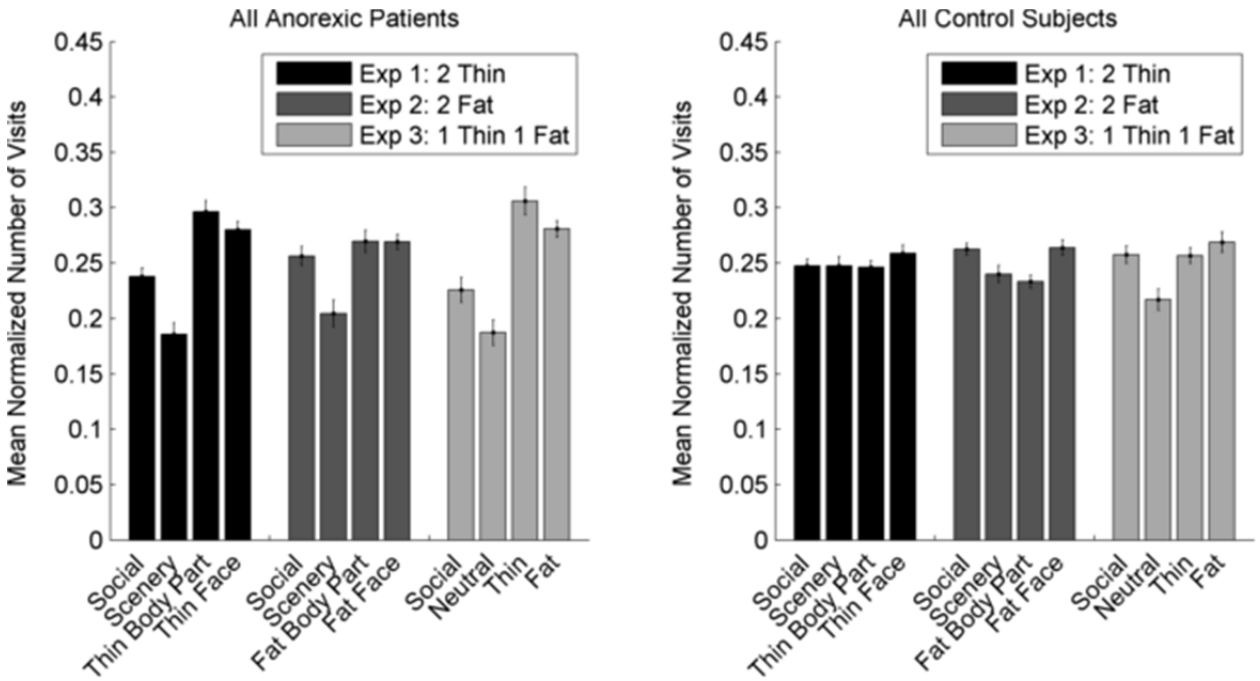


Figure 3.7: Mean Normalized Number of Visits for all Anorexic (left) and Control (right) subject on all image types of all three experiments. See 3.2.2 for a detailed description of images in Experiment 1, Experiment 2, and Experiment 3 test slides. Error bars show standard error of the mean.

This study demonstrates that adolescent AN patients have an attention bias towards both thin and fat images when they are presented alongside social or neutral images and this bias is not present in age-matched controls. Patients with AN spend longer durations looking at weight and shape related images and may return to them more frequently. This confirms and expands the results of previous indirect studies of attention by Stroop and Dot-probe tests, where AN patients were distracted by stimuli with themes fitting with their cognitive pathology. Interestingly, AN patients are also more likely to show attention bias towards thin images as compared to fat images when presented together (in competition) along with social and neutral images. This strongly supports the AN cognitive model of the drive for thinness. When "fat" is in competition with social images alone, AN patients do reveal and attention bias towards fat theme images.

On pro-anorexia web pages, images of fat people are referred to as "reverse thinspiration" and are seen as a tool for helping motivate adherence to weight loss behaviours (7). However, when "fat" is in competition with "thin", the "thin" wins, likely due to its attractiveness to the patients with AN, as it represents what they aspire to become. On the pro-anorexia web pages, the images of thin people are common and referred to as "thinspiration" (Norris et al., 2006). Again these images are used to bolster motivation to maintain weight loss behaviour (Norris et al., 2006). Previous Stroop or Dot-probe have not assessed directly the competition/attention-preference between "thin" and "fat" themes, thus this is a novel comparison and it confirms the importance of the thin ideal as playing a central role in the cognitions of patients with AN.

3.4. Description of Attention Bias Detector

Robust detection of attention biases within individuals requires a method to detect visual scanning behaviour similarities between subjects within a population and differences in VSB between populations (anorexic and control). The visual scanning parameters described thus far have shown clear biases in visual attention between the anorexic and control populations. However, these population biases cannot be extrapolated to individual characteristics without an analysis of the inter-subject variability within a population. For example, Figure 3.8 shows an AN patient who does not exhibit all the biases observed in the population of AN patients. While this AN patient has an attention bias towards thin body shape images in Experiment 1, she does not have a strong bias towards fat body shape images without faces. Additionally, biases between social and body shape images (thin and fat) are not observed in Experiment 3.

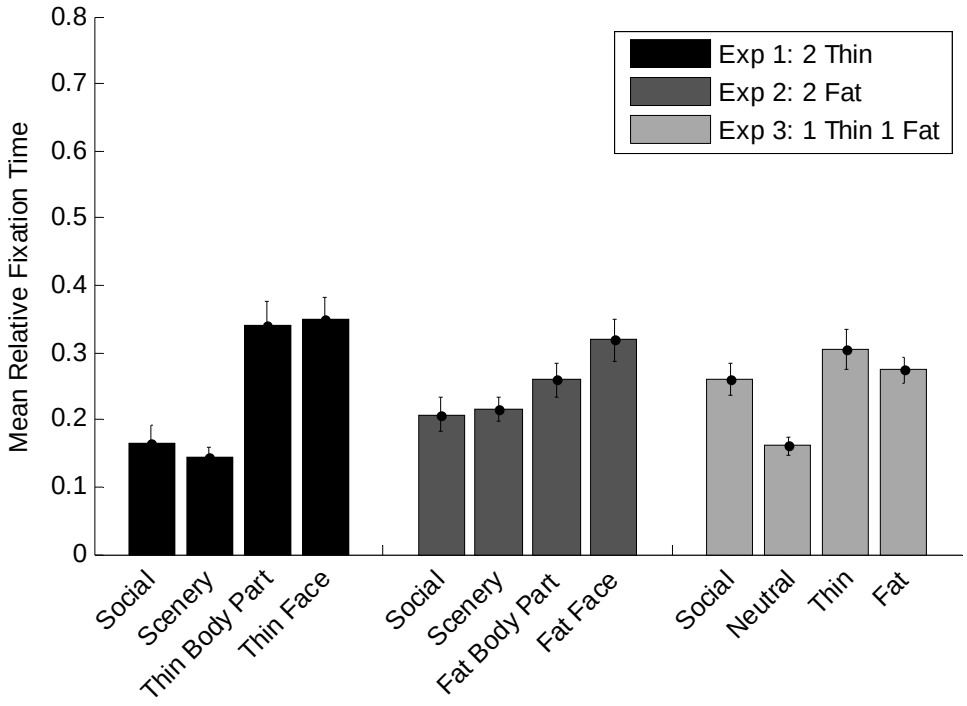


Figure 3.8: Example anorexic patient who does not have similar biases towards fat images as the population of anorexic patients.

3.4.1. Within-Subject Attention Bias as a Detector

To summarize how well VSB bias (i.e. significant differences in RFT between body-shape and social stimuli) classified the subjects, 4 two-tailed t-tests were performed for every subject: 1) between thin and non-thin images of Experiment 1, 2) between fat and non-fat images of Experiment 2, 3) between thin and social images of Experiment 3, and 4) between fat and social images of Experiment 3. The presence of a significant bias between the image types were used separately to predict the group (AN or control) of each subject. The positive predictive value, negative predictive value, sensitivity, and specificity are summarized for each pair of image types in Table 3.5. The bias in VSB that produced the best performance was the bias in scanning between thin body shape images and social and high valence images in Experiment 1. As a result, attention bias towards thin body shapes was chosen as the distinguishing feature to build a detector with the baseline performance having a sensitivity of 100% and a specificity of 70%.

Sensitivity and Specificity of significantly different mean on Exp 1: thin vs social

		Patients with AN		
		Positive	Negative	
Significantly higher mean on body shape image (t-test)	Positive	13	6	0.684211 **Pos PV**
	Negative	0	14	1 **Neg PV**
		1	0.7	
		Sensitivity	**Specificity**	

Sensitivity and Specificity of significantly different mean on Exp 2: fat vs social

		Patients with AN		
		Positive	Negative	
Significantly higher mean on body shape image (t-test)	Positive	9	5	0.642857 **Pos PV**
	Negative	4	15	0.789474 **Neg PV**
		0.69231	0.75	
		Sensitivity	**Specificity**	

Sensitivity and Specificity of significantly different mean on Exp 3: thin vs social

		Patients with AN		
		Positive	Negative	
Significantly higher mean on body shape image (t-test)	Positive	11	5	0.6875 **Pos PV**
	Negative	2	15	0.882353 **Neg PV**
		0.84615	0.75	
		Sensitivity	**Specificity**	

Sensitivity and Specificity of significantly different mean on Exp 3: fat vs social

		Patients with AN		
		Positive	Negative	
Significantly higher mean on body shape image (t-test)	Positive	8	4	0.666667 **Pos PV**
	Negative	5	16	0.761905 **Neg PV**
		0.61538	0.8	
		Sensitivity	**Specificity**	

Table 3.5: Sensitivity and specificity analysis of using presence of attention bias within patient as a classifier between AN or control.

3.4.2. Log-Likelihood Ratio Detector

To construct a detector for VSB biases in patients with AN, the likelihood that measurements of relative fixation time, RFT_j, on thin body forms (without faces) from a patient with AN or from a control subject were calculated by the likelihood ratio Λ_{Tj}:

$$\Lambda_{Tj}(RFT_j) = \frac{P(RFT_j|Class = AN)}{P(RFT_j|Class = Control)} \quad j = 1 \ldots . N$$

Equation 3.1: Calculation of likelihood ratio where, j = 1…N are the test slides, $P(RFT_j|Class = AN)$ and $P(RIT_j|Class = control)$ are the conditional probability densities of RFT_j for patients with AN and for controls, respectively.

When $\Lambda_{Tj}(RFT_j)$ is greater than 1, the measurements are more likely to come from a patient with AN. Under the assumption that VSB on one slide does not affect the VSB on any other slide (i.e., measurements are independent from slide to slide), the log-likelihood ratio that a set of N measurements of relative fixation times were obtained from a patient with AN or from a control subject is:

$$\log(\Lambda_T) = \log(\prod_{j=1}^{N} \Lambda_{Tj})$$

Equation 3.2: Calculation of log-likelihood ratio for a single subject.

When the output of the detector is greater than a threshold (0 = log 1), the processor detects biases in visual scanning behaviour that are consistent with that of patients with AN. The threshold can be adjusted to control the sensitivity and specificity of the processor.

Figure 3.9 shows histograms of the relative fixation time on images with thin body forms for patients with AN (3.9A) and control subjects (3.9B). For control subjects approximately 87.5% of the measurements of RFT_j were lower than 0.36 (the sum of the first two bins of the histogram in Figure 3.9B) compared with only 51.67% for patients with AN (Figure 3.9A). On the other hand, for patients with AN, 22.78% of the measurements of RFT_j were higher than 0.5 (i.e. the sum of the last three bins of the histogram in Figure 3.9A) compared with only 2.14% for control subjects. The differences in the percentage of measurements of both low RFT_j (<0.36) and high

RFT_j (>0.5) between patients with AN and control subjects were statistically significant (RFT < 0.36: $t(31)$= -5.8984, P<0.001; RFT > 0.68: $t(31)$=- 6.2219, P<0.001).

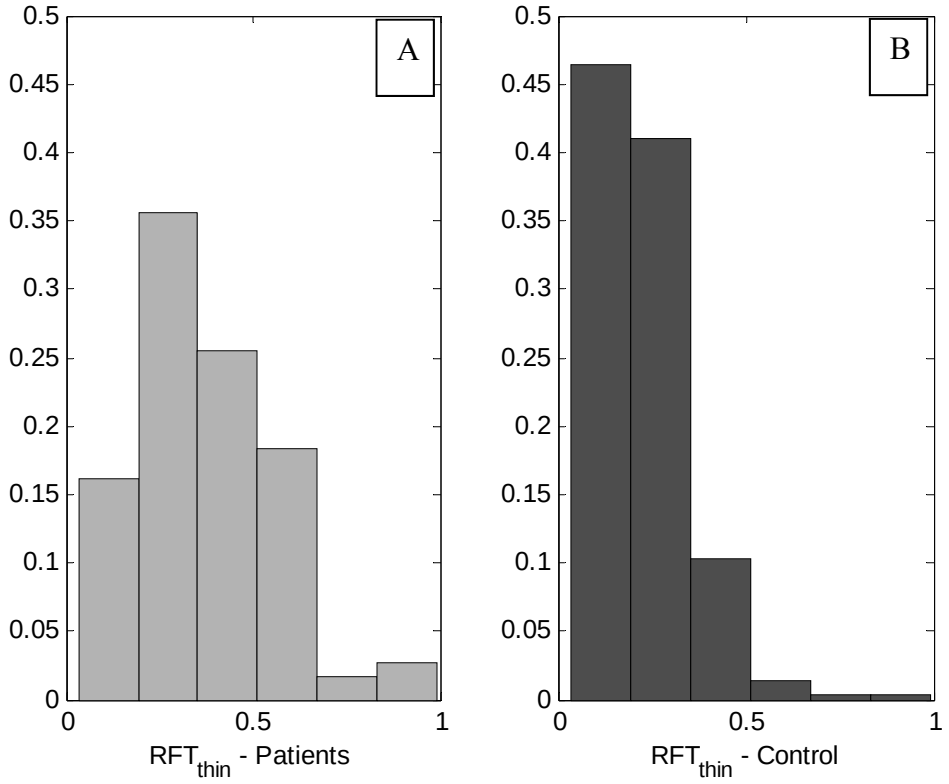


Figure 3.9: Histograms of the relative fixation time on images with thin body forms(RFT) for patients with AN (A) and control subjects (B). The histograms in (A) and (B) have 191 and 294 observations, respectively. The boundaries of the bins in the histograms were determined by dividing the range of the measured RIT_j (0.035 -1) into 6 equal intervals (each bin has at least one observation to facilitate the calculations of the likelihood ratios in Equation 3.2).

The histograms in Figure 3.9 A and B were used in the calculations of the log-likelihood ratios (Equation 3.2) for the measurements of RFT respectively. For each measurement, the conditional probability density was approximated by the height of the bin in the histogram that included that measurement. For each subject/patient the histograms were re-calculated from a data set that excluded the data for this patient/control, so that the calculations of the log-likelihood ratios for each patient/control were not biased by her own data (leave one out procedure). Figure 3.10 shows that when the threshold of the log-likelihood processor is set to 0 (the likelihood ratio is 1), the processor correctly identified 92% of the patients (sensitivity 92%, false negative rate 8%) and 95% of the control subjects (specificity 95%, false positive rate 5%). The receiver operating

curve (ROC, Figure 3.11) shows that by changing the threshold of the processor the performance of the detector can change from having 100% sensitivity and 55% specificity to 77% sensitivity and 100% specificity.

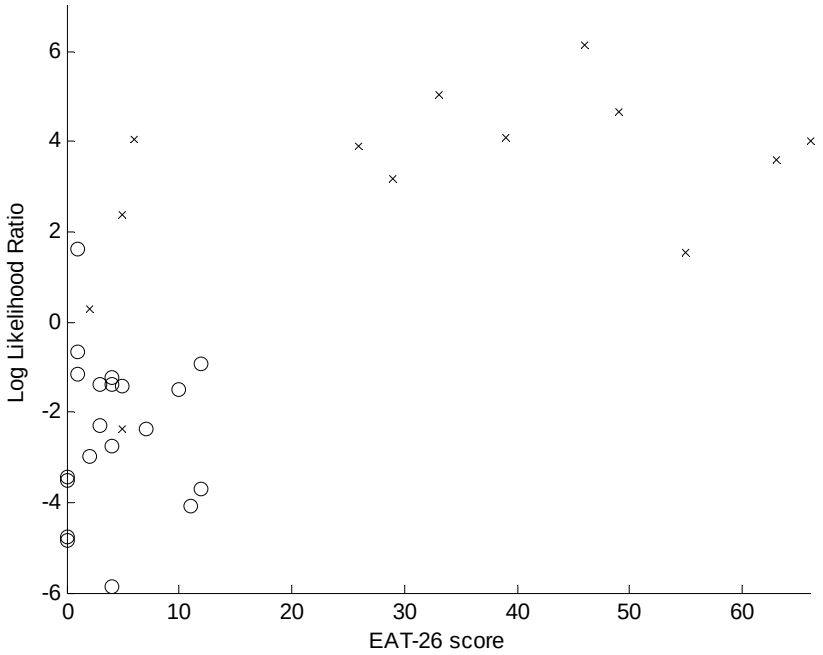


Figure 3.10: Log-likelihood ratio vs. EAT-26 score. X's and O's indicate AN patients and control subjects respectively.

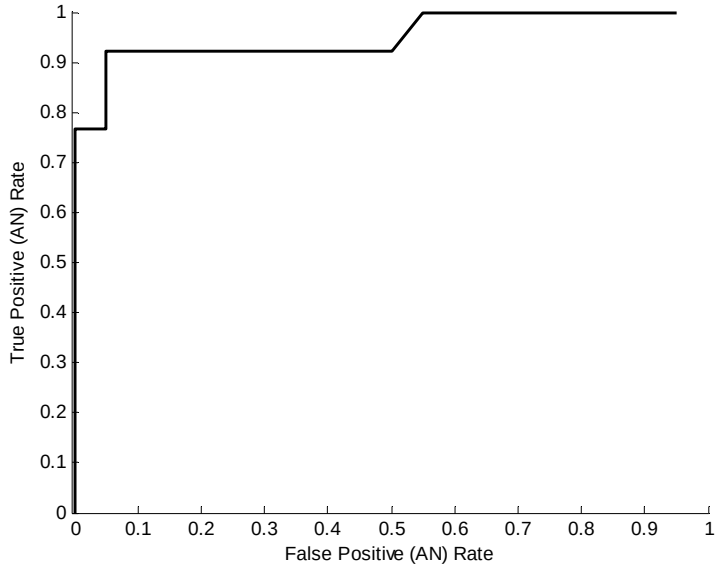


Figure 3.11: The Receiver Operating Curve (ROC) of the processor to detect visual scanning biases in patients with Anorexia Nervosa using RFT on thin body forms (without faces).

To assess the number of images required for good detection, the area under the ROC curve (AUC) was calculated as the number of test slides (observations of RFT on thin body form images) were varied from 1 to 16. For each number of test slides, all possible combinations of the 16 test slides were averaged together. The results are shown in Figure 3.12. The Mean AUC increases fairly rapidly and plateaus at approximately 0.9. This results suggest that increasing the number of images further will not improve the detection of attention bias in AN dramatically. Thus, it appears that there are AN patients (one in this sample 13 of subjects) that cannot be identified by her attention bias (or rather lack of) to thin body shape images without faces (See the single x below 0 in Figure 3.10). This subject may be identified by her attention bias towards other body shape image types and a processor that incorporates more than one dimension of bias to make a decision may have improved performance.

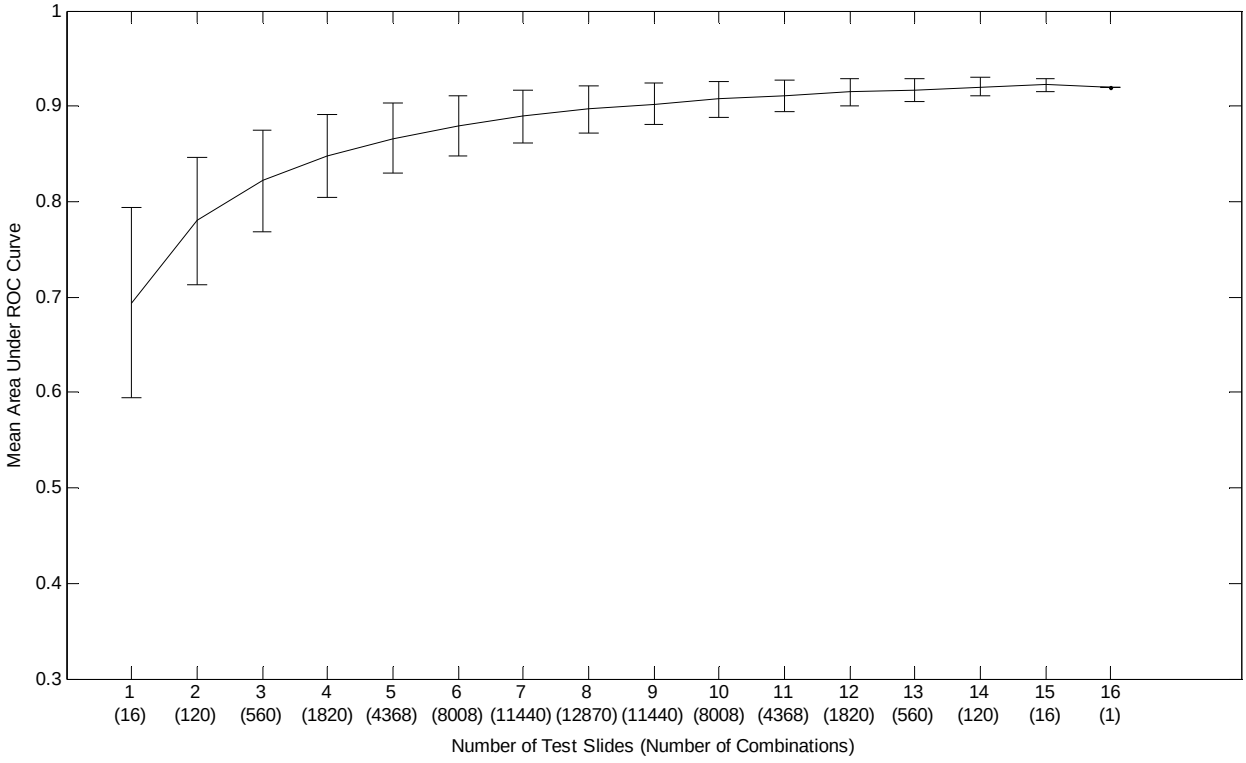


Figure 3.12: Mean Area under ROC curve when using increasing number of test slides. Number in brackets shows number of combinations that went into the calculation of the average. Error bars depict ±1 standard deviation.

3.4.3. Combining Multiple Dimensions of Attention Bias

To demonstrate how multiple dimensions of bias in VSB may be combined together in the log-likelihood processor, bias towards thin body forms and social interaction were chosen based on evidence that these two features may characterize AN well (Long et al., 1994; Garner et al., 1997). Since these two image types appear on the same slide in Experiment 1, two normalized measures were defined: a) the relative interest in images with thin body forms (RIT) and b) the relative interest in images with social interactions (RIS) shown in Equation 3.3 and Equation 3.4 respectively.

$$RIT_j = \frac{RFT_{TBj}}{RFT_{TBj} + RFT_{TFj} + RFT_{HVj}} \quad j = 1 \dots \dots . . N$$

Equation 3.3: Calculation of relative interest in images with thin body forms (RIT). RFT_{TBj}, RFT_{TFj}, RFT_{HVj}, are the Relative Fixation Times on Thin body parts, thin images with faces, and high valence (scenery) images respectively.

$$RIS_j = \frac{RFT_{SOj}}{RFT_{SOj} + RFT_{TFj} + RFT_{HVj}} \quad j = 1 \dots \dots . . N$$

Equation 3.4: Calculation of relative interest in images with social interaction (RIS). RFT_{SOj}, RFT_{TFj}, RFT_{HVj}, are the Relative Fixation Times on social, thin images with faces, and high valence (scenery) images respectively.

The relative interests span the range between [0 – 1]. Note that if all the images on a slide are scanned in a similar manner, $RIT_j = RIS_j = 0.33$.

To construct a detector for VSB biases in patients with AN, the likelihood that measurements of relative interests, $RIT_j and\ RIS_j$, from a patient with AN or from a control subject were calculated by the likelihood ratios ($\Lambda_{Tj}\ and\ \Lambda_{Sj}$):

$$\Lambda_{Tj}(RIT_j) = \frac{P(RIT_j | Class = AN)}{P(RIT_j | Class = Control)} \quad j = 1 \dots . N \qquad (A)$$

$$\Lambda_{Sj}(RIS_j) = \frac{P(RIS_j | Class = AN)}{P(RIS_j | Class = Control)} \quad j = 1 \dots . N \qquad (B)$$

Equation 3.5: Calculation of likelihood ratios where, j = 1…N are the test slides, $P(RIT_j | Class = AN)$ and $P(RIT_j | Class = control)$ are the conditional probability densities of RIT_j for patients with AN and for controls, respectively. $P(RIS_j | Class = AN)$ and $P(RIS_j | Class = control)$ are the conditional probability densities of RIS_j for patients with AN and for controls, respectively.

When $\Lambda_{Tj}(RIT_j)$ and $\Lambda_{Sj}(RIS_j)$ are greater than 1, the measurements are more likely to come from a patient with AN. Under the assumption that VSB on one slide does not affect the VSB on any other slide (i.e., measurements are independent from slide to slide), the log-likelihood ratios

that a set of N measurements of relative interests were obtained from a patient with AN or from a control subject are:

$$\log(\Lambda_T) = \log(\prod_{j=1}^{N} \Lambda_{Tj}) \qquad (A)$$

$$\log(\Lambda_S) = \log(\prod_{j=1}^{N} \Lambda_{Sj}) \qquad (B)$$

Equation 3.6: Calculation of log-likelihood ratio for thin (A) and social interaction (B) images for a single subject.

This processor for the detection of VSB biases in patients with AN combines the log-likelihood ratios for RIT and RIS. When the output of the detector is greater than a threshold (0 = log 1), the processor detects biases in visual scanning behaviour that are consistent with that of patients with AN. The threshold can be adjusted to control the sensitivity and specificity of the processor.

Figure 3.13 shows the difference between the mean relative interest in images with thin body forms(RIT) and the mean relative interest in images with social interactions(RIS), for individual subjects, as a function of the EAT-26 scores. The dashed horizontal line at the mean difference of 0 indicates the expected difference when the VSB on the four categories of images is the same. The dashed vertical line at an EAT-26 score of 20 indicates the clinical cut point. As a group, $(RIT - RIS)$ for AN patients (M = 0.4262, SD = 0.1894, range 0.1516 to 0.7291) is larger (t (31) = 6.8135, $P < 0.001$) than that of the control subjects (M = -0.0193, SD = 0.1797, range -0.3456 to 0.2992). All thirteen patients with AN showed positive differences (i.e., $(RIT - RIS) > 0$) which implies that patients with AN have more interest in images with thin body forms than in images with social interactions. Note that 8 of the control subjects (40.0%) also have more interest in images with thin body forms than in images with social interactions. The mean values of $(RIT - RIS)$ for 9 of the AN patients (69.2 %) are larger than the values of $(RIT - RIS)$ for all the subjects in the control group.

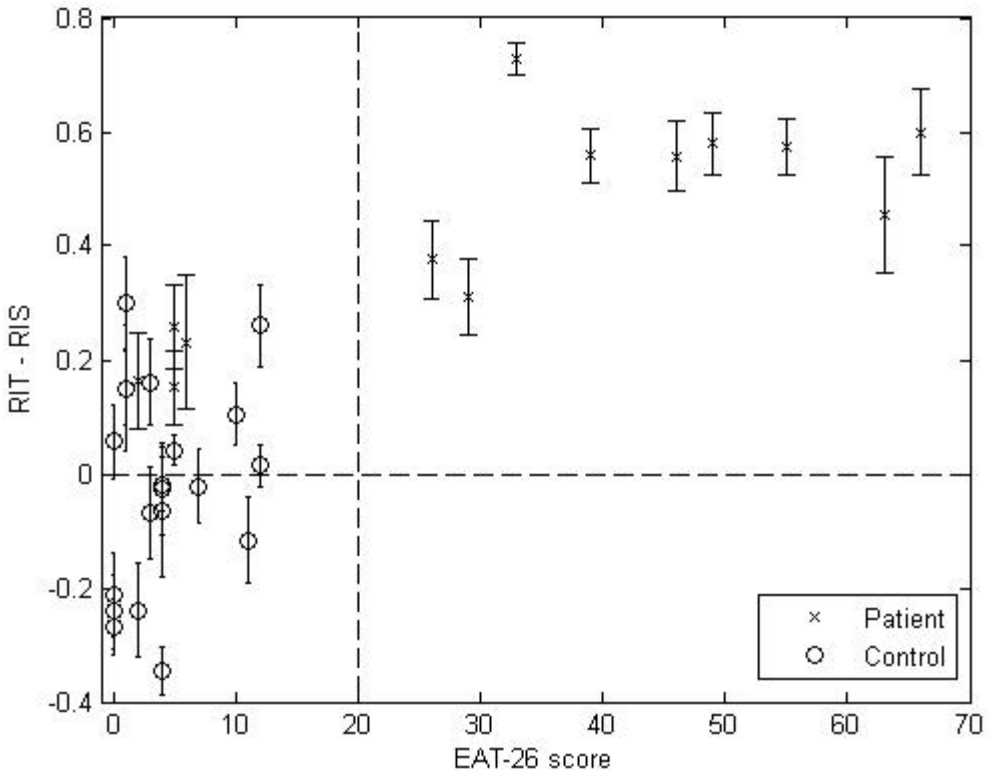


Figure 3.13: Difference between the mean relative interest in images with thin body forms(RIT) and the mean relative interest in images with social interactions(RIS), for individual subjects, as a function of the EAT-26 scores. X's and O's indicate the mean values for patients and controls (lines indicate ± 1 S.D.), respectively.

Figure 3.14 shows histograms of the relative interest in images with thin body forms (RIT) for patients with AN (3.14A) and control subjects (3.14B). For control subjects approximately 63.3% of the measurements of RIT_j were lower than 0.36 (the sum of the first two bins of the histogram in Figure 3.14B) compared with only 18.3% for patients with AN (Figure 3.14A). On the other hand, for patients with AN, 38.2% of the measurements of RIT_j were higher than 0.68 (i.e. the sum of the last two bins of the histogram in Figure 3.14A) compared with only 4.7% for control subjects. The differences in the percentage of measurements of both low RIT_j (<0.36) and high RIT_j (>0.68) between patients with AN and control subjects were statistically significant (RIT < 0.36: $t(31)=7.41$, $P<0.001$; RIT > 0.68: $t(31)=-6.47$, $P<0.001$). The percentage of measurements of RIT_j in the range of 0.33-0.66 (i.e. the sum of the two middle bins of the histograms) for patients with AN and for control subjects were statistically similar ($t(31)=-2.0113$, $P>0.05$).

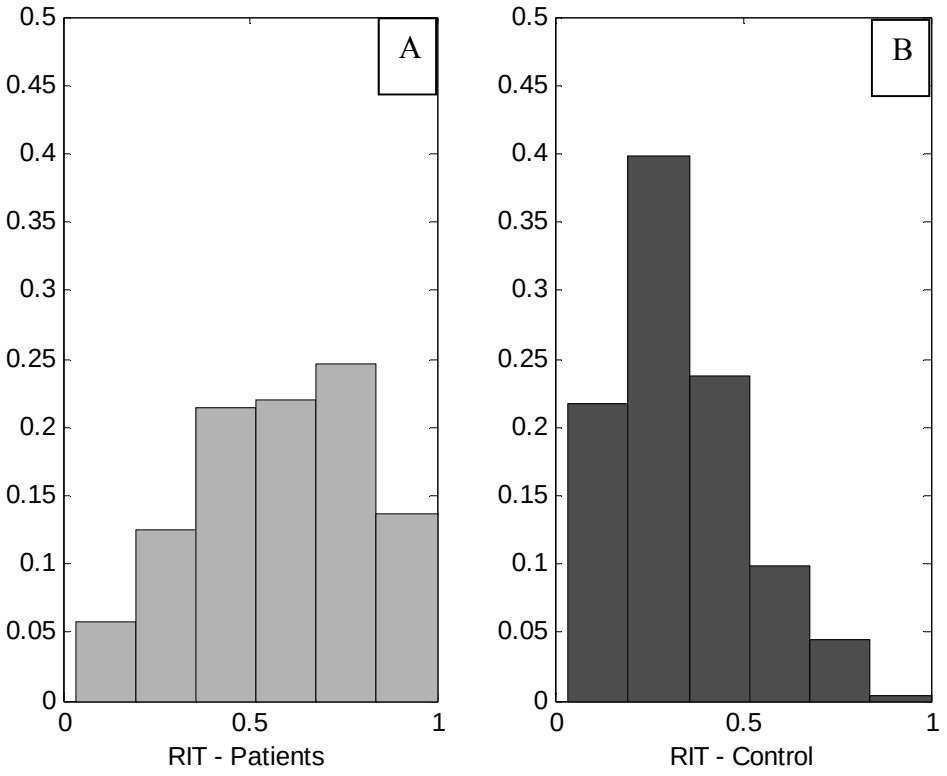


Figure 3.14: Histograms of the relative interest in images with thin body forms(RIT) for patients with AN (A) and control subjects (B). The histograms in (A) and (B) have 191 and 294 observations, respectively. The boundaries of the bins in the histograms were determined by dividing the range of the measured RIT_j (0.035 -1) into 6 equal intervals (each bin has at least one observation to facilitate the calculations of the likelihood ratios in Equation 3.6A).

Figure 3.15 shows histograms of the relative interest in images with social interactions (RIS) for patients with AN (3.15A) and control subjects (3.15B). For patients with AN approximately 72.3% of the measurements of RIS_j were lower than 0.25 (low interest in images with social interactions) compared with only 31.0% for control subjects. The difference in the percentage of measurements of low RIS_j (<0.25) between patients with AN and control subjects was statistically significant ($t(31) = 6.15$, $P< 0.001$). On the other hand, for control subjects, both the % of measurements of RIS_j that were higher than 0.50 (i.e. the sum of the last two bins of the histogram in Figure 3.15A; high interset in social images) and between 0.25-0.5 (medium interest in social images; the sum of the middle two bins of the histogram in Fig 3.15A) were higher than the % of the corresponding measurements in patients with AN ($RIS > 0.50$: $t(31) = 3.14$, $P < 0.004$; $0.25 < RIS < 0.50$: $t(31) = 4.41$, $P< 0.001$). The histograms show that even though the distributions of RIT and RIS for patients with AN and control subjects span

approximately the same range, patients with AN are more likely to have higher RIT values and controls are more likely to have higher RIS values. The histograms also demonstrate that the distributions of RIT and RIS for patients with AN and for control subjects are non-Gaussian.

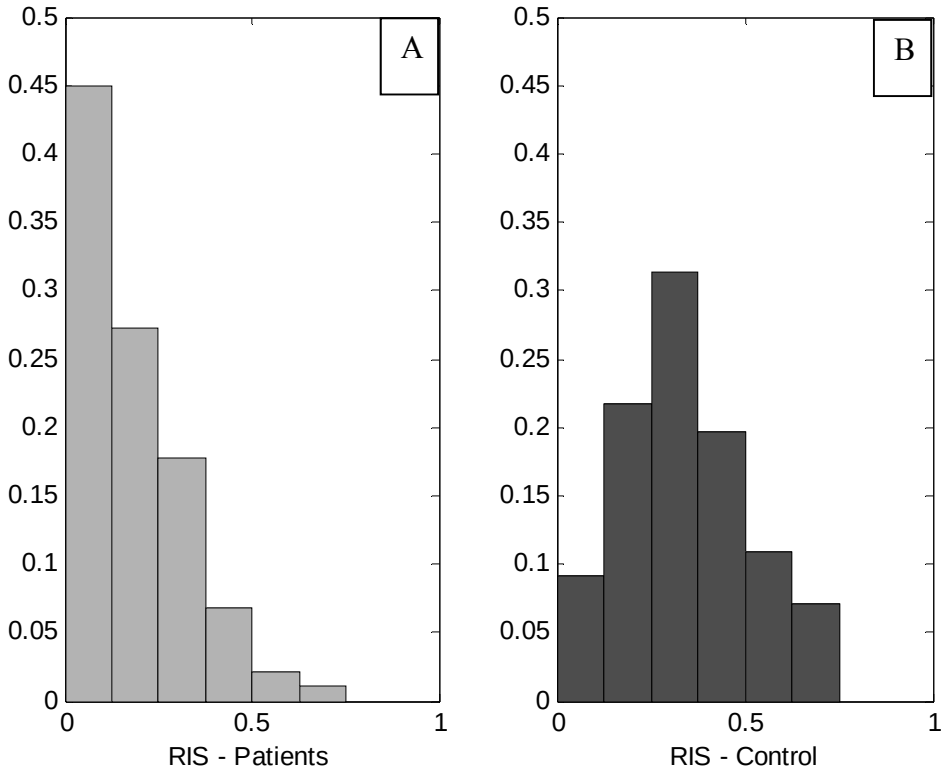


Figure 3.15: Histograms of the relative interest in images with social interactions(RIS) for patients with AN (A) and control subjects (B). The histograms in (A) and (B) have 191 and 279 observations, respectively. The boundaries of the bins in the histograms were determined by dividing the range of the measured RIS_j (0 – 0.75) into 6 equal intervals (each bin has at least one observation to facilitate the calculations of the likelihood ratios in Equation 3.6B).

The histograms in Figure 3.14 and 3.15 were used in the calculations of the log-likelihood ratios (Equation 3.6A and 3.6B) for the measurements of RIT and RIS, respectively. For each measurement, the conditional probability density was approximated by the height of the bin in the histogram that included that measurement. For each subject/patient the histograms were re-calculated from a data set that excluded the data for this patient/control, so that the calculations of the log-likelihood ratios for each patient/control were not biased by her own data (leave one out procedure). When the threshold of the log-likelihood processor is set to 0 (the likelihood ratio is 1), the processor correctly identified 92% of the patients (sensitivity 92%, false negative

rate 8%) and 90% of the control subjects (specificity 90%, false positive rate 10%). The receiver operating curve (ROC, Figure 3.16) shows that by changing the threshold of the processor the performance of the detector can change from having 100% sensitivity and 90% specificity to 54% sensitivity and 100% specificity. If the threshold is adjusted for 100% sensitivity, the processor detects visual scanning behaviour that is consistent with that of patients with AN in all the patients with AN and in 2 of the 20 (10%) control subjects.

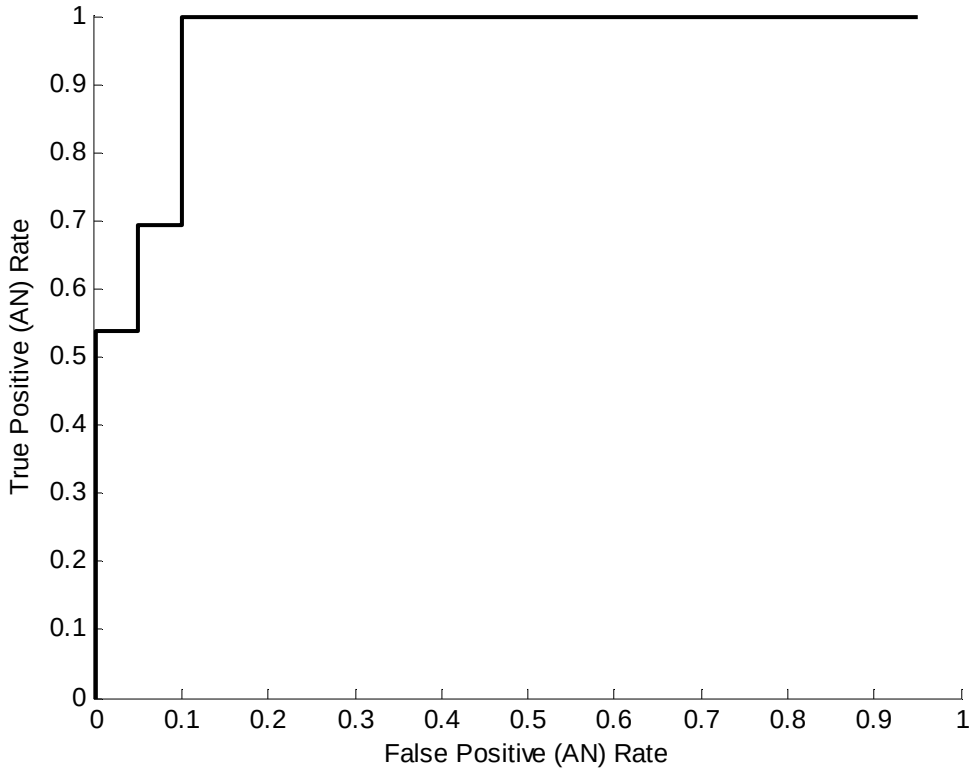


Figure 3.16: The Receiver Operating Curve (ROC) of the processor to detect visual scanning biases in patients with Anorexia Nervosa.

The multidimensional log-likelihood ratio detector has better sensitvity than the single dimension detector at the cost of specificity (Figure 3.16 vs. Figure 3.11). Use of attention bias away from social interaction images in addition to attention bias towards thin body forms provides extra information on the patient that could not be identified as AN using only bias towards thin body forms. The loss of specificity is a result of several control subjects exhibiting avoidance of social interaction images. These results illustrate that detection performance can be reduced if the individual dimensions of attention bias are not sensitive and specific. An optimal design of a

detector may be developed and validated by learning weights between dimensions of attention bias. However, this requires a significantly larger sample of subjects and is outside the scope of this discussion.

In summary, a novel method to objectively detect biases in visual scanning behavior of patients with anorexia nervosa is described. The method measures the subject's relative interest in images with thin body forms (RIT) and in images with social interactions (RIS) and uses a log-likelihood ratio processor to detect VSB biases. When the threshold of the processor is 0, measurements of RIT and RIS from twelve of the thirteen patients with AN, including three of the four AN patients that either minimized or misrepresented their behaviour (patients with EAT-26 scores that are less than 20), will be correctly classified as measurements from patients with AN. At this threshold, measurements from two of the twenty control subjects will be classified incorrectly. The patient that did not test positive was an athlete who was identified very early in the course of illness (duration of symptoms 2 months) and clinically had relatively mild cognitive symptoms of weight and shape concerns. When the threshold of the processor is adjusted to achieve 100% sensitivity (i.e., measurements of RIT and RIS from all AN patients are correctly classified) measurements from two of the control subjects will be incorrectly classified by the processor. By detecting biases in VSB for images with thin body forms and for images with social interactions the detector is able to achieve both high sensitivity and high specificity. Analysis of VSB provides an objective test to identify AN even in patients that do not score in the clinical range on self-report measures. It provides a method to screen adolescents who may be minimizing or misrepresenting the presence of AN cognitions and behaviors. Importantly, it bypasses the volitional component of self-report measures and therefore is less available to conscious manipulation.

The set of parameters that best describes differences in VSB between patients and controls depends on the nature of the disorder (Hannula et.al., 2010). For patients with anxiety, the time delay to the first fixation on threatening visual stimuli is often used to detect biases (Hermans et al., 1999, Rinck and Becker, 2006). In AN, preoccupation with thoughts about shape and weight lead to rumination on images with thin body forms while anhedonia and self-absorption may lead to avoidance of images with social interactions. The two visual scanning parameters: relative

interest in images with thin body forms (RIT) and relative interest in images with social interactions (RIS) provide indirect measurements of the above two processes.

The relative interests in images with thin body shapes and in images with social interactions in patients with AN is correlated with specific patient's characteristics. As a group, AN patients with high $RITs$ and low RISs had been ill and in treatment for over a year, resistant to recovery clinically or were low weight at the time of the study. Clinically, patients that were most rigidly adherent to the cognitive patterns typical to AN had the largest biases in visual scanning behaviour and AN patients with lower $RITs$ and higher RISs began to normalize their eating patterns in treatment more readily. As these characteristics are not always predictable from the EAT-26 scores, it is possible that one can use the analysis of biases in VSB to help in the prognosis and treatment of patients with AN. An objective indicator for attentional biases in AN with high sensitivity and specificity might help to identify subjects who are at risk of developing anorexia nervosa. Such an objective indicator is important since a significant proportion of the adolescent AN population is unable or unwilling to self-identify.

3.4 Conclusions

Attention biases were characterized using visual scanning parameters in the adolescent anorexic population. Compared to controls, anorexic patients exhibited visual attention bias towards thin body shape images and fat body shape images. A preference for thin body shape images was observed when thin and fat body shape images competed for attention simultaneously confirming the importance of the thin idea as playing a central role in the cognitions of AN patients. A processor that detects VSB biases associated with increased interest in images with thin body shapes and decreased interest in images with social interactions is shown to achieve both high sensitivity and high specificity. The log-likelihood ratio processor described in the chapter can be used for reliable detection of VSB biases in individuals with AN even when symptoms are minimized or misrepresented.

Chapter 4. Early Detection of Antidepressant Efficacy in Major Depressive Disorder

4.1. Introduction

Depression is a psychiatric disorder that tends to be chronically recurring and affects about 5% of the Canadian population (Canadian Psychiatric Association, 2001). The world health organization has identified it as one of the most burdensome diseases in the world (WHO, 2002). The incidence of depression as a comorbid condition with other serious illnesses such as stroke, hypertension and cancer has come under increasing public awareness. The evolution of antidepressants in terms of new biological targets has meant a significant improvement in the care and quality of life for people with depression. The clinical treatment for depression has come a long way since the advent of vapours and the water cure. Perhaps the first really effective treatment was electroconvulsive therapy or ECT in the 1930s, and, although it did work remarkably well in some patients, its popularity dropped off as the era of drugs emerged with the use of the "Tricyclics" and then the "MAO inhibitors". Both of these classes of drugs increased the levels of certain brain neurotransmitters (norepinephrine and serotonin) that are hypothesized to be deficient in depressed people, but the side effects were often debilitating and people had difficulty staying on them for any length of time. Then, in the 1980's, came the SSRI's or selective serotonin reuptake inhibitors, such as Prozac. These drugs were a real step forward because they were effective antidepressants and had fewer side effects of the Tricyclics or MAO inhibitors. As a result, patients were more inclined to keep on taking their medication.

However, one of the more enduring and problematic puzzles in the pharmacotherapy of major depression pertains to the time lag between onset of drug-induced central neurochemical changes and the appearance of symptom remission. For example, biochemical studies have shown that it takes about 2-8 hours for SSRI's to stop the uptake of serotonin into the presynaptic neuron, yet the actual therapeutic effect of significant alleviation of depressive symptomatology may not appear until after 2-6 weeks of daily dosing. This pharmacological /clinical disconnect has been the subject of a number of studies that have proposed explanations including autoreceptor desensitization, reversal of serotonin up- regulation (Celada et al., 2004) or neurogenesis

(Malberg, 2004) that are congruent with the time lag for antidepressant action. Regardless of the exact reason for this time lag, its existence poses a significant clinical problem, i.e. if the treatment turns out not to be effective, precious treatment time has been lost, translating into increased risk for serious consequences and increased suffering for the patient. It is obvious that a system that could detect early "microchanges" in mood or mood-change surrogates would be extremely valuable because it could indicate whether a treatment was destined to be efficacious and therefore should be continued--- or if not, should be abandoned early in favour of a different drug or drug class.

One such early mood-change surrogate could be a change in selective attention. Several studies (Dalgleish and Watts, 1990; Mogg and Bradley, 1998) showed that selective attention to negatively valenced information supports and sustains the maladaptive patterns of information processing that are characteristic of depressive states. Moreover, cognitive biases in information processing play an important role in the etiology and maintenance of emotional disorders. (MacLeod et al., 2002).

In a variety of Dot-Probe and Stroop tasks, attention biases towards dysphoric stimuli were observed in depressed populations. These attention biases were observed when stimuli were presented for longer durations but not when the stimuli were presented for shorter durations (Williams et al., 1997; Bradley et al., 1997; Gotlib et al., 2004). Building on these results, Eizenman et al (2003) and Yu and Eizenman, (2004) described and validated a new technology to measure changes in selective attentional biases in people with major depression. Results showed that subjects with depressive disorder spent significantly more time looking at images with dysphoric themes compared to a group of normal controls. It was concluded that subjects with major depressive disorder selectively attend to mood congruent material and that depression appears to influence the elaborative stages of processing when dysphoric images are viewed. Since then, the findings have been repeated in similar eye tracking studies. Kellough et al. (2008) found that the visual scanning biases were maintained over the time course of viewing (up to 30 seconds). Also, Caseras et al. (2007) found that depressed individuals were no more likely than controls to shift their attention toward negative stimuli, but once their attention was focused on negative stimuli, they spent significantly more time looking at these stimuli than did

non-depressed controls. Finally, Sears et al. (2011) observed that dysphoric, never depressed individuals behaved similarly to previously-depressed individuals by spending less time attending to positive images. They also found that dysphoric individuals initially oriented their gaze to dysphoric images more frequently than never-depressed controls.

The study that is described in this chapter aims to expand on previous results by testing the hypothesis that, in major depression, shifts in selective attention away from negatively valenced stimuli, expressed as changes in visual scanning behaviour, occur early on in drug treatment and can predict eventual symptom remission. This, in essence represents an early and sensitive biomarker of antidepressant efficacy. It is hypothesized that this shift can occur before any substantive changes in a score on a depression rating scale or in subjective feelings, but would nevertheless be an early sign of remission response and will be invaluable in optimizing treatment.

The primary objective is to show that, in Major Depressive Disorder, shifts in visual attention away from dysphoric or negatively valenced themes occurring within the first weeks of treatment with an antidepressant and prior to any significant change in Hamilton Depression Rating Scale (HDRS-17)**,** represent an early marker of medication efficacy. We postulate that shifts in visual attention away from negatively valenced themes, as measured by reduction in mean relative fixation time on dysphoric images, will only occur in patients who, after six weeks of medication, demonstrate a significant reduction in depressive symptom scores on the rating scales (responders) compared to baseline (start of treatment) and not in those whose depression scores remain high (non-responders). A significant reduction in depressive symptoms is defined as a change of 50% or greater on the HDRS-17.

4.2. Study Design

To show that changes in visual attention can act as an early marker for antidepressant efficacy, the visual scanning patterns of patients were measured once per week for a period of six weeks following a change in medication management. Changes in visual scanning were monitored in a total of 28 patients with Unipolar Major Depressive Disorder referred for pharmacological treatment. Each session consisted of an evaluation of depressive symptoms using the HDRS-17 rating scale followed by a presentation of slides and recording of visual scanning patterns.

4.2.1. Participants

All subjects were people with Major Depressive Disorder recruited through advertisement at the University of Toronto and the local newspapers. A psychiatrist evaluated the medical and psychiatric history of all prospective patients, explained the study and offered a consent form for their consideration. Twenty-four hours later, the patients were contacted by the study coordinator who determined their willingness to participate and booked an initial visit for those who responded positively and met all the inclusion and exclusion criteria. During this screening visit the study physician assessed whether each patient met DSM-IV-TR criteria for Major Depression using the Mini International Neuropsychiatric Interview (MINI). Throughout the entire study the severity of depressive symptoms were assessed using the HDRS-17.

Inclusion	Exclusion
• Subjects must meet DSM-IV-TR criteria for a diagnosis of Major Depressive Disorder. • Subjects must have a score of at least 20 on the HDRS-17 at time of assessment. • Subjects must be between the ages of 18 and 65. • Able to provide informed written consent.	• Subjects must not suffer from unstable medical conditions. • Previous lack of response to Duloxetine. • Contraindication to Duloxetine. • Currently on Fluoxetine or MAOI antidepressant. • Recent history of substance abuse/dependence (excluding nicotine and caffeine) within the past 6 months as defined by DSM-IV-TR . • History of suicide attempts or self-harm within the last 12 months. • Previous history of non-response to two or more previous adequate antidepressant medication trials. • Anxiety Disorder, Bipolar Affective Disorder, Schizophrenia or other psychotic disorder (including psychotic disorder due to general medical condition, substance-induced psychotic, psychotic disorder not otherwise specified) as defined by the MINI. • Electroconvulsive therapy (ECT) within the 3 months prior to beginning of study. • Inability to communicate in English.

Table 4.1: Inclusion and exclusion criteria for subjects.

4.2.2. Antidepressant Medication

All patients received Duloxetine antidepressant monotherapy. Those patients who were currently receiving other antidepressant medications underwent a washout period of one week (between Visit 1 and 2). Following the baseline visit (V2), patients received 60 mg Duloxetine PO once daily for the duration of the study. Patients were instructed to take the medication with food.

4.2.3. Procedure

VISIT	V1 (Screen) *	V2 Baseline	V3	V4	V5	V6	V7	V8
Week	-1	0	1	2	3	4	5	6
Informed Consent	X							
Medical & Psychiatric Hx	X							
Demographics	X							
Entry Criteria	X							
Physical Exam	X							
Vitals (Wt, Ht)	X							
Urine Drug Screen	X							
MINI	X							
Blood pressure, pulse		X		X		X		X
HDRS-17		X	X	X	X	X	X	X
Visual Scanning		X	X	X	X	X	X	X
Adverse Events		X	X	X	X	X	X	X
Medication Compliance		X	X	X	X	X	X	X
Review of Concurrent Medication	X		X	X	X	X	X	X

Figure 4.1: Study Flowchart.

Prior to the change in medication management, the visual scanning patterns of each patient were measured at the V2 visit (within one week of the first visit), to establish a baseline response to the visual stimuli. Then, following the medication adjustment, each patient had their visual scanning patterns recorded once per week for a further six weeks.

4.2.4. Presentation of Visual Stimuli

Subjects' visual scanning patterns were recorded as they viewed a presentation of visual stimuli. The presentation was created using the processes described in Chapter 2.2. The visual stimuli were organized as a series of slides, each containing four images. Images were selected from the International Affective Picture System (IAPS) [CSEA-NIMH (1999)], a standardized library of images rated for valence, arousal, and dominance. Valence is a measure of a subject's relative pleasure in viewing an image, while arousal classifies subject reaction to an image in the

continuum from relaxed to excited. Dominance rates the subject's feeling of control when viewing an image. The IAPS numerical values for these quantities have been assigned through tests on normal populations [Lang *et al* (1999)]. Since the focus of the current study is on attention to dysphoric stimuli, valence and thematic content were the criteria used to select the images. The four images on each slide included two dysphoric images and two control-pleasant images whose positions on the slide are assigned randomly. Dysphoric images displayed themes of loss and sadness while control-pleasant images presented themes of interpersonal attachment and social content. Dysphoric images were selected to have valence ratings below 4, while control-pleasant images were rated above 6. Each test slide contained two high valence images and two low valence images. Images on the same slide were matched for similar arousal ratings. An example of a test slide is shown in Figure 4.2.

Figure 4.2: Example test slide with two dysphoric (low valence) and two social (high valence) images.

A total of 35 slides were presented to subjects while their visual scan paths were recorded at each visit: 15 test slides and 20 filler slides. The images on the test slides were selected according to the criteria described above while filler slides were used to allow subjects to acclimate to the

display protocol and to ensure that dysphoric images did not appear on every slide. Filler slides changed from session to session while the images on the test slides remained the same. The positions of the four images on each test slide were randomly changed between sessions. By using a large number of filler slides (different filler slides for each session) and a randomized order for the position of the four images on each test slide, the possibility that patients' scanning behavior was affected by recall from previous presentations was minimized. Each slide was presented for 10.5 seconds for a total presentation time of 8 minutes and 45 seconds. The presentation was shown and the visual scan paths recorded using VAST in the conditions described in Chapter 2.3.

4.2.5. Parameter Estimation and Analysis

Visual scanning parameters were generated using the automated system as described in Chapter 2.4. Factors such as the subject looking away from the screen, moving their head outside of the estimation range, or disruption by eyelashes or hair inhibited proper gaze estimation on occasion. Slides were removed and excluded from analysis where the number of valid estimates was less than 80% of the presentation time (8.4s). Visits with less than 50% of total test slides (< 8 slides) were not analyzed and subjects without enough slides at baseline (Visit 2) were removed.

4.3. Results

Of the 28 patients that have completed the study thus far, 22 remained after removal of subjects due to insufficient data. Twelve patients responded to the medication (had a HDRS-17 of half their initial value in their last visit) and 10 patients did not respond to the medication. Table 3.2 summarizes the total HDRS-17 scores over the 8 visits.

Subject ID	Age	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8
Non-Responders									
2D	**49**	20	16	*	*	*	*	*	*
4D	**57**	33	25	21	15	18	19	12	15
5D	**23**	N/A	27	25	29	24	*	20	16
6D	**47**	22	21	20	15	18	18	17	12
9D	**49**	28	25	21	23	28	31	31	31
14D	**34**	24	16	6	*	26	24	*	*
16D	**34**	N/A	21	22	11	18	20	*	20
17D	**55**	N/A	28	12	19	16	21	*	21
24D	**56**	N/A	23	25	28	20	*	*	*
64D	**42**	29	26	20	26	25	29	*	27
Responders									
8D	**56**	N/A	27	16	12	13	9	5	8
19D	**39**	N/A	29	27	16	*	12	11	8
20D	**62**	N/A	21	15	12	9	6	6	9
21D	**62**	N/A	25	9	7	4	6	5	9
1D	**43**	N/A	22	17	19	12	10	8	5
3D	**47**	N/A	21	16	5	3	2	1	1
7D	**23**	16	22	10	7	11	9	8	5
10D	**50**	N/A	21	14	10	8	7	8	6
22D	**53**	N/A	20	12	10	7	6	5	4
47D	**31**	N/A	23	*	17	14	18	6	5
54D	**46**	N/A	16	15	9	8	9	4	6
68D	**48**	N/A	21	9	11	12	8	3	*

Table 4.2: Total HDRS-17 scores for all participants diagnosed with Major Depressive Disorder over the 8 Visits. Note subjects without a Visit 1 did not require a washout period because they were not previously on any medication. Visits with an asterisk (*) indicate missed visits or patient drop-out.

For the analysis of the data, the Relative Time on Image of the two dysphoric images on each test slide were added together to produce a Relative Time on Dysphoric Images (RTDI) parameter. Since every test slide has two dysphoric and two social images, the Relative Time on Social Images (RTSI) are the compliment of RTDI such that RTSI + RTDI = 1. The mean RTDI of all slides recorded from responders and non-responders are plotted for each visit in Figure 4.3. A two-tailed equal variance Student's T-Test was performed between the RTDI for each visit. The resulting p-values, degrees of freedom, and T-statistic are shown in Table 4.3

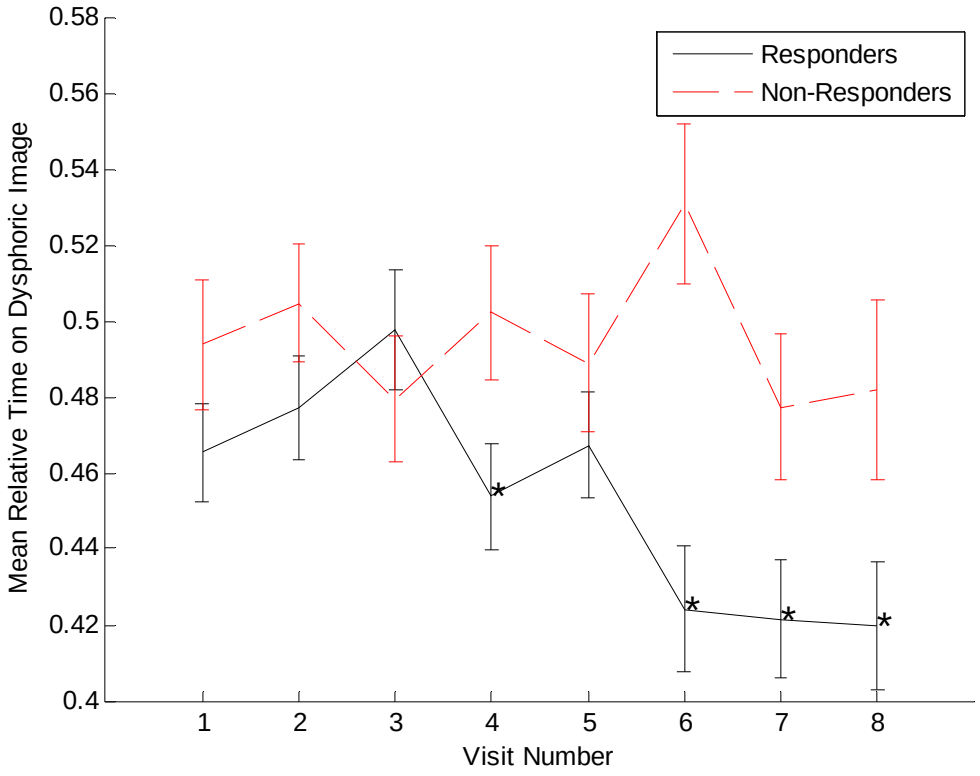


Figure 4.3: Mean Relative Time on Image by visit on Dysphoric IAPS Images grouped by Responders and Non-Responders to medication. Antidepressant treatment occurs immediately after visit 2. Error bars show standard error of the mean and asterisks (*) denote mean RTDI of responders that are significantly different (P<0.05) from mean RTDI of non-responders.

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	Visit 7	Visit 8
P	0.1738	0.1831	0.4271	0.0294	0.3411	0.0001	0.0439	0.0302
df	254	307	258	259	262	215	205	237
T Stat	-1.364	-1.3341	0.7954	-2.1904	-0.9537	-4.0537	-2.0279	-2.1801

Table 4.3: Two-tailed equal variances Student's T-Test with between Responders and Non-Responders for each visit.

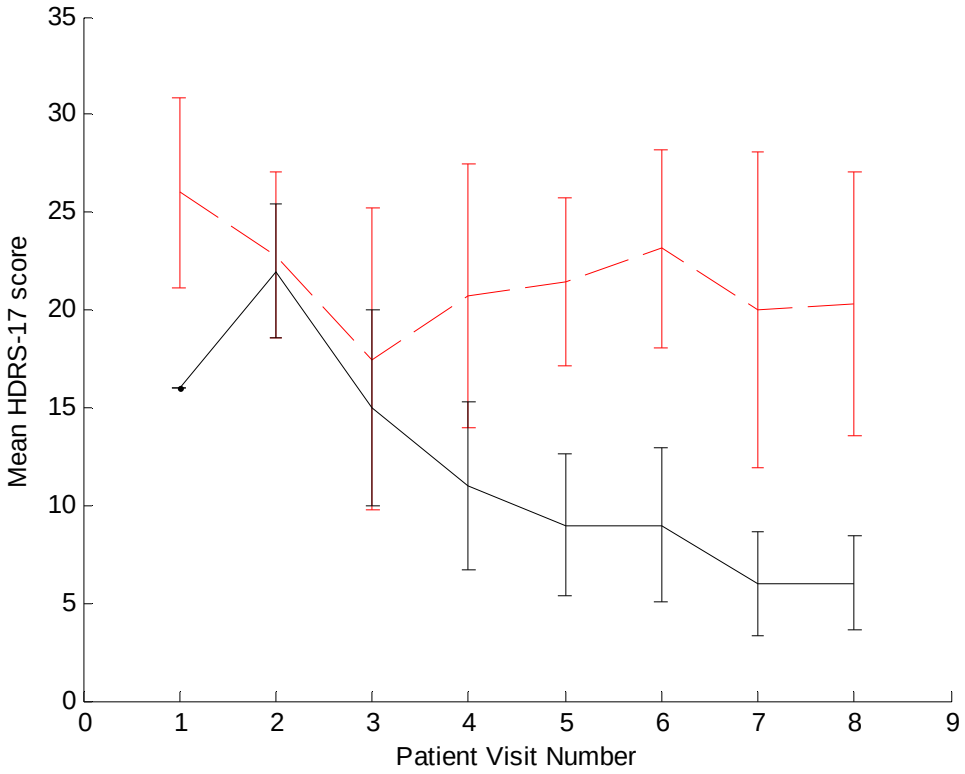


Figure 4.4: Mean HDRS-17 score for responders and non-responders over the 8 patient visits.

A significant difference ($p<0.05$) between responders and non-responders first occurs two weeks following the start of medication and at Visits 4, 6, 7 and 8. The delay in detecting differences between responders and non-responders through the HDRS-17 questionnaire is similar to the delay in detecting differences between the groups through the analysis of RTDI (Figure 4.3 and 4.4). The lack of a significant difference between the groups in Visit 5 may be an indication of relatively large inter-subject variability in a small sample size.

To assess individual responses of subjects, the mean RTDI were plotted for all subjects along with their total HDRS-17 scores by visit. In the group of 10 non-responders, 7 subjects had no change in RTDI over the 8 visits, one subject had an increase, and two had a decrease in at least one visit following Visit 2. Of the 12 responders, 4 subjects had a significant decrease in RTDI in at least one of the visits following Visit 2. Two responders had a significant increase in RTDI in at least one of the weeks following Visit 2 and the other 6 subjects had no significant change in RTDI over the 8 visits. Interestingly, two subjects, one responder and one non-responder, had

significant decreases in RTDI in visits following Visit 2. In the following paragraphs, individual patients, one from each of the above categories of responses, are discussed.

4.3.1 Non-Responders

Subject 9D, whose HAMD scores and the relative time on dysphoric images are shown in Figure 4.5, is representative of the data from the majority of the non-responders (7 out of 10) where no change was observed in RTDI over the 8 visits. An absence of change when no medication response is observed on the HDRS-17 is consistent with the hypothesis. Within the non-responders groups (10 patients) there were three exceptions: Subjects 5D, 4D and 16D. Figure 4.6 shows that Subject 5D had an HDRS-17 score of 28 at baseline and after Visit 4, has a continuous decrease in HDRS-17 score until in Visit 8 a score of 16 was observed. While this subject is a non-responder (according to the criteria that his/her score in Visit 8 should be lower than 50% of baseline) the HDRS-17 at Visit 8 is only 2 points higher than what is required to be classified as a responder. The visual scanning data for subject 5D shows significant decreases in RTDI following the start of medication. The data for Subject 4D are similar to that of subject 5D. Following medication, there is a continuous decrease in HDRS-17 and in Visit 8 the score is only 3 points higher than what is required to be classified as a responder. These changes in HDRS-17 scores are accompanied by a significant decrease in RTDI. It is possible that for these patients the eventual symptom alleviation may not have been tracked within the scope of this study. Subject 16D, is the last of the three non-responding patients that showed a significant change in visual scanning behaviour. Figure 4.7 shows a continuous gradual increase from baseline RTDI that become significant in Visit 6. According to the HDRS-17, the patient had an initial positive response to the treatment, but this initial positive response was reversed in visits 5 to 8.

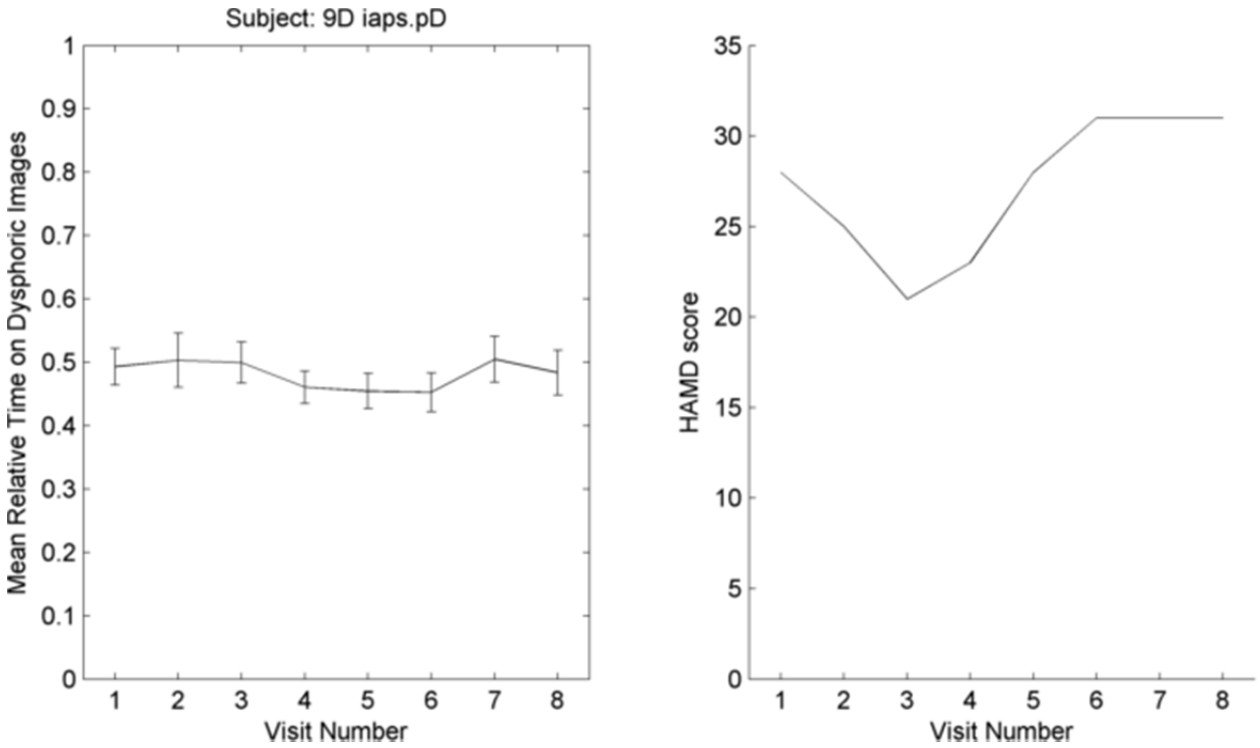


Figure 4.5: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for subject 9D.

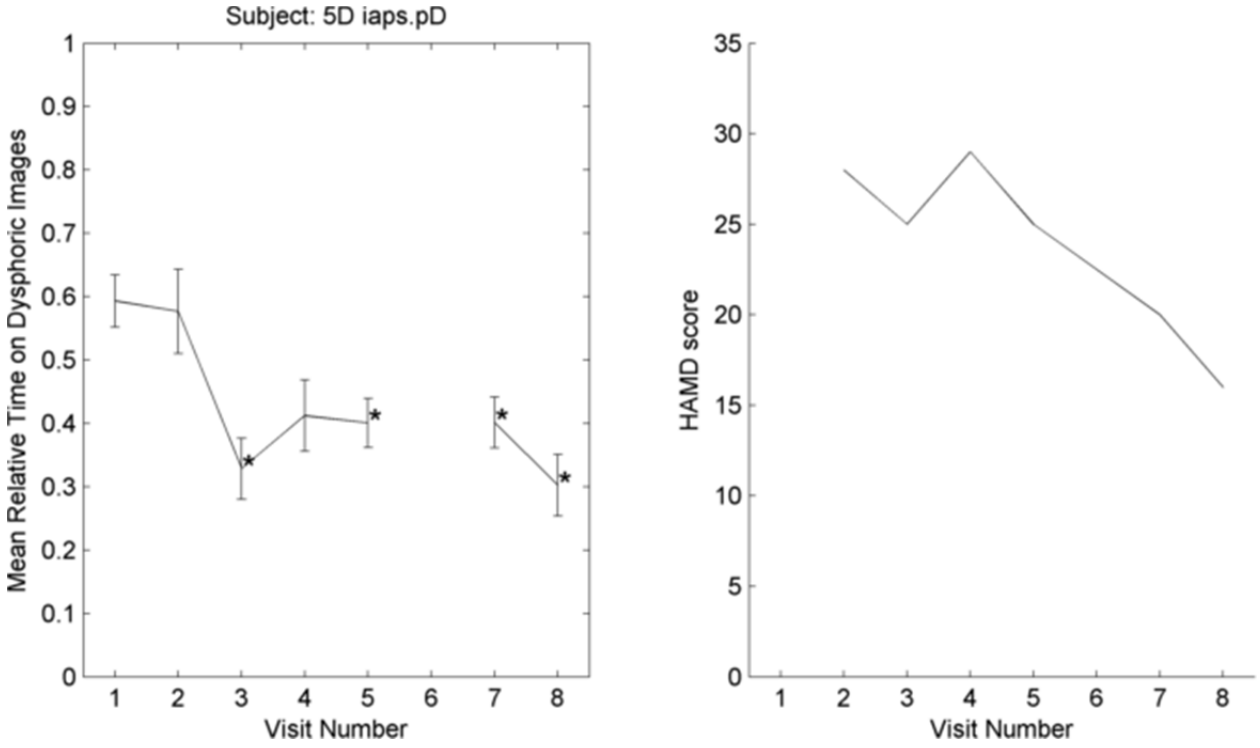


Figure 4.6: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for subject 5D. A subject with decreased RTDI while being a non-responder.

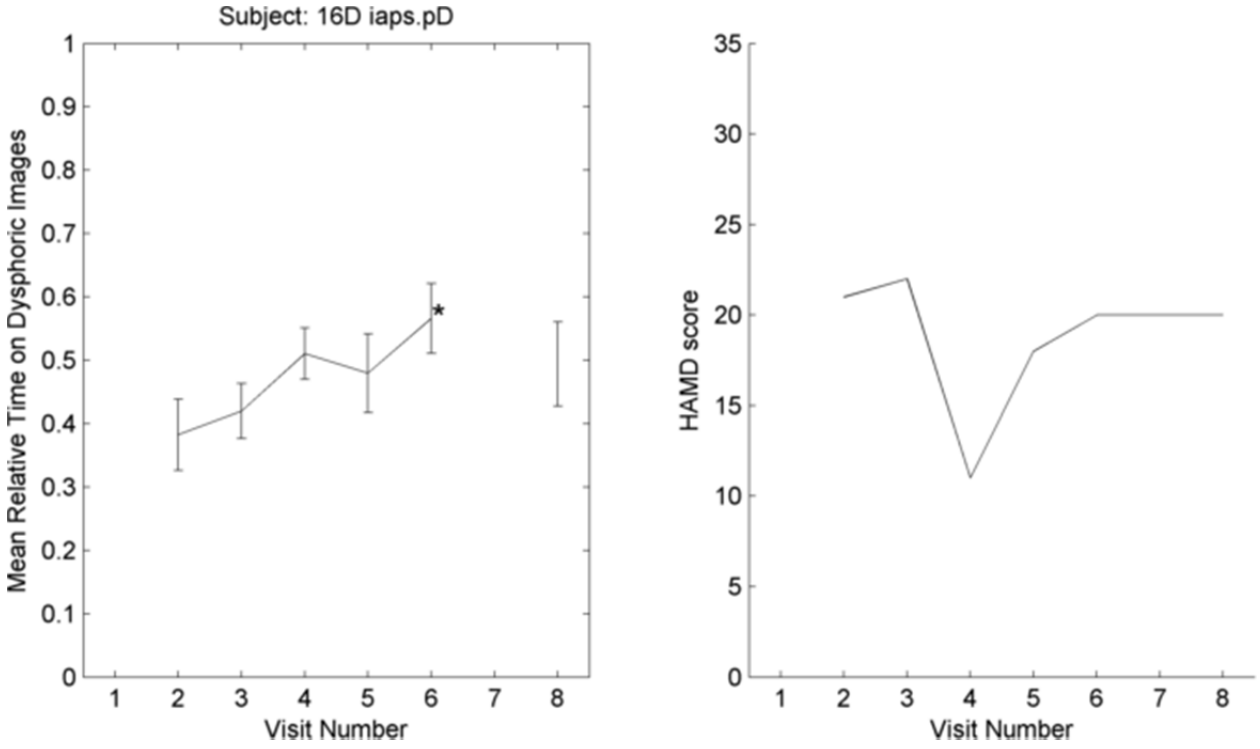


Figure 4.7: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for subject 16D. A subject with increased RTDI while being a non-responder.

4.3.2 Responders

Subject 20D, whose HDRS-17 scores and relative time on dysphoric images are shown in Figure 4.8, is a responder who exhibits a significant change in RTDI after starting Duloxetine Monotherapy (i.e., after Visit 2). This patient demonstrates a shift away from dysphoric images as his/her clinical condition, as measured by HAMD, improves. Unfortunately, the data collected at visit 3 was too poor to be analyzed and therefore it is uncertain whether the change in visual scanning behaviour occurred earlier. Interestingly, Subject 20D's RTDI at Visit 2 is 0.4037 ± 0.0339 which is less than the mean RTDI for all other patients (0.4844 ± 0.0102. An RTDI of 0.5 is expected if the subject has no bias towards either dysphoric or social images. Thus, Subject 20D demonstrated positive bias towards social images at baseline (i.e. before medication) and this bias increased as the patient responded to medication. The data of subject 20D suggests that each subject has his/her own criteria for interpreting low and high valence images. This leads to large variability in baseline values of VSB parameters, for similar scores on the HDRS-17 questionnaire.

Figure 4.9 shows the HDRS-17 scores and the relative time on dysphoric images for Subject 1D who is a responder with no observable change in RTDI over the 8 weeks. For this subject, changes in clinical symptoms as measured by HDRS-17 are not expressed in RTDI. With a mean RTDI approximately equal to 0.5 every week, this subject views both dysphoric and social images an equal amount of time. This result can be interpreted as either the subject did not have a bias or the bias was not observable through the RTDI parameter.

Figure 4.10 shows the HAMD scores and the relative time on dysphoric images for Subject 19D who demonstrates an increase in RTDI when responding to medication in three of the visits after starting Duloxetine Monotherapy. The change in visual scanning behaviour is in the opposite direction to the one postulated by our hypothesis. With a baseline (Visit 2) RTDI of 0.2841 ± 0.0298, which is the lowest of all responders, the data from Subject 19D suggest that avoiding dysphoric images during depressive states might characterize the visual scanning bias in this patient. This possibility is discussed below.

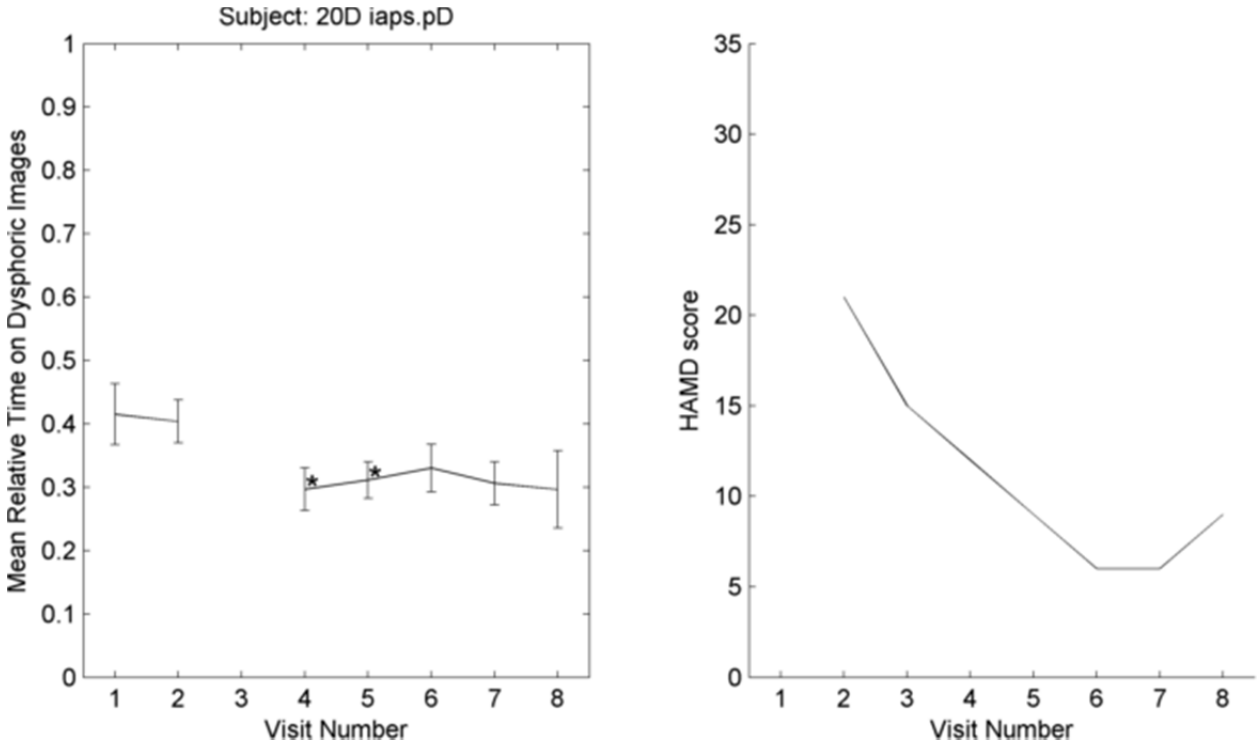


Figure 4.8: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for example subject 20D. An asterisks (*) denotes that the mean is significantly different ($P<0.05$) from Visit 2.

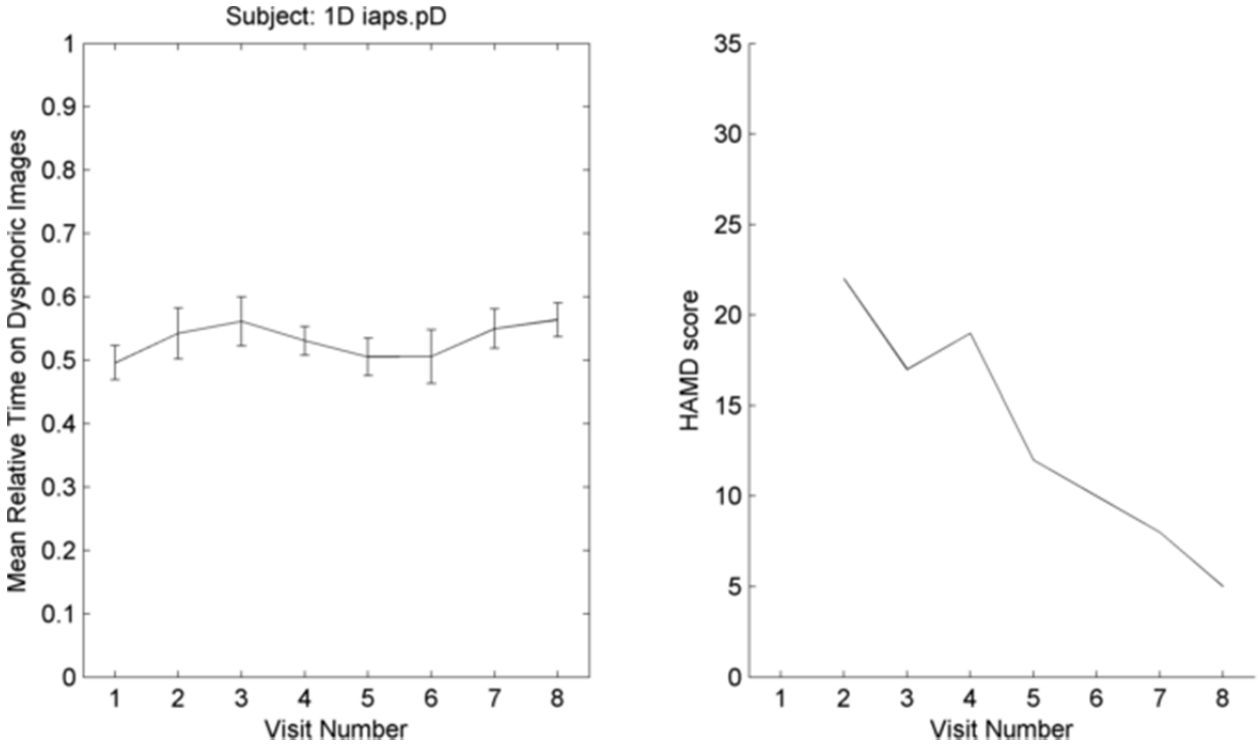


Figure 4.9: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for subject 1D.

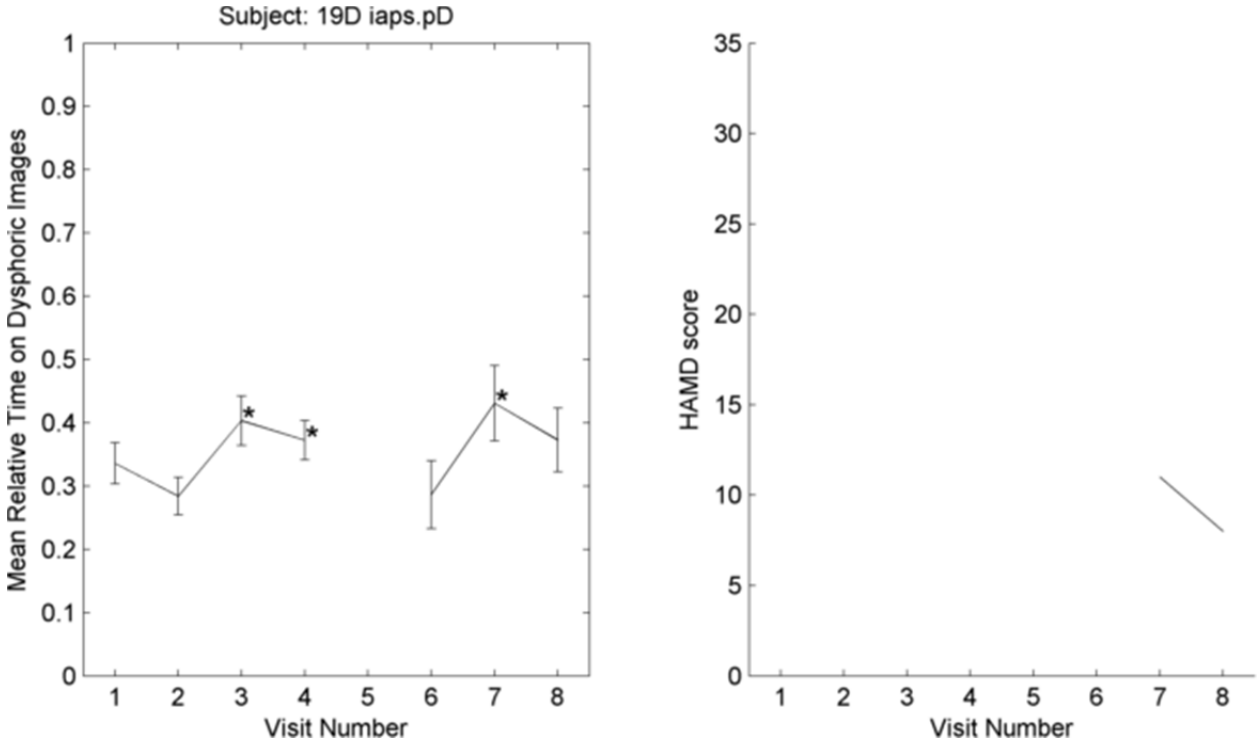


Figure 4.10: A) Mean Relative Time on Image on dysphoric images of IAPS slides and B) Total HDRS-17 score by patient visit number for subject 19D. An asterisks (*) denotes that the mean is significantly different ($P<0.05$) from Visit 2.

RTDI is a relatively stable parameter that incorporates visual scanning behaviour over the entire duration of a slide. As a result, it might suffers from a lack of sensitivity and specificity. In the development of the visual scanning paradigm in 2003, Eizenman et al. showed a significant difference between control and dysphoric subjects using the Relative Time on Image parameter. Given the large inter-subject variability in this study, it is evident that a more sensitive visual scanning parameter is required if the biases are to be measured in individuals and used to predict medication response.

4.4. Discussion

As a population, shifts in visual attention away from dysphoric stimuli after antidepressant treatment characterize responders from non-responders. However, large inter-subject variability is observed in this study. For example, the range of mean Relative Fixation Time at Visit 2 (baseline) is [0.2738,0.8009]. Evidently, the attention bias observed in the population of depressed subjects (Eizenman et al., 2003; Caseras et al., 2007; Kellough et al., 2008; Sears et al., 2011) is not representative of the behaviour of a depressed individual. These differences between individuals may be understood in terms of individual positivity offset and negativity bias. In many physiological systems there are different response curves for negative and positive stimuli (low valence and high valence). Positive stimuli usually lead to stronger responses at low activation (low arousal) while negative stimuli lead to stronger activation at high activation (high arousal). The differences between the positive and negative response curves at low and high activations are called positivity offset and negativity bias, respectively. Ito and Cacioppo (2005) found support for both a positivity offset and negativity bias using visual stimuli selected from IAPS. Moreover, they noticed large variations in individual responses (the two parameters were normally distributed), which suggests that when viewing visual stimuli each person has its own unique set-point. The set point can be defined as a point along the activation axis (arousal) for which the response to positive stimuli is equal to the response to a negative stimuli. For an individual, visual stimuli with arousal values that are lower than his/her set-point will result in positive response bias (i.e, the individual will spend more time looking at high valence stimuli), while visual stimuli with arousal values that are higher than his/her set-point will result in negative response bias. As the set-point for control subjects can vary significantly (one can obtain some idea about these changes from Ito and Cacioppo, 2005) relative to the limited range of arousal values that we use in our study it is possible that control subjects will generate both

positive and negative biases in VSB (i.e., the observed behaviour). It is likely that as a group, the set-point for patients with depression is lower than that of control subjects, and therefore as a group we can detect a negative bias. However, the overlap between the ranges of the set-points for the two groups makes it difficult to separate between the responses of individual controls and patients. In the context of treatment, individual set-points might be used as an indication of response. As individuals respond to antidepressants, shifts in the set-point may break the maladaptive patterns that sustain the disorder. Those whose set-points are not affected by medication or shift in a manner that sustain the disorder do not respond to medication. Therefore, it appears crucial to consider arousal in the design of the visual stimuli. For example, the set-point may be estimated per person by varying the arousal between images.

Another mechanism that may explain some of the variability observed between individuals is avoidance. Models of approach and avoidance suggest that avoidance functions to inhibit of ongoing behaviour that produces negative outcomes and assists in resolving conflicting goals. However, it may also play a role in limiting access to positive reinforcement which contributes to the onset and maintenance of depression (Trew et al., 2011). While avoidance leads to short term relief, depressed individuals who avoid do not move closer to solving the problems that have led to the depression and may exacerbate unresolved problems (Ottenbreit & Dobson, 2004). Avoidance has been observed in visual scanning studies. In a study by Hermans, Vansteenwegen, and Eelen (1999), control participants showed a stable pattern of eye gaze over the time course while spider anxious participants spent significantly longer looking at the spider picture than the flower at the beginning of visual presentation. This pattern, however, gradually shifted away from the spiders with more time (Hermans, Vansteenwegen, & Eelen, 1999). These results support a "Vigilance-Avoidance Model", which has also been observed in a separate study of spider fearful subjects (Rinck & Becker, 2006). In another study of children with separation anxiety disorder, anxious children spent significantly more time gazing at separating pictures than non-anxious children at the beginning of visual presentation but showed a reversed pattern after approximately 3 seconds (In-Albon, Kossowsky, & Schneider, 2010). These studies show that avoidance is another possible visual scanning response when depressed individuals are presented with dysphoric stimuli. Following this hypothesis, Subject 19D, who had the lowest mean RTDI of all participants, may be explained as a patient who avoided dysphoric stimuli

while in a depressive state (Visit 2) but was able to shift attention towards dysphoric stimuli after responding to medication. In light of these observations, it may be appropriate to reconsider the hypothesis that depression is characterized by shifts in attention towards dysphoric stimuli (rumination). A model of depressed visual scanning behaviour that includes avoidance as a mechanism may explain some of the large inter-subject variability.

4.5. Conclusions

A significant shift in visual attention away from dysphoric stimuli occurs as early as 2 weeks following Duloxetine monotherapy in the population of 13 responders. This shift is not present in the population of 10 non-responders. However, large inter-subject variability prevents Relative Time on Dysphoric Images (RTDI) from being used as an early indicator of response to antidepressants. Future research into how the arousal of the visual stimuli and avoidance mechanisms in depression affect VSB will provide understanding into the inter-subject variability and subsequently the development of early and objective indicators of drug efficacy.

Chapter 5. Detection of Apathy in Alzheimer Patients

5.1. Introduction

Biases in visual attention, as measured by eye-tracking technology, might be used to differentiate Alzheimer's disease (AD) patients with depression from those with apathy. In patients without AD, a bias in visual attention towards dysphoric images has been associated with depression (Olin et al., 2002), while apathetic patients with AD may lack responsiveness (attentional bias) towards emotional stimuli (Robert et al., 2009; Drago et al., 2010). The goal is to use visual scanning behaviours (VSBs) to identify patients with apathy from those with depression. To provide context, these visual attention measures are compared between AD patients and those of healthy elderly controls with neither apathy nor depression. Using an objective, readily quantifiable technique to distinguish apathy from depression might facilitate differential diagnoses and choices between pharmacological treatments of these very common and often indistinguishable disturbances in patients with AD.

Based on previous studies (Eizenman et al., 2003; Daffner, Mesulam, and Scinto, 1999), it is hypothesized that AD participants with depression will show different biases in visual attention than AD participants with apathy. Specifically:

1) When presented with visual stimuli that include emotional content (both dysphoric and social) and emotionless content (neutral), non-apathetic participants (including depressed AD patients) will display an attentional bias towards emotional images whereas apathetic AD participants will display less or no bias.
2) When presented with visual stimuli that include dysphoric, social and neutral images, depressed patients will display an attentional bias towards dysphoric images (i.e. increased relative fixation time) whereas non-depressed AD patients (apathetic or not) will not display such a bias.

5.2. Study Design

5.2.1. Participants

Participants diagnosed with AD were considered for inclusion for each of the three groups: apathetic, depressed and control. Additionally, elderly non-AD participants within the age range

of the AD participants were included as a control. Within each group, the spectrum of disease severity was represented (Mini-Mental Status Examination (MMSE) scores ranging from 10-30). AD patients were screened for apathy using the Neuropsychiatric Inventory (NPI) apathy subscale, and depression was diagnosed based on Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV-TR) criteria for major depressive episode (MDE) and the NPI depression subscale (see Table 5.1). The Modified Mini Screen (MMS) was administered to healthy controls to exclude those with neuropsychiatric disturbances. Study candidates were recruited from the Psychiatry Department at Sunnybrook Health Sciences Centre as well as from outside referrals.

5.2.3. Visual Stimuli

The visual stimuli include a series of slides each containing four images presented on a computer screen. Images were selected from the standardized International Affective Picture System (IAPS), classified as neutral, dysphoric or social. The images in each slide series were chosen based on the IAPS 9-point rating scale for valence (feelings of displeasure vs. pleasure, with higher values indicating more pleasure) and arousal (feelings of calm vs. excitement, with higher values indicating more excitement). The IAPS numerical values for these quantities have been assigned through tests on normal populations (Lang et al., 1999). Neutral images have an approximate valence of 5 and low arousal, social images ranged from 6 to 8 while dysphoric images have valence ratings of 2 to 4. Images were organized into slides of 2 types: "Emotion" (1 dysphoric, 1 social and 2 neutral images show in Figure 5.1) and "Emotionless" (4 neutral images shown in Figure 5.2). Sixteen slides of each type for a total of 32 were presented. The four images on each slide are arranged 2 by 2, with each category of image appearing in each specific spatial position equal number of times (equally distributed) throughout the 16 slides.

The two slide types were used to compare the visual scanning behaviour when neutral images compete for attention with emotional images and when neutral images compete for attention with other neutral images. It is expected that non-apathetic participants will scan the Emotion slides differently than the Emotionless slides while apathetic participants will scan the Emotion and Emotionless slides similarly. Depressed participants are expected to have additional attention bias towards the dysphoric image.

Figure 5.1: Example of an Emotion slide where 1 dysphoric (bottom-right), 1 social (bottom-left), and 2 neutral (top-left and top-right) images are shown together.

Figure 5.2: Example of an "Emotionless" slide where all 4 images are neutral.

5.2.2. Procedure

Informed consent was obtained for all participants. Patients were assessed for inclusion and exclusion criteria and the MMS was administered to healthy controls to screen for psychiatric disorders. Diagnostic criteria for depression (see Table 5.1) and the NPI apathy subscale (≥4) were used to dichotomize AD subjects and to screen for other behavioural disturbances. Eligible participants then had their visual scanning patterns and pupil-sizes recorded and analyzed by EL-MAR's Visual Attention Scanning Technology (VAST, EL-MAR Inc., Toronto, Ontario, Canada) as described in Chapter 2. Briefly, following a short calibration procedure, subjects viewed 5 filler slides to become familiar with the procedure. The study slides were then presented sequentially, with a 10.5-second viewing time allocated to each with a 2 second pause in between slides. The test time for each subject was divided into 2 sessions of approximately 7 minutes each (16 slides per session plus 10 filler slides). In between the two sessions the subject had a 5-minute break. To measure severity of target symptoms, the Apathy Evaluation Scale (AES) and Cornell Scale for Depression in Dementia (CSDD) were administered. Other assessments of overall attention (Wechsler Adult Intelligence Scale - Digit Span and Conners' CPT), severity of dementia (MMSE and Global Deterioration Scale) were administered as possible covariates. Patient demographics including age, gender, medical history and concomitant medications were recorded as possible covariates.

Factors such as the subject looking away from the screen, moving their head outside of the estimation range, or disruption by eyelashes or hair inhibited proper gaze estimation on occasion. Slides were removed and excluded from analysis where the number of valid estimates was less than 80% of the presentation time (8.4s). Subjects with less than 50% of total test slides (< 9 slides in Emotion or Emotionless slides) were removed.

Inclusion (non-AD Controls)	*Exclusion (All)*
• No current diagnosis of AD or dementia and no history of functional impairment (MMSE≥26) • A minimum of 50 years of age (within age range of AD participants) • No indication of psychological disturbances (MMS<6)	• clinically significant behavioural disturbances on NPI (e.g. agitation/aggression, delusions, hallucinations) in which separate domain scores are ≥4 • change in anti-dementia medications less than 1 month prior to study day (AD groups only) • significant eye pathology (e.g. cataracts), visual acuity problems and communicative impairments • on medication for depression or apathy at screening and/or current use of psychotropic medications (except cognitive enhancers for AD) • current severe medical conditions, including uncontrolled hypertension, diabetes and stroke • presence or history of other psychiatric disorders or neurological illnesses or traumatic brain injury
Inclusion (AD Controls)	
• Diagnosis of possible or probable AD based on DSM-IV-TR5 and NINCDS-ADRDA criteria6 • No clinically significant apathy or depression (NPI<4 in each subscale)	
Inclusion (AD Apathy)	
• Diagnosis of possible or probable AD based on DSM-IV-TR5 and NINCDS-ADRDA criteria6 • Apathy on NPI is "very frequent" or "frequent/often" and severity is "moderate" or "marked" or NPI apathy subscore≥4 • No clinically significant depression based on DSM-IV-TR criteria	
Inclusion (AD Depression)	
• Diagnosis of possible or probable AD based on DSM-IV-TR5 and NINCDS-ADRDA criteria6 • Diagnosis of depression according to DSM-IV-TR criteria and NPI ≥4 for depression domain	

Table 5.1: Participant Eligibility Criteria

5.3. Preliminary Results and Discussion

5.3.1. Objective Measure of Apathy

Sixteen participants have completed the study thus far: 2 apathetic Alzheimer's patients, 2 depressed Alzheimer's patients, 6 Alzheimer's patients, and 6 age-matched controls. This small sample size prevents population behaviour to be analyzed. However, a preliminary discussion about the behaviour of individuals within the four groups of participants can be carried out.

It is hypothesized that non-apathetic participants, including depressed patients, will exhibit attention bias towards emotional images. A simple way to measure this attention bias is to compare the total time spent on emotion images (social and dyspohric) with the total time spent on emotionless image (neutral images). The Relative Fixation Time on the social image and the dysphoric image were added together to produce Relative Time on Emotion Images (RTEI). The Relative Fixation Time on the two neutral images per slide were added together to produce Relative Time on Neutral Images (RTNI) such that RTEI + RTNI = 1. The mean RTEI of all 16 Emotion slides are shown in Figure 5.3 for each of the 16 participants. All participants had a mean RTEI greater than 0.5 indicating that they all spent more time viewing emotion images in comparison to neutral images. This observation is likely a result of difference in the amount of detail between images on the Emotion slides. In the example of the Emotion slide shown in Figure 5.1, the social and dysphoric images have a greater number of details in the image than the neutral images. Therefore, it requires extra fixations and time to view emotion images in comparison to the neutral images. Despite this lack of normalization in the Emotion slides, hypothesis 1 predicts that the two apathetic patients should have the lowest RTEI of all the participants. Contrary to this hypothesis, 8 non-apathetic participants are shown to have lower RTEI than subject 1 AAD (the apathetic patient with the higher RTEI) in Figure 5.3.

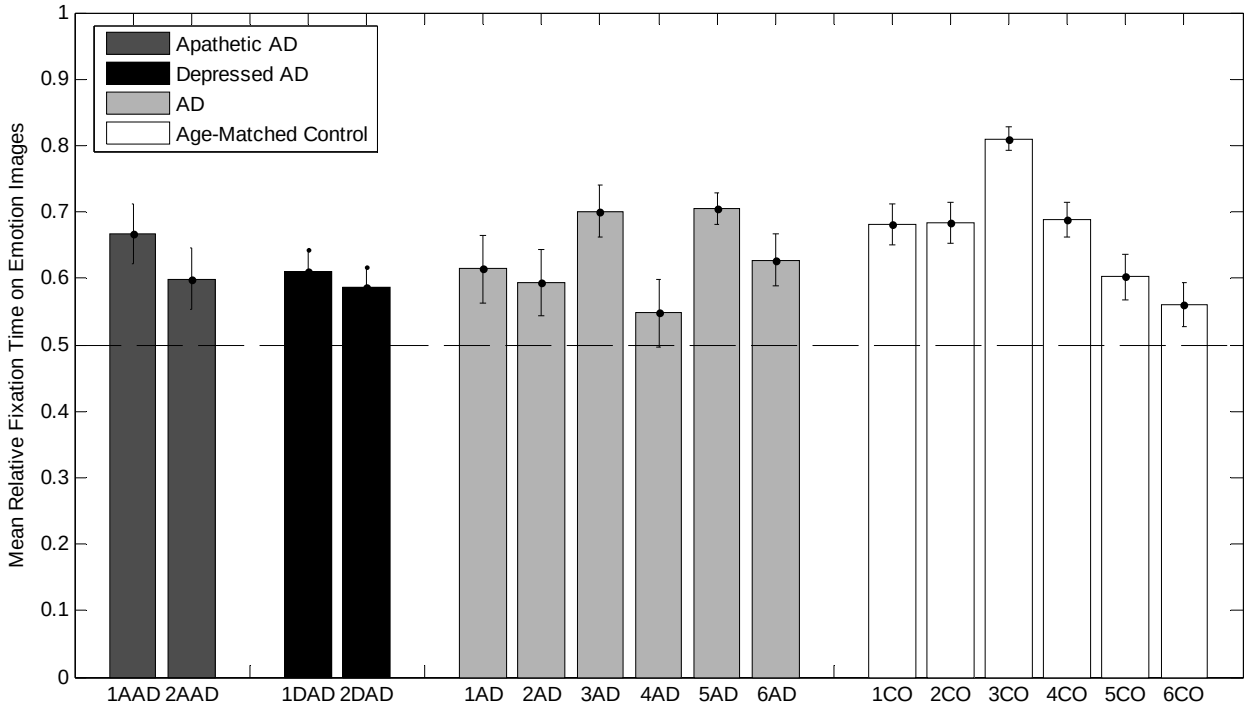


Figure 5.3: Mean Relative Fixation Time on Emotion Images for individual subjects of each group: apathetic Alzheimer's patients (AAD), depressed Alzheimer's patients (DAD), Alzheimer patients without apathy or depression (AD), age-matched controls (CO). Error bars show standard error of the mean.

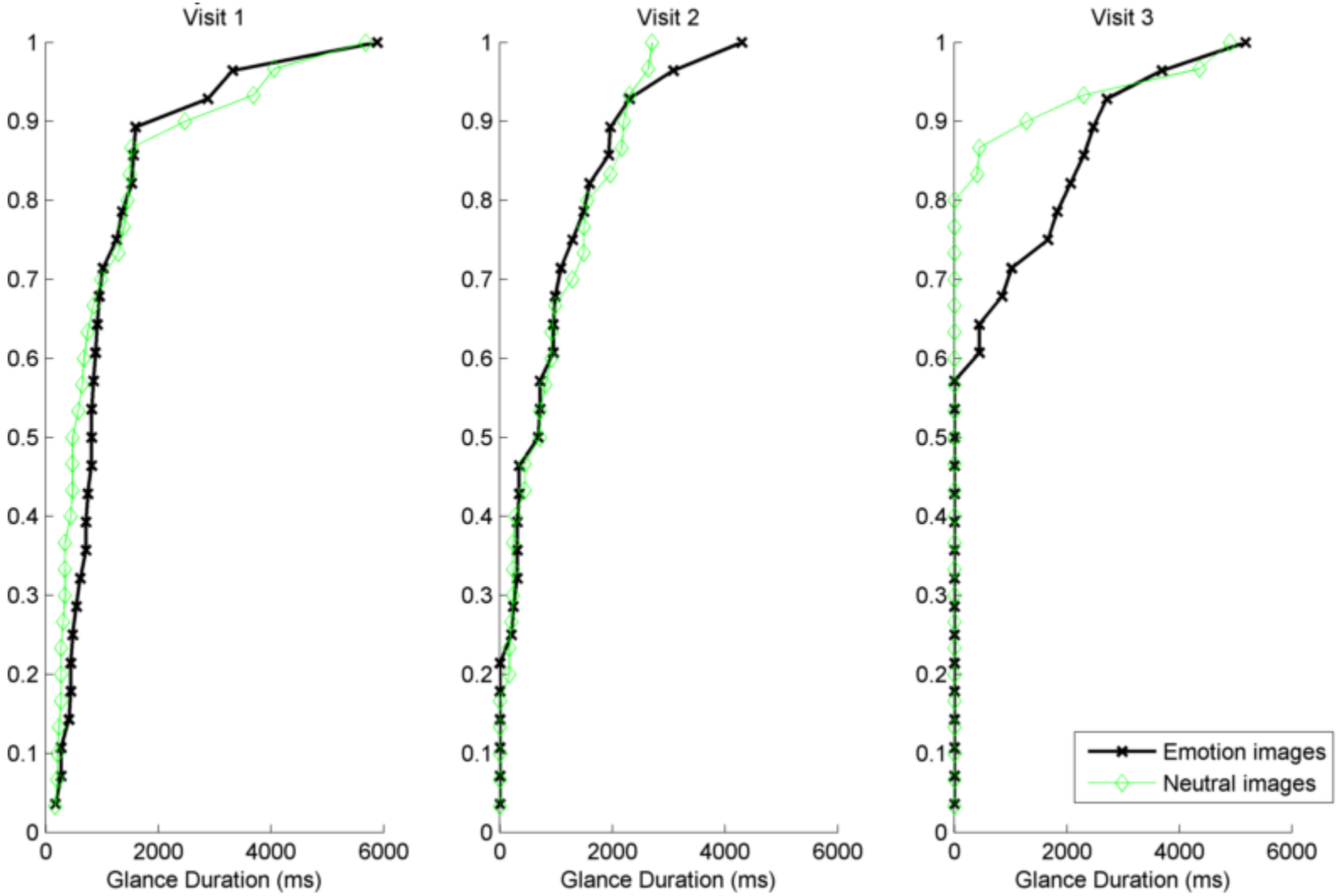
The time-sequence of visual scanning from the participants were assessed by analyzing Glance Duration by Visits. As described in Chapter 2, this visual scanning parameter measures the amount of time spent on an image each time a person visits the image. Figures 5.4A-5.4D show Glance Duration on Visit 1 (GD1), Glance Duration on Visit 2 (GD2), and Glance Duration on Visit 3 (GD3) for four selected subjects as cumulative distribution functions (CDFs). CDFs were chosen to represent the distributions to bypass the need to choose probability distribution function parameters (e.g. number of bins in a histogram). The CDFs of glance durations on emotion images (both social and dysphoric) on every Emotion slide and two randomly selected neutral images on every Emotionless slide are shown together in the same graph. Deviations between the CDFs of the two image types suggest that they view the image types differently. A rightward shift in CDF means that the subject has a larger number of longer glance duration on that image type while a shift in CDF towards the left indicates shorter glances durations. Comparing Figure 5.4C with Figure 5.4D, a similar observation may be made as the one described in Chapter 2 with the two Anorexic subjects. Figure 5.4C shows that approximately 90% of GD1 are shorter than 1000ms for subject 2AD (on either image type). In Figure 5.4D,

Subject 5CO only has 30% of GD1 shorter than 1000ms which suggests that Subject 5CO spends more time viewing the images on the first visit than 2AD. These differences in scanning are not likely to be indicative of apathy as both apathetic and non-apathetic participants can demonstrate "faster" (less time spent on visit 1) and "slower" (more time spent in visit 1) scanning. However, it does reveal that comparison of glance durations between individuals require normalization. For example, a GD1 of 2000ms is unusually long for subject 2AD but within the normal range for subject 5CO. Normalizing glance durations may be achieved by comparing distributions of different image types for the same individual.

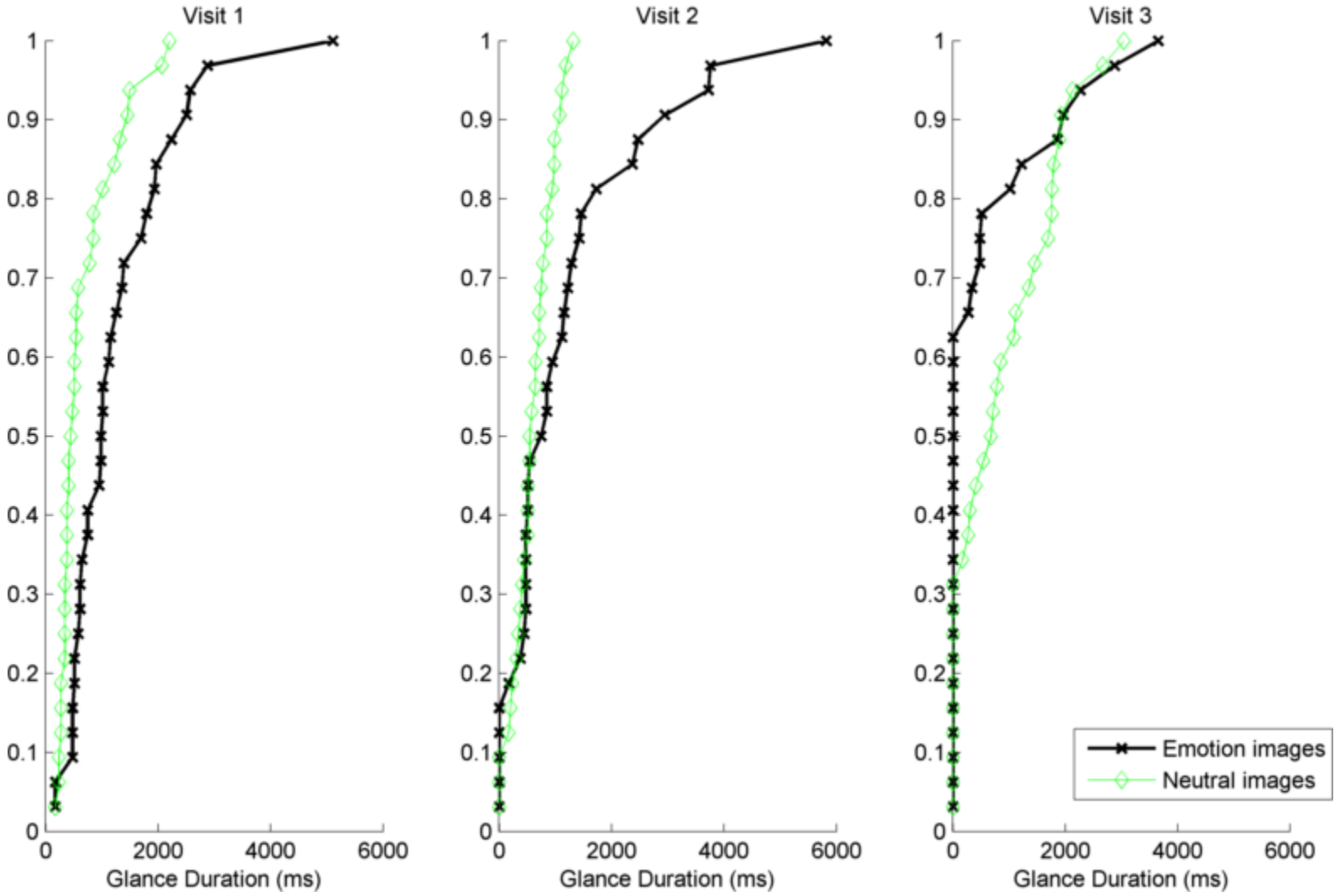
Considering Figures 5.4A – 5.4D, GD1 and GD2 appear to provide a good measure of apathy within an individual. Apathetic Subject 1AAD had a fairly high mean RTEI in Figure 5.3. In Figure 5.4A, the CDFs of glance durations on emotion images and neutral images completely overlap in both GD1 and GD2 suggesting that Subject 1AAD did not differentiate between the two image types. In contrast, the other three subjects shown in Figures 5.4B-5.4D (1DAD, 2AD, and 5CO respectively) had a smaller mean RTEI than Subject 1AAD but have differences between emotion and neutral images in GD1 and GD2. In Figure 5.4B, deviations in CDFs are observed between emotion images and neutral images in both GD1 and GD2. This observation suggests that similarities in GD1 and GD2 between emotion and non-emotional images may be exclusive to apathetic patients. Subject 2AD and subject 5CO are shown in Figure 5.4C and Figure 5.5D respectively. Although these two subjects have different approaches to scanning the images ("fast" vs. "slow"), they similarly have attention bias towards emotion images during the second visit but not during the first visit.

The glance durations on the third visit to images show a different story than in GD1 and GD2. Figures 5.4B and 5.4C show that the depressed Alzheimer's and non-depressed Alzheimer's patients have longer glance durations on neutral images than on emotion images. This difference demonstrates correlation between visits. Both Subject 1DAD and 2AD spent longer durations on emotion images in GD1 and GD2 and thus had less time to view images in visit 3 in comparison to the neutral images on the Emotionless slides. Likely, some glance duration estimates were shortened from the slides switching after 10.5 seconds of presentation. For example, if the average glance duration on a single image in each visit is 1 second, visiting all four images twice

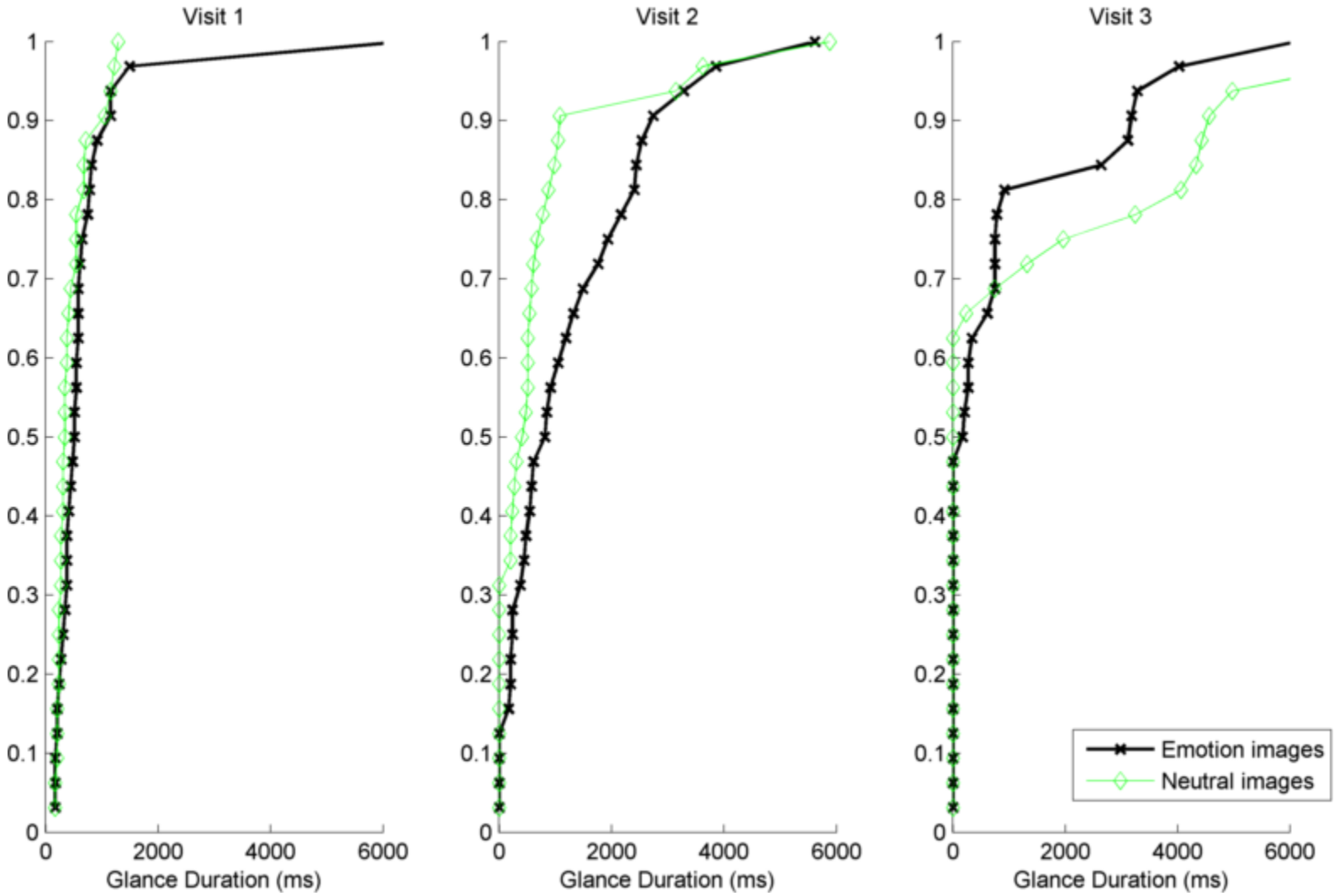
would leave 2 seconds for the third visit to the images. This phenomenon can be observed by the point where the CDFs cross the y-axis. For the "slow" scanner Subject 5CO (Figure 5.4D), only 20% of images are visited a third time regardless of image type. In contrast, the "fast" scanner Subject 2AD (Figure 5.4C) visits 38% of neutral images and 52% of emotion images a third time. Although Subject 1AAD appears to show differentiation between emotion and neutral images in GD3, the lack of observations arguably make this difference less significant than if it were to occur in GD1 and GD2.



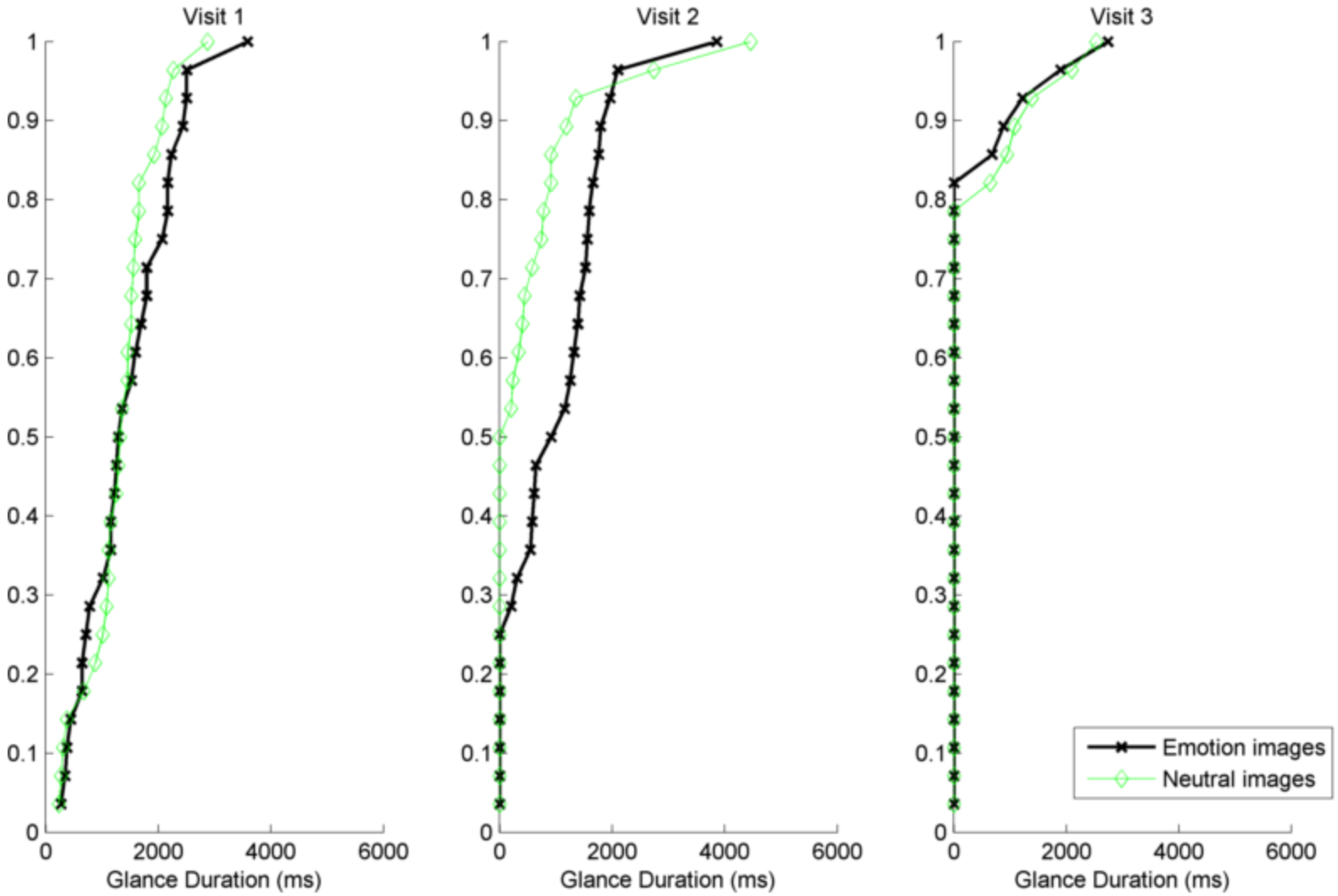
Figures 5.4A: The cumulative distribution function of glance durations by visit 1, 2, and 3 are shown for an apathetic Alzheimer's patient (1AAD). The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.



Figures 5.4B: The cumulative distribution function of glance durations by visit 1, 2, and 3 are shown for a depressed Alzheimer's patient (1DAD). The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.



Figures 5.4C: The cumulative distribution function of glance durations by visit 1, 2, and 3 are shown for an Alzheimer's patient without apathy or depression (2AD). The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.



Figures 5.4D: The cumulative distribution function of glance durations by visit 1, 2, and 3 are shown for an age-matched control (5CO). The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.

Comparing glance durations between slides is a way to keep the observations of glance durations independent between emotion and neutral images. Observations on the same slide are negatively correlated. Spending more time on one image, decreases the amount of time you can spend on the other images. This may be observed clearly in Figure 5.5 which plots the CDF of neutral images on the Emotion slides in addition to the other two image types plotted previously in Figures 5.4A-D. The CDF of neutral images on the Emotion slides does not overlap with the CDF of the neutral images on the Emotionless slides because the increased glance duration on the emotion images of the Emotion slides leaves less time to view the neutral images. By comparing the two emotion images on the Emotion slides to the two randomly selected neutral images on the Emotionless slides, a comparison is made between the behavior exhibited when emotion images compete for attention against emotionless images and the behavior exhibited when emotionless images compete for attention between themselves. This comparison is an improved measure of attention bias over RTEI because information about the natural scanning pattern of an individual (when viewing emotionless images) is considered.

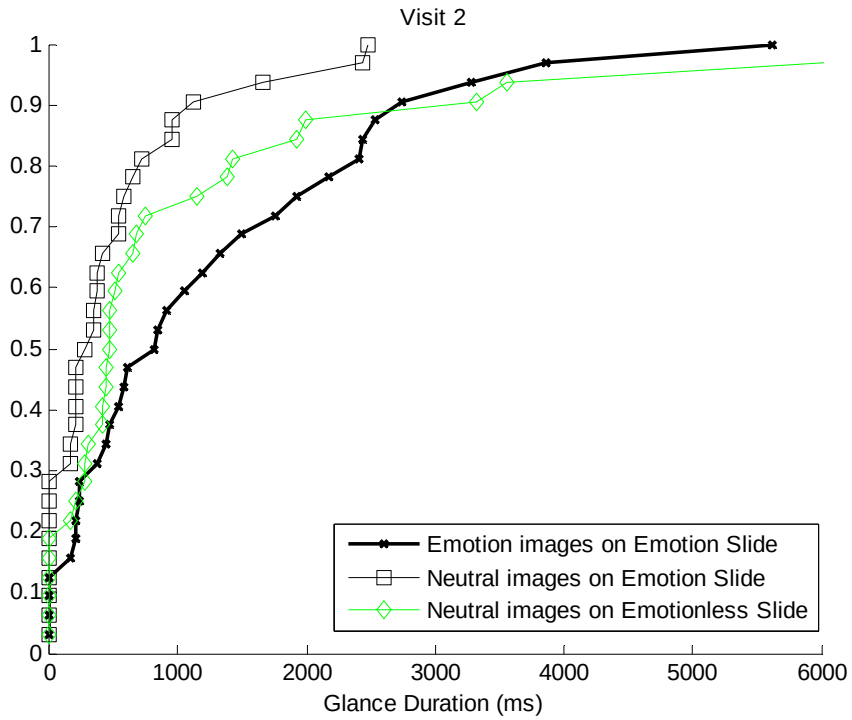


Figure 5.5: The cumulative distribution function of glance durations visit 2 of subject 2AD. The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. The dark line with square points represent glance durations on neutral images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.

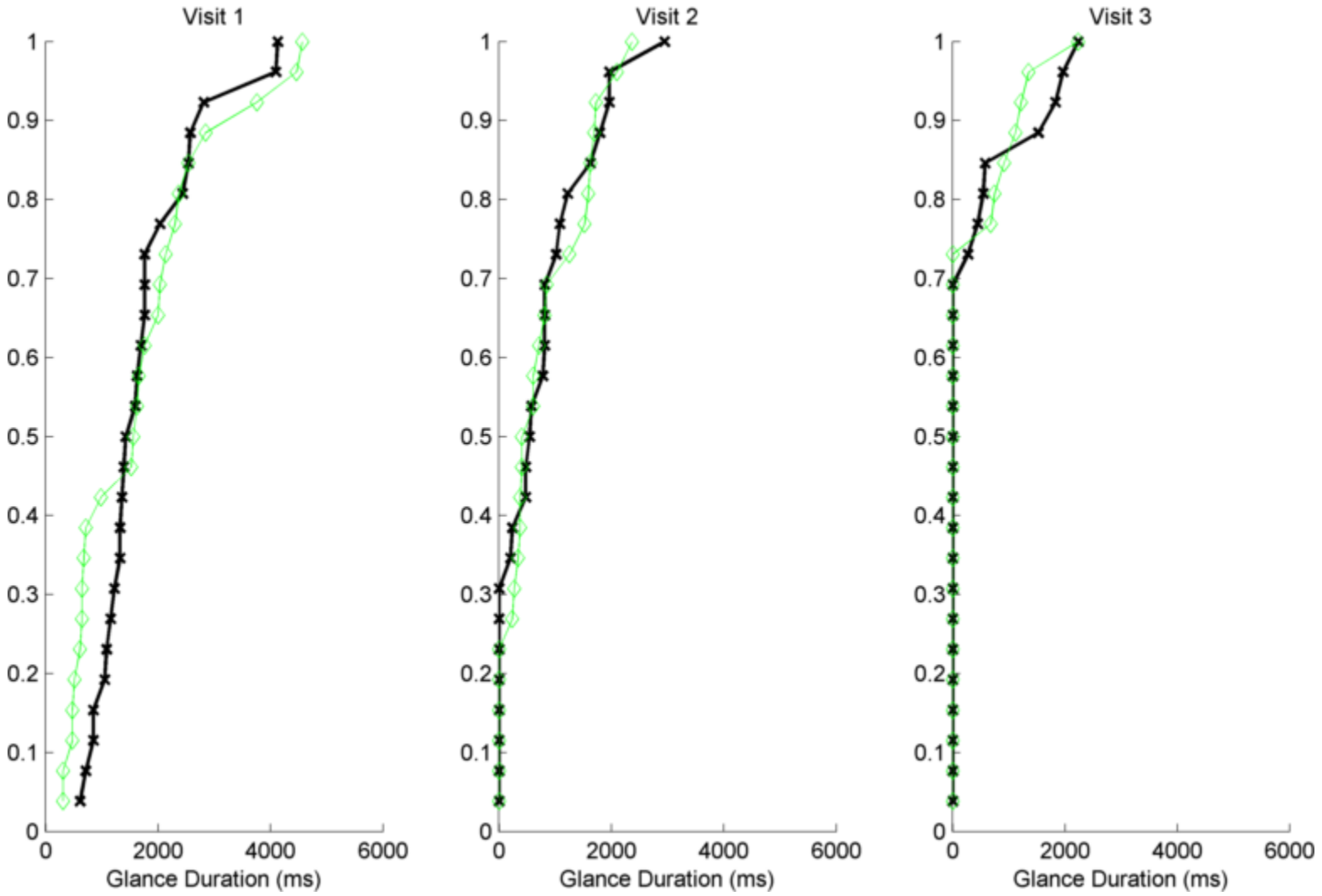
To formalize the comparisons between the distributions of Glance Duration by Visits, the Kolmogorov-Smirnov test for equality was performed between the sum of GD1 and GD2 on emotion images of the Emotion slides and the sum of GD1 and GD2 on neutral images of the Emotionless slides. The results are summarized in Table 5.2. The KS-statistic is the maximum absolute vertical difference between the two CDFs being compared. Table 5.2 shows that two Apathetic subjects (1AAD and 2AAD) have the smallest difference between the two CDFs with the exception of Subject 6CO. When transformed into p-values using the Kolmogorov distribution, these subjects have the highest p-values indicating that the distributions of glance durations on the two image types are equal. The next largest p-value aside from these three subjects is 0.3215 from Subject 4AD. Though, this value does not fall below the 0.05 significance level, the large gap in p-value separating the apathetic subjects and all other subjects (except 6CO) suggest that this measure of apathy might perform well on this set of 16 participants. If the significance level is set at a p-value anywhere in the gap between apathetic and non-apathetic participants, this measure will identify both apathetic patients as apathetic and misclassify only 1 out of 12 non-apathetic participants. It is expected that the significance level be adjusted upward to account for the extra detail of the emotion images in comparison to the neutral images. The CDFs of GD1, GD2, and GD3 of the one mislabelled age-matched control participant (6CO) are shown in Figure 5.6. Strangely, this age-matched control shows no differentiation between emotion and neutral images in any of the visits. This lack of differentiation demonstrates that there is still variability between individuals that are not explained by Glance Duration by Visit. Such variability between individuals may be better understood in combination with the other evaluations recorded in this study (e.g. MMSE, CCSD, AES, etc.).

The selection of using visits 1 and 2 together is based on observations of the CDFs between the individuals. Since the current sample size is small, the performance of the visual scanning apathy test is not necessarily representative of its predictive value. This test was done to demonstrate how within-subject normalizations might be performed to create an improved indicator of apathy. With a larger number of participants, a more sophisticated model of selecting visits may be developed and subsequently validated.

	AD Apathetic		AD Controls					
	1AAD	2AAD	1AD	2AD	3AD	4AD	5AD	6AD
KS Stat	0.1905	0.25	0.4375	0.625	0.6	0.3462	0.5	0.4455
p-value	0.9295	0.6325	0.0657	0.0019	0.0047	0.3215	0.0231	0.1822

	AD Depressed		non-AD (age-matched) Controls					
	1DAD	2DAD	1CO	2CO	3CO	4CO	5CO	6CO
KS Stat	0.375	0.5	0.375	0.4375	0.7802	0.6524	0.5	0.1538
p-value	0.1625	0.0656	0.1625	0.0657	0.0002	0.002	0.0389	0.995

Table 5.2: KS-statistic and p-values of Kolmogorov-Smirnov test for equality between GD1+GD2 on emotion images of the Emotion slides and GD1+GD2 on neutral images of the Emotionless slides.



Figures 5.6: The cumulative distribution function of glance durations by visit 1, 2, and 3 are shown for an age-matched control (6CO). The thick dark line with 'x' points represent glance durations on emotion images on Emotion slides. Light line with diamond points represent glance durations on two randomly selected neutral images on Emotionless slides.

5.3.2. Measure of Depression

The second hypothesis postulates that depressed participants will exhibit attention bias towards dysphoric images in comparison to controls. Figure 5.3 illustrated that all participants have a bias towards emotion images in comparison to neutral images (i.e. mean RTEI > 0.5 for all subjects). To compare the attention bias towards dysphoric images with the bias towards social images, a normalization procedure similar to the one shown in Chapter 3 was performed to increase the independence between the dysphoric and social observations. The two resulting measures were named Relative Interest in Dysphoric Images (RID) and Relative Interest in Social Images (RIS) and are defined by Equations 5.1A and 5.1B below.

$$\text{A) } RID = \frac{RFT_{dys}}{RFT_{dys}+RFT_{neu1}+RFT_{neu2}}$$

$$\text{B) } RIS = \frac{RFT_{soc}}{RFT_{soc}+RFT_{neu1}+RFT_{neu2}}$$

Equation 5.1: Calculation of A) Relative Interest in Dysphoric Images (RID) and B) Relative Interest in Social Images where RFT_{dys}, RFT_{soc}, RFT_{neu1}, RFT_{neu2} are the Relative Fixation Times on dysphoric, social, and the two neutral images on the Emotion slides.

In Figure 5.7, the mean RID - RIS is shown for every participant. Five participants (1 apathetic, 2 Alzheimer's control, and 2 age-matched control) had mean values below 0 indicating that these participants had an attention bias towards social images over dysphoric images. Seven participants (1 apathetic, 4 Alzheimer control, 2 age-matched control) had mean values approximately equal to 0 indicating that these participants did not have attention bias towards either dysphoric or social images. Lastly, four participants demonstrated attention bias towards dysphoric images over social images (2 depressed, 2 age-matched control). Interestingly, subjects 5CO and 6CO who also scored higher on the visual scanning apathy indicator (i.e. had higher p-values for the K-S equality of distributions test) demonstrated bias towards dysphoric stimuli. It may by indicative to study the similarities of these two subjects on the other possible covariates that were recorded (e.g. age, education, MMSE, etc.). Unfortunately, a discussion of individual subject scores or other personally identifiable information is against the privacy policy of the current Sunnybrook Hospital REB approval [REB# 242-2011]. With a larger sample size, aggregate data may be used to identify correlations between these possible covariates. Excluding

Subjects 5CO and 6CO, mean RID - RIS appears to be a good indicator of depression. Once again, given the small sample size, analysis of the predictive value of these visual scanning indicators cannot be performed at this time.

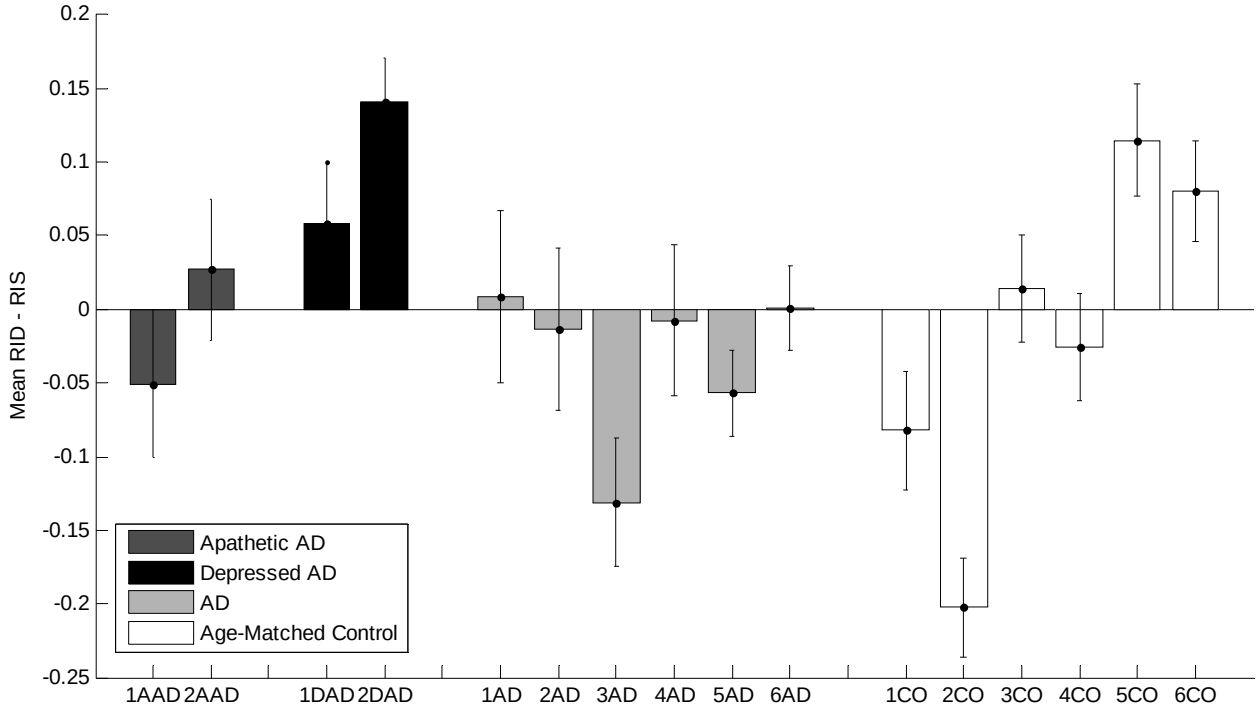


Figure 5.7: Mean RID - RIS for each participant. Error bars show standard error in the mean.

5.4. Conclusions

A method of using VSB as an objective indicator of apathy and depression was described in this chapter. Relative Time on Emotion Images was shown to be a poor measure of apathy because sensitivity was lost when behaviour of an entire slide is averaged together. Glance Duration by Visits characterized VSB on Emotion and Emotionless slides. Comparing the distributions of glance durations on emotion images of Emotion slides with glance durations on neutral images of Emotionless slides provided a good indicator of apathy. Using the non-parametric Kolmogorov -Smirnov test, apathetic participants were identified with high p-values in comparison to all other non-apathetic participants except for one age-matched control. Depressed participants could be detected using Relative Interest in Dysphoric Images (RID) and Relative Interest in Social Images (RIS) which are two normalized measures. While the

depressed subjects had significantly higher RID than RIS, all other participants except for two age-matched controls did not. Future efforts to expand on these preliminary results should validate the methods using larger sample sizes and improvements may be made by understanding the possible covariates affecting the clarification of patients and subjects in the different groups.

Chapter 6. Conclusions

6.1 Summary

The main contributions of this thesis are the following:

- Development of a general framework to study VSB in psychiatric disorders

  Presentation creation, data recording, and development of visual scanning parameters were discussed together as a general framework for studies in psychiatric disorder. Guidance for selection, organization, and presentation of visual stimuli was provided. A set of visual scanning parameters that characterize VSBs was described. These parameters included a novel subset that organizes VSB into visits which provides a method to study temporal aspects of VSBs.

- Analysis of VSB in adolescent onset Anorexia Nervosa

  Central psychopathological concerns often automatically bias attentional processing and VSB. VSBs demonstrated that patients with AN exhibited bias towards thin and fat body shapes images which were not present in age-matched control subjects. When thin and fat images competed for attention simultaneously (images are presented together), AN patients demonstrated stronger bias towards thin body shapes, thus highlighting the importance of the thin ideal as playing a central role in the cognitions of patients with AN.

- Development and characterization of a log-likelihood ratio processor

  It was demonstrated that a log-likelihood processor can be made from multiple observations of a visual scanning parameter in an individual. This processor was expanded upon by using multiple dimensions of attention bias (desire for thinness and social withdrawal) resulting in a detector with high sensitivity (92%) and specificity (90%). Analysis of biases in VSB provides an objective test to identify AN even in patients that do not score in the clinical range on self-report measures.

- Investigation of VSB as an early indicator of antidepressant efficacy in patients with Major Depressive Disorder

  Attention biases towards dysphoric stimuli have been observed in Major Depressive Disorder. A significant shift in VSB away from dysphoric stimuli was observed as early as 2 weeks following the start of Duloxetine Monotherapy in responders but not in non-responders. Large inter-subject variability was observed and discussed. Procedures for future improvement of the study were suggested.

- Pilot study for objectively distinguishing apathy from depression in elderly Alzheimer's patients using VSBs

  Apathy, in part, has been characterized by a lack of response to emotional stimuli. This feature of apathy was not observed in a small group of participants when Relative Fixation Time was used to characterize VSB. However, significant differences between apathetic and non-apathetic participants, including depressed patients, were identified using Glance Durations by Visit. These results demonstrate the utility of this novel visual scanning parameter and highlight the value of quantifying VSBs with greater temporal resolution.

6.2. Future Work

6.2.1. Normalization

Visual scanning behaviour (VSB) is controlled by both low-level perception processes (e.g., temporal and spatial characteristics of the visual stimuli) and high-level cognitive processes, which are driven by the subject's memories, emotions, expectations and goals. As low-level perceptual processes (processes associated with features like colour, contrast, edges etc.) are usually similar in psychiatric patients and controls, differences in VSB reflect mainly differences in high-level cognitive processes. Nevertheless, low-level perception processes affect the observed visual scanning parameters and therefore these measures reflect both low and high level processes. In future studies the variability of the measured VSB parameters may be reduced by normalizing the measurements with respect to low-level perception processes (e.g., using

saliency maps). Such normalization procedures can further improve the sensitivity and specificity of objective indicators by explaining away variability.

6.2.2. Glance duration by Visit Order

In the visual scanning parameters Glance Durations by Visit, Fixations and Number of Fixations by Visit, an implicit assumption was made that the n^{th} visit to an image is comparable to the n^{th} visit to any other image on the same slide. This assumption disregards the order in which the images are visited. The parameter Glance duration by Visit Order provides this information. Figure 6.1 plots Glance Duration of Visit 1 by Visit Order for all the control participants of the Anorexia Nervosa study when "filler slides" were viewed (i.e., the content of the images on each slide is similar). Glance Duration Visit 1, Order 1 is approximately 800 ms, Glance duration Visit 1, order 2 and 3 are approximately 900ms and glance durations with higher orders are approximately 1300 ms. Clearly, the order of visiting the images on a slide affects the glance duration. By estimating the scanning behaviour by visit order per person, the Glance Duration by Visits may be normalized to increase the sensitivity of this parameter to detect biases in VSB.

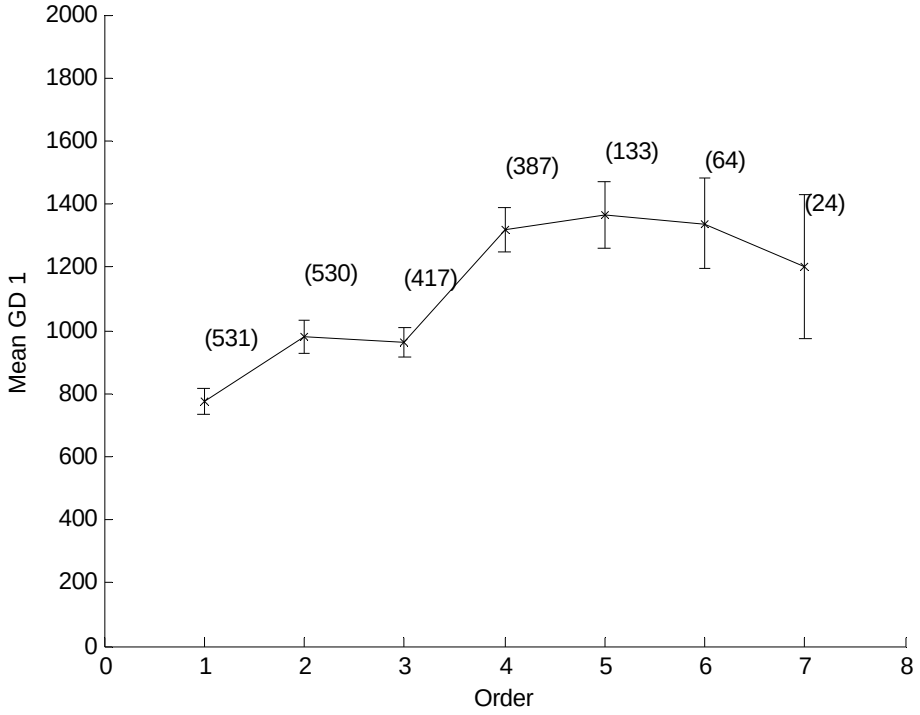


Figure 6.1: Mean Glance Duration Visit 1 by Visit Order. Error bars show standard error of the mean and numbers beside the data points indicate number of observations. Visit orders above 4 are possible when there are only 4 images when there is a revisit of an image before all images are seen. The approach to scanning the images is subject-dependent thus there are an unequal number of observations between visit orders.

References

Ambrosi-Randic N, Pokrajac-Bulian A., "Psychometric properties of the eating attitudes test and children's eating attitudes test in Croatia," *Eat Weight Disord*, Vol. 10, No. 4, Pp. 76-82, 2005.

American Psychiatric Association Committee on Nomenclature and Statistics: Diagnostic and Statistical Manual of Mental Disorders. 3rd edition. Washington, D.C.: *American Psychiatric Association*, 1980.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders: Fourth Edition. Text Revision (DSM-IV). Washington, DC: *American Psychiatric Association*, 2000.

Ballard, D., Hayhoe, M., and Pook, P., "Deictic codes for the embodiment of cognition," *Behavioral and Brain Sciences*, Vol. 20, Pp. 723–767, 1997.

Bauser, D.A.S., Suchan, B., Duan, I., "Differences between perception of human faces and body shapes: Evidence from the composite illusion," *Vision Research*, Vol. 51, Pp. 195-202, 2011.

Becker, M. W., Detweiler-Bedell, B., "Early detection and avoidance of threatening faces during passive viewing," *The Quarterly Journal of Experimental Psychology*, Vol. 62, No. 7, Pp. 1257-1264, 2009.

Canadian Psychiatric Association. "Canadian clinical practice guidelines for the treatment of depressive disorders," *Can J Psychiatry*, Vol. 46, Supp1, 2001.

Cassel, D., An Objective Methodology to Assess Visual Acuity using Visual Scanning Parameters. University of Toronto M.A.Sc. Thesis. 2009.

Caseras, X., Garner, M., Bradley, B.P., Mogg, K., "Biases in Visual Orienting to Negative and Positive Scenes in Dysphoria: An Eye Movement Study," *Journal of Abnormal Psychology*, Vol. 116, No. 3, Pp. 491-497, 2007.

Celada, P., Puig, M. V., Amargos-Bosch, M., Adell, A., Artigas, F. "The therapeutic role of 5-HT_{1A} and 5-HT_{2A} receptors in depression," *J Psychiatry Neurosc,*. Vol. 29, No. 4, Pp. 252-265, 2004.

Daffner, K. R. , Mesulam, M. M., Cohen, L. G., Scinto, L. F., "Mechanisms underlying diminished novelty-seeking behavior in patients with probable Alzheimer's disease," *Neuropsychiatry Neuropsychol Behav Neurol*, Vol. 12, Pp. 58-66, 1999.

Davis, C., Woodside, D.B., "Sensitivity to the rewarding effects of food and exercise in the eating disorders," *Comparative Psychiatry*, Vol. 43, No. 3, Pp. 189-194, 2002.

Dewa, C.S., Chau, N., Dermer, S., "Examining the Comparative Incidence and Cost of Physical and Mental Health-Related Disabilities in an Employed Population," *Journal of Occupational and Environmental Medicine*, Vol. 52, No. 7, Pp. 758-762, 2010.

Dobson, K.S., Dozois, D.J., "Attentional biases in eating disorders: a meta-analytic review of Stroop performance," *Clin Psychol Rev*, Vol. 23, No. 8, Pp. 1001-1022, 2004.

Downing, P.E., Jiamg, Y., Shuman, M., Kanwisher, N., "A cortical area selective for visual processing of the human body," *Science*, Vol. 293, Pp. 2470-2473, 2001.

Drago, V., Foster, P.S., Chanei, L., Rembisz, J., Meador, K., Finney, G., Heilman, K.M., "Emotional Indifference in Alzheimer's Disease," *J Neuropsychiatry Clin Neurosci*, Vol. 22, No. 2, Pp. 236-242, 2010.

Dunker KL, Philippi ST. "Differences in diet composition of Brazilian adolescent girls with positive or negative score in the Eating Attitudes Test," *Eat Weight Disord*, Vol. 10, No. 3, Pp. 70-75, 2005.

Eizenman M, Yu LH, Grupp L, Eizenman E, Ellenbogen M, Gemar M, Levitan RD, "A naturalistic visual scanning approach to assess selective attention in major depressive disorder," *Psychiatry Research* Vol. 118, Pp. 117-128, 2003.

Fairburn C.G., Cooper Z., Shafran R., "Cognitive Behaviour therapy for eating disorders: a 'transdiagnostic' theory and treatment," *Behaviour Research and Therapy*, Vol. 41, Pp. 509-528. 2003.

Garner, D.M., Olmsted, M.P., Bohr, Y., Garfinkel, P.E., "The eating attitudes test: psychometric features and clinical correlates," *Psychological Medicine*, Vol. 12, No. 4, Pp. 871-878, 1982.

Garner, D.M., Vitousek, K.M., Pike, K.M. "Cognitive-behavioral therapy for anorexia nervosa." In D.M. Garner & P.E. Garfinkel (Eds.), Handbook of treatment for eating disorders (pp. 94-144). New York: Guilford Press, 1997.

Giel, K.H., Friederich, H.C., Teufel, M., Hautzinger, M., Enck, P. and Zipfel, S., "Attentional processing of food pictures in Individuals with Anorexia Nervosa – An Eye-Tracking study," *Biological Psychiatry.* Vol. 69, Pp. 661-667, 2011a.

Giel, K.H., Teufel, M., Friederich, H.C., Hautzinger, M., Enck, P. and Zipfel, S., "Processing of pictorial food stimuli in patients with eating disorders – a systematic review," *Int. J. Eating Disorders*. Vol. 44, No. 2, Pp. 105-117, 2011b.

Gottesman, I., "The Endophenotype Concept in Psychiatry: Etymology and Strategic Intentions," *Am J Psychiatry*, Vol. 160, No. 4, Pp.636-645, 2003.

Green, M.W., McKenna, F.P., De Silva, M.S.L., "Habituation patterns to color naming of eating-related stimuli in anorexics and non-clinical controls," *British J. Clinical Psychology*, Vol. 33, Pp. 499-508, 1994.

Guestrin, E.D., Eizenman, M., "General theory of remote gaze estimation using the pupil center and corneal reflections," *IEEE Trans Biomed Eng*, Vol. 53, No. 6, Pp. 1124-33, 2006.

Guestrin, E.D., Eizenman M., "Remote point-of-gaze estimation requiring a single-point calibration for applications with infants," *Proceedings of the 2008 symposium on Eye tracking research applications,* Savannah, Georgia: ACM; 2008.

Harbottle EJ, Birmingham CL, Sayani F. "Anorexia nervosa: a survival analysis. Eating & Weight Disorders: EWD2008," Vol. 13, No. 2, Pp. 32-34, 2008.

Harris, E., Barraclough, B., "Excess mortality of mental disorder," *The British Journal of Psychiatry*, Vol. 173, No. 1, Pp. 11-53, 1998.

Hannerz, H., Borga, P., Borritz, M., "Life expectancies for individuals with psychiatric diagnoses," *Public Health*, Vol. 115, No. 5, Pp. 328-37, 2001.

Hannula, D.E., Althoff, R.R., Warren, D.E., Riggs, L., Cohen N.J., and Ryan J.D., "Worth a glance: using eye movements to investigate the cognitive neuroscience of memory," *Frontiers in human Neuroscience*, Vol. 4, Article 166, Pp. 1-13. 2010.

Health Canada. A Report on Mental Illness in Canada, 2002.

Hoek HW, van Hoeken D., "Review of the prevalence and incidence of eating disorders," *International Journal of Eating Disorders*, Vol. 34, No. 4. Pp. 383-396, 2003.

Hermans D, Vansteenwegen D, Eelen P. "Eye movement registration as a continous index of attention deployment: data from a group of spider anxious students," *Cognition and Emotion*, Vol. 13, No. 4, Pp. 419-434, 1999.

Hyman, S.E., "Can neuroscience be integrated into the DSM-V?" *Nat Rev Neurosci,* Vol. No. 8. Pp. 725-732, 2007.

In-Albon, T., Kossowsky, J., & Schneider, S., "Vigilance and Avoidance of Threat in the Eye Movements of Children with Separation Anxiety Disorder," *Journal of Abnormal Children Psychology*, Pp. 225-235, 2010.

Insel, T.R., Wang, P.S., "Rethinking mental illness," *JAMA*, Vol. 303, Pp. 1970-1971, 2010.

Janelle, C.M., Hausenblas, H.A., Fallon, E.A., Gardner, R.E., "A visual search examination of attentional biases among individuals with high and low drive for thinness," *Eating and Weight Disorders*, Vol. 8, Pp. 138-144, 2003.

Jansen A, Nederkoorn C, Mulkens S. "Selective visual attention for ugly and beautiful body parts in eating disorders," *Behaviour Research and Therapy,* Vol. 43, Pp. 183-196, 2005.

Johansson, L., Ghaderi, A., Andersson, G., "Stroop interference for food- and body-related words: a meta-analysis," *Eating Behaviors,* Vol. 6, No. 3, Pp. 271-281, 2005.

Kellough, J.L., Beevers, C.G., Ellis, A.J., Wells, T.T., "Time course of selective attention in clinically depressed young adults: An eye tracking study," *Behaviour Research and Therapy,* Vol. 46, No. 11, Pp. 1238-1243, 2008.

Kendell, R., Jablensky, A. "Distinguishing between the validity and utility of psychiatric diagnoses," *Am J Psychiatry*, Vol. 160, Pp. 4-12, 2003.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J., "Major depression and generalized anxiety disorder. Same genes, (partly) different environments?" *Arch. Gen. Psychiatry.* Vol. 49, Pp. 716–722, 1992.

Kessler, R.C., Chiu, W.T., Demler, O., Merikangas, K., Walters, E.E., "Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R)," *Arch Gen Psychiatry*. Vol. 62, Pp. 617-627, 2005.

Lee, M., Shafran, R., "Processing biases in eating disorders: the impact of temporal factors," *International Journal of Eating Disorders*, Vol. 41, No. 4, Pp. 372-375, 2008.

Lim, Jacobs, Ohinmaa, Scheopflocher, Dewa, C.S., "A new population-based measure of the burden of mental illness in Canada," *Chronic Diseases in Canada*, 2008.

Long, C.G., Hinton, C., Gillespie, N.K., "Selective processing of food and body size words: application of the Stroop Test with obese restrained eaters, anorexics, and normals." *International Journal of Eating Disorders*, Vol. 15, No. 3, Pp. 279-83, 1994.

MacLeod, C., Mathews, A., and Tata, P., "Attentional Bias in Emotional Disorders."*Journal of Abnormal Psychology*, Vol. 96, No. 1, Pp. 15-20, 1986.

Malberg, J.E., "Implications of adult hippocampal neurogenesis in antidepressant action." *J Psychiatry Neurosci*. Vol. 29, No. 3, Pp. 196-205, 2004.

Mathews, A., and MacLeod, C., "Cognitive approaches to emotion and emotional disorders." *Annual Review of Psychology,* Vol. 45, Pp. 25-50, 1994.

Matthews A, MacLeod C., "Cognitive vulnerability to emotional disorders." *Annual Review of Psychology*, Vol. 1, 167-195, 2005.

Nesse, R., Stein, D.J., "Towards a genuinely medical model for psychiatric nosology." *BMC Medicine*. Vol. 10, No. 5, 2012.

Norris, M.L., Boydell, K.M., Pinhas, L., Katzman, D.K., "Ana and the Internet: a review of pro-anorexia websites." *International Journal of Eating Disorders*, Vol. 39, No. 6, Pp. 443-7, 2006.

Olin, J.T., Katz, I.R., Meyers, B.S., Schneider, L.S., Lebowitz, B.D., "Provisional Diagnostic Criteria for Depression of Alzheimer's Disease." *American Journal of Geriatric Psychiatry*. Vol. 10, Pp. 129–141, 2002a.

Ottenbreit, N. D., Dobson, K. S., "Avoidance and depression: The construction of the Cognitive-Behavioral Avoidance Scale." *Behaviour Research and Therapy*, Vol. 42, Pp. 293–313, 2004.

Overduin, J., Jansen, A., Louwerse, E., "Stroop interference and food intake." *International Journal of Eating Disorders*, Vol. 18, No, 3, Pp. 277-85, 1995.

Pinhas, L., & Bondy, S. J., "Epidemiology of eating disorders in Children and Adolescents." In J. Lock (Ed.), *The Oxford Handbook of Child and Adolescent Eating Disorders: Developmental Perspectives* (Vol. 1). New York: Oxford University Press. 2010.

Preti, A., Girolamo, G., Vilagut, G., Alonso, J., Graaf, R., Bruffaerts, R., et al., "The epidemiology of eating disorders in six European countries: results of the ESEMeD-WMH project." *Journal of Psychiatric Research,* Vol. 43, No. 14, Pp. 1125-1132, 2009.

Rawal, A., Park, R.J., Williams, J.M.G., "Rumination, experiential avoidance, and dysfunctional thinking in eating disorders." *Behav Res Ther*. Vol. 48, No. 9, Pp. 851-859. 2010.

Robert, P., Onyike, C.U., Leentjens, A.F.G., et al. "Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders." *European Psychiatry*. Vol. 24, No. 2, Pp. 98–104, 2009.

Royal College of Psychiatrists, "Research priorities for psychiatry," 2011.

Richardson, D. C., Dale, R., & Spivey, M. J., "Eye movements in language and cognition: A brief introduction." In M. Gonzalez- Marquez, I. Mittelberg, S. Coulson, & M. J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 323-344). Amsterdam: John Benjamins. 2007.

Rinck M & Becker ES., "Spider fearfuls attend to threat then quickly avoid it: evidence from eye movements." *Journal of Abnormal Psychology*, Vol. 115, No. 2, Pp. 231-238, 2006.

Roefs, A., Jansen, A., Moresi, S., Willems, P., Grootel, S.V., Borgh, A., "Looking good. BMI, attractiveness bias and visual attention." *Appetite*, Vol. 51, Pp. 552-555, 2008.

Rosenhan, D. "On being sane in insane places." *Science,* Pp. 179:250, 1973.

Sackville, T., Schotte, D.E., Touyz, S.W., Griffiths, R., Beumont, P.J., "Conscious and preconscious processing of food, body weight and shape, and emotion-related words in women with anorexia nervosa." *International Journal of Eating Disorders*, Vol. 23, No. 1, Pp. 77-82. 1998.

Sears, C.R., Kristin, R.N., Ference, J.D., Thomas, C.L., "Attention to Emotional Images in Previously Depressed Individuals: An Eye-tracking Study," *Cogn Ther Res*, Vol. 35, Pp. 517-528, 2011.

Shafran R., Cooper Z., and Fairburn C.G., "Clinical perfectionism: a cognitive-behavioural analysis." *Behaviour Research and Therapy*, Vol. 40, Pp. 773-791, 2002.

Shafran, R., Lee, M., Cooper, Z., Palmer, R.L., Fairburn, C.G., "Attentional bias in eating disorders." *International Journal of Eating Disorders*, Vol. 40, No. 4, Pp. 369-80, 2007.

Shafran, R., Lee, M., Cooper, Z., Palmer, R.L., Fairburn, C.G., "Effect of Psychological Treatment on Attentional Bias in Eating Disorders." *International Journal of Eating Disorders*, Vol. 41, No. 4, Pp. 348-354, 2008.

Smeets, E., Jansen, A., Roefs, A., "Bias for the (un)Attractive Self: On the Role of Attention in Causing Body (dis)Satisfaction." *Health Psychology*, Vol. 30, No. 3, Pp. 360-367, 2011.

Steinhausen, H.C., "Outcome of eating disorders." *Child & Adolescent Psychiatric Clinics of North America*. Vol. 18, No. 1, Pp. 225-42, 2009.

Spitzer, R.L., Williams, J.B., Gibbon, M., First, M.B. "The structured clinical interview for DSM-III-R (SCID). I: History, rationale, and description." *Arch Gen Psychiatry*, Vol. 49, No. 8, Pp. 624-629, 1992.

Spivey, M., Tanenhaus, M.K., Eberhard, K.M., Sedivy, J.C., "Eye movements and spoken language comprehension: effects of visual context on syntactic ambiguity resolution."Vol. 45, No. 4, Pp. 447-481, 2000.

Stroop, John Ridley. "Studies of interference in serial verbal reactions". *Journal of Experimental Psychology*, Vol. 18, No. 6, Pp. 643–662, 1935.

Trew, J. "Exploring the roles of approach and avoidance in depression: An integrative model." *Clinical Psychology Review*. Vol. 31, Pp. 1156-1168, 2011.

Toh, W.L., Rossell, S.L., Castle, D.J., "Current visual scanpath research: a review of investigations into the psychotic, anxiety, and mood disorders." *Comprehensive Psychiatry*, Vol. 52, Pp. 567-579, 2011.

Wietersheim J., Kunzi, F., Hoffmann, H., Glaub, J., Rottler E., Rraue C., "Selective Attention of patients with Anorexia Nervosa While looking at pictures of their own body and bodies of others: An Exploratory Study." *Psychosomatic Medicine*, Vol. 74. Pp. 107-113, 2012.

Williamson, D.A., White, M.A., York-Crowe, E., Stewart, T.M., "Cognitive-behavioural theories of eating disorders." *Behaviour Modifications* Vol. 28, Pp. 711-738. 2004.

Wilson, M., "DSM-III and the transformation of American psychiatry: a history." *Am J Psychiatry,* Vol. 150**,** Pp**.** 399-410, 1993.

World Health Organization. "The World Health Report 2002: Reducing Risks, Promoting Healthy Life." Geneva, Switzerland: World Health Organization, 2002.

Yu, L.H., Eizenman, M., "A new methodology for determining point-of-gaze in head-mounted eye tracking systems." *IEEE Trans Biomed Eng*, Vol. 51, No. 10, Pp. 1765-73, 2004.