CALIBRATION OF TWO DIMENSIONAL SACCADIC ELECTRO-OCULOGRAMS USING ARTIFICIAL NEURAL NETWORKS.

MICHAEL J. COUGHLIN
B. Sc., Grad. Dip. Ed., B. A. (Hons)

SCHOOL OF APPLIED PSYCHOLOGY
GRIFFITH UNIVERSITY
QUEENSLAND

Submitted in fulfilment of the degree of Doctor of Philosophy

July 2002
Abstract

The electro-oculogram (EOG) is the most widely used technique for recording eye movements in clinical settings. It is inexpensive, practical, and non-invasive. Use of EOG is usually restricted to horizontal recordings as vertical EOG contains eyelid artefact (Oster & Stern, 1980) and blinks. The ability to analyse two dimensional (2D) eye movements may provide additional diagnostic information on pathologies, and further insights into the nature of brain functioning. Simultaneous recording of both horizontal and vertical EOG also introduces other difficulties into calibration of the eye movements, such as different gains in the two signals, and misalignment of electrodes producing crosstalk. These transformations of the signals create problems in relating the two dimensional EOG to actual rotations of the eyes. The application of an artificial neural network (ANN) that could map 2D recordings into 2D eye positions would overcome this problem and improve the utility of EOG. To determine whether ANNs are capable of correctly calibrating the saccadic eye movement data from 2D EOG (i.e. performing the necessary inverse transformation), the ANNs were first tested on data generated from mathematical models of saccadic eye movements. Multi-layer perceptrons (MLPs) with non-linear activation functions and trained with back propagation proved to be capable of calibrating simulated EOG data to a mean accuracy of 0.33° of visual angle ($SE = 0.01$). Linear perceptrons (LPs) were only nearly half as accurate. For five subjects performing a saccadic eye movement task in the upper right quadrant of the visual field, the mean accuracy provided by the MLPs was 1.07° of visual angle ($SE = 0.01$) for EOG data, and 0.95° of visual
angle (SE = 0.03) for infrared limbus reflection (IRIS®) data. MLPs enabled calibration of 2D saccadic EOG to an accuracy not significantly different to that obtained with the infrared limbus tracking data.
TABLE OF CONTENTS

Abstract ii
Table of Contents iv
List of Figures ix
List of Tables x
Acknowledgements xi
Statement of Originality xii
Introduction xiii

CHAPTER 1 1
Eye Movements 1
- Saccadic Eye Movements 1
- Smooth Pursuit Eye Movements 3
- Vergence Eye Movements 3
- Torsional Eye Movements 4
- Fixation Eye Movements 4
- Nystagmoid Eye Movements 4

CHAPTER 2 7
Measurement of Eye Movements 7
- Electro-oculography (EOG) 9
  - Background 10
  - Theory and General Procedure 11
  - Relative Strengths of the EOG Method 13
  - Limitations of the EOG Method 14
  - Calibration of EOG Signals 17
- Infrared Limbus Tracking System 18
  - Background 18
CHAPTER 3

Artificial Neural Networks (ANNs)

Origins of Parallel Distributed Processing
Background and Theory

A General Framework for PDP Models

A Set of Processing Units
The State of Activation
Output of the Units
The Pattern of Connectivity
The Rule of Propagation
Activation Rule

Modifying Patterns of Connectivity as a Function of Experience (learning rules)

Representation of the Environment

Some Important Developments

Simple Linear Models
Non-linear Units and Multiple Layers
Backpropagation - Learning internal representations by error propagation

The Generalised Delta Rule
The Sigmoid Function and Improving Learning

Application Performance

Conclusion

CHAPTER 4

Application of ANNs to the Measurement of 2D Saccadic EOG.

Transformations Inherent in Recording 2D EOG
Calibration of 2D EOG
CHAPTER 5

Experiment 1

Determination of the MLP Hidden Layer

Method

Participants

Materials

Electrodes and Hardware

Stimuli

Artificial Neural Networks

Procedure

Subject Preparation

Saccadic Task

Data Recording

Data Screening

ANN Architecture and Training

The closest-target estimate of calibration accuracy.

Results

Discussion

CHAPTER 6

Experiment 2

Simulation

Method

Saccadic Simulation Engine (SSE)

Noise Added to Simulated data

Fixation error.
Gain differences. 74
Rotation and translation of recording axes. 75
Physiological noise. 75
Simulated Waveforms 77
Artificial Neural Networks 80
The Simple Classifier (SC) 82
Simulated Data Sets 82
Calibration Accuracy 83

Results 83

Discussion 90
Summary 93

CHAPTER 7 94
Experiment 3 94
Subject Data 94
Method 94
Participants 94
Materials 95
Electrodes and Hardware 95
Head Restraint 95
Stimuli. 95
Procedure 97
EOG 97
IRIS® Infrared Limbus Tracker 98
Saccadic Task 98
Data Recording 99
Data Pre-processing 99
Effectiveness of Automatic Trial Rejection 99
List of Figures

Figure 1. The principal axes of eye rotation ................................................................. 2
Figure 2. The optokinetic reflex ..................................................................................... 5
Figure 3. The eye and its primary components ............................................................. 7
Figure 4. Schematic representations of some typical EOG waveforms ...................... 17
Figure 5. The IRIS head-mounted sensor device ......................................................... 19
Figure 6. A multilayer network .................................................................................... 28
Figure 7. An individual node and representative transfer functions............................ 31
Figure 8. A simple backpropagation network .............................................................. 40
Figure 9. Electrode placement and recording channels ................................................ 49
Figure 10. Schematic EOG waveforms and a multi-layer perceptron (MLP) ............... 54
Figure 11. Distance on a flat screen versus increasing visual angle ................................. 57
Figure 12. Approximate location of dot stimuli ............................................................ 59
Figure 13. Window of data points for saccades ............................................................ 62
Figure 14. Possible fixation errors ............................................................................... 73
Figure 15. Histogram of physiological noise .................................................................. 76
Figure 16. Examples of simulated and transformed waveforms ..................................... 79
Figure 17. Training curves for simulated data ............................................................... 87
Figure 18. Mean calibration output and errors for simulated data ................................. 89
Figure 19. Custom made head restraint ....................................................................... 96
Figure 20. Mean calibration error interaction .............................................................. 104
Figure 21. Training curves for subject data ................................................................. 107
Figure 22. Mean calibration output and errors for subject data .................................... 111
Figure B1. Flowchart for screening subject data (clean_M.m) ..................................... 137
Figure B2. Flowchart for pre-processing subject data (autocross2.m) ....................... 140
List of Tables

Table 1. Calibration Accuracy (percent correct) for Different Hidden Layers 64
Table 2. Calibration Accuracy (degrees of visual angle) for Simulated EOG Data 84
Table 3. Accepted and rejected trials for automatic and manual pre-processing 101
Table 4. Calibration Accuracy (degrees of visual angle) for Subject Data 102
Acknowledgements

Many thanks to my supervisors Dr. Tim Cutmore and Dr. Trevor Hine for their guidance and assistance over the years.

Special thanks to my wife Narelle, and children Reilly and Rebecca for their support, patience, and acceptance of a poorer lifestyle for many years.

This work was funded in part by an Australian Research Council grant A79601859.
Statement of Originality

“This work has not been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.”

Signed

________________________________________________

Date

________________________________________________
INTRODUCTION

Eye movements are an important source of behavioural information, subserved by a variety of brain mechanisms (Leigh & Zee, 1991; Masson & Mestre, 1998; Oster & Stern, 1980). There are many types of brain pathology and diseases of the central nervous system that can give rise to anomalous eye movements (Gambini, O., Abbruzzese, M., & Scarone, S. 1993: Hutton & Kennard, 1998; Karoumi, Ventre Dominey, Vighetto, Dalery, & d'Amato, 1998; Kennard, Crawford, & Henderson, 1994; Levy, Holzman, Matthysse, & Mendell, 1993; Opgenoorth & Baldaszti, 1992; Pallanti, Grecu, Gangemi, & Massi, 1996; Pallanti, Quercioli, Zaccara, Ramacciotti, & Arnetoli, 1998; Tien, Ross, Pearlson, & Strauss, 1996; Trillenberg et al., 1998).

The electro-oculogram (EOG) is the most widely used technique for recording eye movements in clinical settings. Of the 109 patient studies reviewed by Levy et al. (1993), 74 used EOG, and most of the remainder used infrared techniques. EOG is inexpensive, practical, non-invasive to the eye, and provides data in the form of a signature that may be read by the trained eye to diagnose pathology (Oster & Stern, 1980). Although there is considerable research being conducted on eye movements and their diagnostic potential, most recordings have been restricted to the horizontal dimension. For example, all the research articles mentioned above were based on horizontal recordings, and in the Levy et al. (1993) review, the authors report only two vertical eye movements in their results. Research by Previc (1982, 1990), Previc and Blume (1993), and Previc, Breitmeyer, and Weinstein (1995) has indicated that visual asymmetries, in relation to pattern recognition,
functional specialisation, search, and discriminability, exists between the left, right, upper, and lower visual fields. This suggests that the ability to accurately record and analyse two dimensional (2D) eye movements may provide additional diagnostic information on pathologies, and further insights into the nature of brain functioning.

Use of EOG is usually restricted to horizontal recordings, as vertical EOG contains eyelid artefact and blinks, as well as the physiological noise also present in horizontal recordings. Simultaneous recording of both horizontal and vertical EOG also introduces other difficulties into calibration of the eye movements, such as different gains in the two signals, and misalignment of electrodes producing crosstalk. These transformations of the signals create problems in relating the 2D EOG to actual rotations of the eyes.

The application of an artificial neural network (ANN) that could map 2D recordings into 2D eye positions would overcome this problem and improve the utility of EOG. It would be valuable in its potential to provide an accessible tool for diagnostic purposes, and for research into 2D eye movements. This application of an ANN would not be restricted to EOG, but would also be of use with optical eye trackers, such as infrared limbus trackers.

This thesis is about the calibration of two-dimensional (2D) saccadic eye movements using artificial neural networks (ANNs). Chapter 1 provides an overview of the basic anatomy of the eye and the different types of possible eye movements. Chapter 2 gives a brief introduction to the measurement of eye movements with emphasis on the two eye tracking systems used in the research: electro-oculography (EOG) and an infrared limbus tracker (IRIS®).
Chapter 3 introduces ANNs with an overview of their structure and operation. Chapter 4 looks at the difficulties associated with calibration of 2D saccadic EOG and the types of ANNs that may be suitable for this purpose. Chapter 5 describes an experiment using subject data to determine a suitable number of hidden nodes for one of the ANNs, the multi-layer perceptron (MLP). Chapter 6 describes an experiment that develops simulated 2D saccadic EOG waveforms to confirm the suitability of the chosen ANNs to perform the transformations necessary to calibrate 2D saccadic EOG without possible additional confounds introduced by human subjects. Chapter 7 describes an experiment that tests the ANNs as calibrators of 2D saccadic eye movement data, both EOG and IRIS®. Finally, Chapter 8 provides overall conclusions and suggestions for further developments.

Programmes written in Matlab® version 5 (The MathWorks, Inc., 1997) were used to process the data and the Matlab® Neural Network Toolbox version 2.0 (The MathWorks, Inc., 1994), was used to create and train the ANNs. Data from the three experiments, and programmes written for the research can be found on the accompanying CD. The directory structure in Appendix A lists the content of the CD and the heading hierarchy in Appendix A reflects the hierarchy of the Matlab® m files.
CHAPTER 1

Eye Movements

The eyeball rotates about three axes that intersect at the centre of the eyeball, and to a good approximation, intersect at a point that is fixed in both the eye and the orbit (Figure 1). Horizontal rotations (about a natural vertical axis) to and from the nose are referred to as adduction and abduction respectively; vertical rotation (about a natural horizontal axis) as elevation (up) and depression (down); and rotations of the top of the cornea (about the line of gaze) towards and away from the nose as intorsion and extorsion respectively.

Five neuronal systems are used in the control of eye position. The saccadic system points the fovea towards objects of interest and the smooth pursuit system keeps moving targets in the fovea. The vestibular-ocular and optokinetic reflexes compensate for head movements and the vergence movements system aligns the eyes to look at targets at different depths (Goldberg, Eggers, & Gouras, 1991). The final common pathways of all these systems are the motor neurons that innervate three complementary pairs of muscles attached to the eyeballs. The five neuronal systems give rise to six classes of eye movement: saccadic, smooth pursuit, vergence, torsional, fixation, and nystagmus. Unless referenced otherwise the following descriptions are from Young and Sheena (1975).

*Saccadic Eye Movements*

These are voluntary movements by which fixation is changed from one
Figure 1. The principal axes of eye rotation.

point to another. They are rapid and conjugate and have high initial and final acceleration (up to 40000 degrees of visual angle/s²) with a displacement-dependent peak velocity as high as 400 to 600 degrees of visual angle/s. Their duration is dependent on displacement and varies from 30 to 120 ms. Motion of the head is often involved when the saccade displacement exceeds 30 degrees of visual angle. Latencies are in the range 100 to 300 ms in response to a visual target that jumps about in the visual field, with a minimum delay of 100 to 200 ms (refractory period) occurring between successive saccades. A torsional component is associated with vertical or oblique saccades due to the arrangements of the six extraocular muscles (four rectus and two oblique). The horizontal movements by the medial and lateral recti are the only pure rotations, as these are the only pair of muscles attached on opposite sides of a great circle of the eye.

Smooth Pursuit Eye Movements

These are conjugate movements for tracking slowly moving visual targets (1-30 degrees of visual angle/s) and are not generally under voluntary control. That is, they usually require the existence of a moving visual stimulus for their execution and the system calculates how fast the target is moving and moves the eyes accordingly in a predictive manner. The smooth pursuit system operates with limited velocity and acceleration.

Vergence Eye Movements

These movements are disjunctive and involve slow rotation of the eyes in opposite directions to allow binocular fixation on an object moving away or towards the eyes. They are stimulated by focusing error and binocular disparity and have maximum velocities of 10 degrees of visual angle/s over a
range of nearly 15 degrees of visual angle. Slower and smoother than conjugate movements they also appear to be non-predictive.

**Torsional Eye Movements**

These are also disjunctive movements and involve compensatory slow rotations about the line of gaze. These rolling motions are limited to about 10 degrees of visual angle of rotation and are necessary for the perceptual stability of horizontal lines. Without compensatory torsional movements, lines that appear horizontal in some positions of gaze would be perceived to tilt in others (Goldberg et al., 1991).

**Fixation Eye Movements**

These are a variety of miniature eye movements, generally less than 1 degree of visual angle in displacement, which occur during attempted steady fixation of a target. Drift is a slow random motion away from the target at speeds of only a few minutes of arc per second. Flicks, or microsaccades, are small rapid movements of the same dynamic nature as larger voluntary saccades. These may help to re-fixate the target and have magnitudes as large as 1 degree of visual angle and occur at intervals as little as 30 ms. For any particular individual, drifts and flicks tend to occur along a single preferred axis. Also when fixating, normal individuals exhibit a high-frequency tremor (30-150 Hz) with peak displacements of approximately 30 arcsec of visual angle at about 70 Hz. Because of these fixation movements, accuracies of about 0.5 to 1.0 degrees of visual angle are often sufficient when trying to show what part of the visual field is being fixated.

**Nystagmoid Eye Movements**

This is a general term for oscillatory movements that includes both
Figure 2. The optokinetic reflex. Horizontal eye position as the subject sits still inside a vertically striped drum rotating slowly to their right. The eye position plot against shows that during the slow phase the eyes move in the same direction as the striped drum in order to keep the image of the drum still on the retina.

Saccadic and smooth movements that are normally of a compensatory nature. Optokinetic nystagmus, or train nystagmus, is a sawtooth pattern of eye movements (Figure 2) elicited by a moving patterned visual field. It has a slow phase, as the eye fixates and follows with slow phase eye movement some part of the moving field, and a fast phase consisting of a return saccadic jump, in which the eye fixates on another part of the visual field. Optokinetic nystagmus has a maximum frequency of about 5 Hz and a displacement of from 1 to 10 degrees of visual angle. Vestibular nystagmus is an oscillatory eye motion, similar to optokinetic nystagmus with both slow and fast phases,
attributable to stimulation of the semi-circular canals during brief head rotation. It also has a saw tooth pattern and is used as a test of the semi-circular canal function.

In summary, although there are several types of eye movements, the eye movements to be investigated in this dissertation are 2 dimensional (2D) saccadic eye movements. Saccadic eye movements arise from systems which are the easiest to isolate and record in laboratory testing. They are also eye movements which can be affected by damage to the cerebral hemispheres and whose functional deficits are associated with the major psychoses (Pierrotdeseilligny, 1994; Muir, St.Clair, Blackwood, & Roxburgh, 1992). For example, aberrations in smooth pursuit and saccadic eye movements are frequently observed in people with schizophrenia (Abel, Levin, & Holzman, 1992; Radant & Hommer, 1992; Levy, Holzman, Matthysse, & Mendell, 1994).
CHAPTER 2

Measurement of Eye Movements

There are various physical characteristics of the eye described by Young and Sheena (1975) that are used in eye movement measurement. There is the retina, the cornea, the limbus, the pupil, and the corneo-retinal potential, all of which move with the eye (Figure 3). Reflections also occur from various surfaces within the eye.

Figure 3. The eye and its primary components.
The first of these physical characteristics, the retina, was used in some of the earliest techniques for quantitative, though subjective assessment of eye movements. These techniques involved the use of after-images on the retina. Although rarely used, some techniques employed the use of the corneal bulge to determine eye movement. The cornea, the central area has a considerably smaller diameter than the eyeball itself and also projects forward. However, the cornea slips a small amount with respect to the sclera (the white of the eye) when forces are applied and possibly even during acceleration of the eyeball, so these techniques were not very accurate. More commonly, corneal reflections, and reflections from other optical curvatures in the eye (Purkinje Images) provide the basis of monitoring eye position. Some techniques also make use of the pupil and/or scleral blood vessels or folds of the iris. Electro-oculography (EOG) makes use of the corneo-retinal potential (see details in EOG section) and the infrared limbus tracker utilises the difference in corneal reflection on both sides of the eye to track the iris-scleral boundary, the limbus (see details in IRIS® section). Other techniques use video systems to track the limbus or pupil. Some other less used methods are the Contact Lens Method, the Point of Regard Measurement, and the Double Purkinje Image Method. All these techniques and others are reviewed by Young and Sheena (1975).

Important considerations for an eye measurement system are expense, portability, convenience, and subject comfort. Some of these may be traded-off for accuracy and precision. The ideal eye movement recording system should have electronic and mechanical stability, easy non-traumatic application, low cost, ample linear properties, wide measuring range, high spatial and temporal resolution, and preferably non-contact to the eye. No
existing method satisfies all these criteria, however the most important (non-contact to the eye) is met by both methods of interest here: the EOG and the infrared limbus tracker. The EOG is the focus of this research, as it appears to have the best compromise of the above considerations. It is due to its low cost and simplicity of use that it is the most widely used technique for recording eye movements in clinical settings. Of the 109 patient studies reviewed by Levy, Holzman, Matthysse, & Mendell, (1994), 74 used EOG, and most of the remainder used infrared techniques. EOG is practical, non-invasive, and provides data in the form of a signature that may be read by the trained eye to diagnose pathology (Oster & Stern, 1980). The infrared limbus tracker is of interest because it is commonly used in eye movement research and should prove invaluable as a verification tool, as the development of the EOG and artificial neural network system progresses.

**Electro-oculography (EOG)**

A potential difference, the corneo-retinal potential, exists between the cornea and the retina. The cornea is about 1mv positive with respect to the back of the eye. When the eyes move, movement of the electrical dipole within the eye can be recorded from skin electrodes placed proximal to the eye. The potential varies diurnally and with light adaptation (decreasing, following periods in the dark). For stable measurements, 30 to 60 minutes adaptation is needed prior to an experiment. Eye movements produce signals at skin electrodes (the electro-oculogram, EOG) that can be used to infer the eye movements taking place.
**Background**

Oster and Stern (1980) have provided the following background on the development of the EOG technique. It was not until the 1920s that recordings of human eye movements using the standing potential were available, although its existence had been known since du Bois-Reymond (1849). In the 1800s the standing potential's light dependency was observed in animal preparations, with the potential increasing with light stimulation and decreasing with darkness. Difficulties encountered with baseline drift saw most work with the EOG used for the recording of nystagmus where a.c. amplifiers were adequate, although there was some study on reading and evaluation of stress and fatigue. The availability of relatively drift free amplifiers in the late 1950s and the use of the light- and dark-induced EOG responses as clinical tools for the evaluation of retinal and choroidal integrity led to a renewed interest in electro-oculography. This renewed interest was also aided by developments in sleep research and visual information processing. The EOG is the only eye movement technique that can be used with the eyelids closed, for example, during sleep.

According to Shackel (1967), the EOG has been used in a wide range of applications including studies of oculomotor function, measuring performance during various practical tasks, and sleep studies. Oster and Stern (1980) add that some other uses have been to determine how good and poor readers differentially process the written word, how schizophrenics and normals differ in their performance on cognitive tasks, and how to enhance the study of complex behaviours. EOG continues being used to study the reading process (e.g. Poblano, Cordoba de Caballero, Castillo, & Cortes, 1996), and is
still the method most frequently used to examine eye movements in schizophrenia (Calkins, Katsanis, Hammer, & Iacono, 2001). It has been a powerful tool for psychologists and others allowing the study of more complex stimuli and tasks traditionally in the domain of cognition and perception (e.g., Galley, 1993; Muir et al., 1992).

Shackel (1967) listed some of the more practical tasks for which the EOG has been utilised, including simulated car driving, inspection in industry, and reading of instrument dials. EOG is also the primary method used in studies of sleep and dream activities due to its simplicity and ability to record with the eyes closed (e.g. Tsuji, Satoh, Itoh, Sekiguchi, & Nagasawa, 2000). The EOG is used for studies wanting to record blinks, giving typical spike waveform with amplitudes of ½ to 1 mV (e.g. Kobayashi, Hara, & Goi, 1996; Kong & Wilson, 1998). The first application to clinical problems was probably nystagmus recordings (electro-nystagmography), followed later by applications to vestibular function studies (Shackel, 1967). Researchers are still using EOG for nystagmus recordings (e.g. Katayama & Mori, 2001). Many other neurological syndromes also require EOG recordings for diagnosis (e.g. Behrens & Weiss, 1999; Kawasaki & Tamura, 1987; Pinkers, Cuypers, & Aandeker, 1996; Spada, Bisti, Colucci, & Balestrazzi, 1999).

Hence it is important to find ways to maximise the potential of the data collected from this widely applied technique.

Theory and General Procedure

The operation of the EOG system is based on the principle that as the eye rotates so does the electric dipole within the eye, and electrodes placed on the skin surface near the eye socket will detect the potential field, with the
voltage changes being linearly related to the angle of eye rotation for excursions at least up to 30 degrees of visual angle away from the centre (Shackel, 1967). Young and Sheena (1975) have reported that theoretically, as the eye and its electrical dipole rotates, the potential difference, as measured in a plane normal to the principal axis, varies with the sine of the angle of rotation. Oster and Stern (1980) reported both relationships. The resultant voltage differences between a pair of electrodes placed along the appropriate plane of rotation is essentially linearly related to the angle of gaze for ±30 degrees of visual angle and to the sine on the angle from ±30 up to ± 60 degrees of visual angle.

Oster and Stern (1980) suggested the following procedure for preparation and measurement. Electrodes should be placed as close to the eyes as is comfortable and the skin prepared by briskly rubbing with alcohol. Electrodes placed at the outer canthi are used for binocular horizontal recording of conjugate eye movements and electrodes above and below the eye for vertical eye movements. Being as close to the eye as possible is important, as larger distances give smaller potentials. When looking straight ahead, the planes of rotation should intersect at right angles through the pupil. Orthogonal planes of rotation are essential to minimise crosstalk, although Shackel (1967) reported that some crosstalk occurs due to subjects’ non-linear potential fields, which can only be overcome by careful subject selection or complete calibration of the field for the entire rotation range of interest. He also warned of the problem of unbalanced potentials for equal but opposite eye movements if the electrodes are not positioned symmetrically. Young and Sheena (1975) recommend silver-silver chloride electrodes as they provide
minimum discomfort and resist excessive polarisation over many minutes of use. They further recommended using shielded cables and a grounded ear electrode to help minimise electro-magnetic pickup. All the above authors also suggested low pass filtering of the analogue signal. This is discussed in more detail below in the section on limitations of the EOG.

Stern, Ray and Davis (1980) supported the selection of silver-silver chloride electrodes. They reported on an experiment in which pairs of electrodes of different metals were placed in a saline solution and the potential difference between each metal pair (the bias potential) was measured over time. Platinum produced a bias potential of 320 mV; silver, 94 mV; zinc, 100 mV; and silver-silver chloride, only 2.5 mV. The silver-silver chloride electrodes not only have the smallest initial error in their bias potential, but they also show a relatively small drift of potential with use, that is, their polarisation potential is minimal. Polarisation is the building up of a counter electromotive force that effectively appears as an increase in the subjects' resistance. It can be thought of as an unequal distribution of ions on the electrode surfaces as a result of the passage of current through the electrolyte (Stern, Ray, & Davis, 1980).

Relative Strengths of the EOG Method

The advantages of the EOG for recording eye movements are its non-invasive nature, the low cost, and its ability to work with the eyelids closed. No equipment needs to be in actual contact with the eye, and the setup time for positioning the electrodes is only a few minutes.

Oster and Stern (1980) reported that the EOG method is the only technique where the signal is generated by a bioelectric event as opposed to a
signal energy of reflected light. Because the source is within the eye, head movements do not hinder accuracy unless the absolute eye position within the visual field is required. The voltage output is more simply related to the eye's angular deviations than it is to the metrics of reflected light from optical systems. The frequency response is also superior, not being limited by camera frame speed as in video based systems. They also claimed that its physiological basis enhances its effective range with ±30 degrees of visual angle or since visualisation of the eye is not necessary, even ±80 degrees of visual angle available on both horizontal and vertical channels. Accuracies were in the range of ±1 degree of visual angle for the horizontal channel. Furthermore, blinks in the vertical channel did not cause discontinuities in the vertical recording. Although optical methods may be more accurate, they have limited effective range (25-30 degrees of visual angle), and eye closure and blinking interrupt recording. Young and Sheena (1975) claimed a slightly less effective range of ±70 degrees of visual angle and slightly lower accuracy’s of ±1.5-2 degrees of visual angle. They also reported that linearity progressively gets worse above 30 degrees of visual angle.

Another important advantage of the EOG method relates to its cost. The necessary standard equipment is relatively inexpensive and is available in laboratories and hospitals where basic physiological responses are routinely recorded.

Limitations of the EOG Method

The weaknesses are a lack of resolution (relative to optical methods) and susceptibility to external sources of artefact. EMG artefact can be from extra-ocular muscle potentials such as facial muscles, body movement, and
even the eye muscles themselves. Skin potential response (SPR) signals and brain potentials (EEG) may also be recorded.

Resolution seems to be limited to about ±1 degree of visual angle compared to about ±1 minute of arc of visual angle for magnetic induction or the infrared limbus tracking methods (Reulen et al., 1988). Although this claimed degree of accuracy for the infrared limbus tracking is probably somewhat optimistic. Carmody, Kundel, and Nodine (1980) reported mean spatial fixation accuracies for infrared limbus tracking using a chinrest and over the entire visual field of 0.58 degrees of visual angle horizontally and 0.51 degrees of visual angle vertically). Shackel (1967) reported that SPR signals are within the same general frequency band as EOG signal frequencies, and that the transition from ionic current flow in the body to electronic current flow in the amplifier may produce noise in the form of slow drift. He claimed that noise problems from EMG, except from teeth clenching, can be filtered out by using a bandwidth between 1 and 40 Hz. Furthermore, EEG interference on most subjects is less than about a half of a degree of visual angle. Oster and Stern (1980) confirmed these findings. They reported that interference can come from muscle activity (EMG), skin potential responses, skin conductance changes, and cortical activity (EEG), as their frequencies overlap EOG signal frequencies and will be algebraically added. Although skin conductance changes are less predictable, the others can be attenuated. Only intense EMG cannot be eliminated by electrode placement, but can be low-pass filtered.

Both Shackel (1967) and Young and Sheena (1975) reported an average potential of about 20 μV/degree of rotation. Oster and Stern (1980)
reported on the variation in the corneo-retinal potential, with up to 10µV/deg having been recorded associated with age, sex, blood sugar, diurnal variations, globe protrusions and myopia, and illumination transients. Illumination transients involved potential changes within the first 60 to 80s if a dark-adapted eye was suddenly illuminated, rising steeply at 1 to 5 min and reaching an apex in about 8 to 10 min. The opposite also occurred with light reduction. These latencies and amplitudes are widely used measures of retinal and choroidal integrity. Changes in illumination were therefore a problem for calibration and can require 29 to 52 min for dark to light adaptation and 17 to 51 min for light to dark adaptation. Potential levels were significantly higher for women with minimum and maximum differences of 1.67 µV/deg and 3.33 µV/deg. These values were taken throughout an entire course of dark-light adaptation. Potential levels also tended to decline from age 11 to 32 years.

Although blinks do not stop EOG recordings, eyelid movement does pose its own particular problem for EOG recordings of vertical eye movements. Oster and Stern (1980) concluded that the upper eyelid acts as a sliding resistor and extension enhances conductivity. Downward movement of the lid increased the positive signal on the electrode above the eye. The same effect obtained if the eyeball had rotated up. Movement of the upper lid relative to the eyeball was the source of the eyelid artefact (Figure 4), which occurred when the EOG potential rises above the level established by the succeeding steady-state potential. This only occurred in the vertical recordings where the distance between the eyelid and electrode was variable. Monopolar recording with an electrode below the eye greatly reduced this artefact but reduced signal strength. The blink waveform (Figure 4) has been attributed to
this effect of the eye lid as well. Shackel (1967) also noted this effect when referring to the spike at the end of an upward saccade being due to upper eyelid movement relative to the eyeball.

![Figure 4](image_url)

*Figure 4.* Schematic representations of some typical waveforms obtained from EOG recordings: (a) a horizontal saccade, (b) a blink (not to scale), (c) a vertical saccade showing eyelid artefacts.

**Calibration of EOG Signals**

Determination of eye position using EOG is usually based on the amplitude of the recorded voltage at the completion of the saccade compared to the amplitude at the previous fixation (Shackel, 1967; Oster & Stern, 1980). This amplitude is represented by the difference in height between the two flat sections of the EOG saccadic waveform (see Figure 4a). This voltage is
compared to voltages recorded in a series of calibration trials in which the eyes have performed a series of saccades to targets of known coordinates.

*Infrared Limbus Tracking System*

Although the EOG method is the primary focus of the research, the IRIS® system was also used simultaneously to record 2D saccadic eye movements. The IRIS® system (model 6500) is an infrared limbus tracker produced by Skalar Medical (Figure 5). The main aim of incorporating the IRIS® is to use it as a performance benchmark for the 2D EOG-ANN combination. For horizontal saccades, the IRIS® performs to a high degree of accuracy (Reulen et al., 1988). Because of the difficulties in recording vertical eye movements, the recording of 2D eye movement tasks is an area in which little work has been done using either system. Following is an outline of the IRIS® system and its performance, summarised from Reulen et al.’s (1988) evaluation paper.

*Background*

Infrared limbus trackers are based on the reflection of infrared light by the sharp boundary between the iris and the sclera (Young & Sheena, 1975). This sharp boundary is the limbus (Figure 3). Infrared-light emitting diodes and infrared-sensitive detectors are positioned in front of the eye, such that on both sides of the eye (nasal and temporal) their receptive fields match the iris/sclera transition. With horizontal abduction of the eye, the nasally positioned detector will record an increased reflection whereas the temporally positioned detector will record a decrease in reflection. Subtraction of the signals gives eye position with respect to the head. Repositioning a second set of emitters and detectors also allows measurement of vertical eye movements.
Figure 5. The IRIS head-mounted sensor device.

Reulen et al. (1988) report that prior to the release of the IRIS®, the already existing differential reflection techniques had serious drawbacks such as limited linear range, poor mechanical stability of the device with respect to the eye, and lengthy complicated installation and calibration procedures. The IRIS® differential reflection technique addresses these drawbacks and following is an outline of its general principles and properties.

Theory and General Procedure

The Infrared-light transducer has a maximum infrared-light emission at 950 nm and optimal sensitivity of its detectors at 850 nm. Beamwidth (angle between the peak response and half-power point) is 24° for the emitters and 14° for the detectors. The detector array, consisting of nine phototransistors is
positioned above the transmitter array of nine infrared-light-emitting diodes. This arrangement means that a fixed portion of the radiant infrared light, constrained within a narrow beamwidth, covers a selected reflector area, so the active area of each transistor is matched with a particular part of the limbus. The arrays have been fitted into an anodised aluminium case that can be clipped to the lightweight head support in two ways. This enables recording of either horizontal or vertical eye movements. Field of view for the horizontal is limited to about 35 degrees of visual angle from the central position and 20 degrees of visual angle for the vertical. To ensure mechanical stability the frames are rigidly attached to an adjustable head support, and for optimal alignment, independent adjustment is possible in all three directions by means of rotative mechanisms (Figure 5). This provides easy and non-traumatic application of the method especially for subjects with little experience or motivation.

Eye-movement signal detection with minimised interference from ambient light is possible due to the chopped emission of infrared light. The chopped emission refers to the generation of a square wave signal. Combined with synchronous signal detection this also provides an improved signal-to-noise ratio. Chopping also allows a higher level of energy emission of infrared light, compared to continuous emission, while still maintaining the same time-integrated amount of infrared-light energy upon the eye. The chopping process considerably reduces signal interference, which may be caused by changes in ambient light.

The oscillator generates a square-wave signal of 2.5 kHz that drives a 50 mA current through each of the nine infrared-light-emitting diodes that are
connected in parallel. For horizontal measurements the two most laterally
located detector signals (nasally 1 & 2 and temporally 8 & 9) are summed and
then subtracted from each other. For vertical measurements 2 & 3 and 7 & 8
are used as these combinations reduce eyelid artefacts. Low pass filters are
used to further reduce any interference. The noise level of the system
corresponds to an eye movement of about 1 min of arc. Linearity of the signal
was measured for both horizontal and vertical movements. Deviation from
linearity as a maximum deviation from the range (28 degrees of visual angle for
the horizontal, and 16 degrees of visual angle up and 18 degrees of visual angle
down for the vertical) was 3 percent and 2 percent respectively. Cross
coupling, which is unavoidable in systems using light reflection techniques,
was with optimal alignment, about 1 in 10. That is, an eye rotation of 10
degrees of visual angle in either the horizontal or vertical direction would
produce a cross-coupled signal of 1 degree of visual angle in the

corresponding orthogonal direction. This crosstalk would be problematic in
calibrating 2D eye movements.

In a comparison of ENG (electronystagmography) and MI (magnetic
induction) against IRIS® for 15 degrees of visual angle saccades either side of
a central spot, the IRIS® signal is almost free of drift and has a much higher
resolution (1 min of arc compared to 1-2 degrees of visual angle for the ENG).
The ENG also shows small pendular eye movements with a frequency of
about 10 Hz, which may be caused by EEG signals. The IRIS® almost matches
the MI systems for linearity, resolution, and other basic properties for purely
horizontal or vertical eye movements, but the MI system is superior for
oblique or torsional movements. However the MI system's invasive nature
(requiring a medical professional to administer a local anaesthetic to the eye
prior to attaching by suction the annulus in which the wire coil is embedded)
makes it far from ideal for clinical applications.

The main improvements of the IRIS® over existing reflection methods
lies in the use of chopped input and the application of an array of narrow beam
detectors and emitters. The somewhat limited dynamic range is due to changes
in pupil diameter giving rise to signals in the laterally positioned detectors
beyond 30 degrees of visual angle to the left or right. At this point natural
pupillary oscillations, which vary with subjects, may interfere with recordings.
Eyelid margins can also obscure the limbus in the superior and inferior
positions making vertical recordings difficult. An area needing improvement is
that of simultaneous recording of both vertical and horizontal eye movements
and correction for the crosstalk.

*Calibration of IRIS® Recordings*

Stampe (1993) in outlining a reliable calibration method for video-
based pupil-tracking systems provides a good summary of calibration methods
applied to eye trackers in general. He reported that converting eye-tracker data
to screen coordinates requires a mapping function that determines how
distortions between screen and tracker data are corrected. The coefficients of
the mapping function are computed through a calibration process in which a
set of known position targets are displayed and their eye-tracker position data
recorded. The choice of the mapping function determines the number of
calibration points required: anywhere from 3 points for a non-linear one-
dimensional calibration to 25 points for an extreme example of piecewise
linear calibration. The average mapping error is a U-shaped function of the
number of points used for calibration. A low number necessitates the use of a simplistic mapping function, that may not be able to correct all distortions, whereas with more points the special noise caused by inexact fixation of calibration points by the subject increases. McConkie (1981) suggested repeating each calibration point several times and using mean recorded position, but the subject can become habituated, and the longer time increases the likelihood of head movement during the calibration trials.

Stampe (1993) reported that the literature addresses two types of mapping function for 2D data, piecewise and non-linear. The piecewise divides the screen into a grid of cells with a calibration point presented at each grid junction. The tracker then defines a grid of quadrilaterals, each of which needs to be separately mapped back onto its original rectangular grid cell (Kliegl & Olsen, 1981; McConkie, 1981). Problems arise in that abrupt changes in scaling and distortion may occur at the boundaries between grid cells, and although adding more cells to the grid reduces the changes, more calibration points need to be collected.

Smooth changes in mapping across the screen can be achieved by non-linear mappings. The most common non-linear mapping is the biquadratic introduced by Sheena and Borah (1981). This requires 9 calibration points, which is close to the optimum number for minimum mapping error (Stampe, 1993). The non-linear terms of the function allow smooth scale changes to occur across the screen and can correct curved distortions. Small errors in calibration-point fixations or head movements can result in large errors in screen gaze position near the edges of the screen as squared terms in the equations become very large near the edges of the screen. Hence, it is
suggested that the area outside the calibration grid not be used for test targets.
Stampe presented the centre calibration point again at the end of the
calibration trial, and used this final fixation to adjust for any head movement
that may have occurred during the calibration session.

Although Stampe (1993) was reporting on the use of the bi-quadratic
method for video-based pupil-tracking systems, other have adapted the method
for use with infrared limbus trackers. For example, Carmody, Kundel, and
Nodine (1980) used two polynomials (quadratic) and a 25 point calibration
grid, whereas Nodine, Kundel, Toto, and Krupinski (1992) used a similar
mapping function and a 9 point calibration grid. Both groups supplemented the
pre-trial calibration grid with a 5 point post-test grid (corner points and centre)
to check for gross head movement.

**Summary**

Although the IRIS® is a more accurate eye movement measuring
system free of biological interference, the EOG system remains a versatile
technique for recordings greater than ±1 degree of visual angle and is easily
operated and portable (Oster & Stern, 1980). Despite a few disadvantages, the
EOG still provides an inexpensive, accurate, and reliable method of measuring
slow-phase eye movements (Schmid-Priscoveanu & Allum, 1999) and
saccades.

Although there is considerable research being conducted on eye
movements and their diagnostic potential, most recordings have been
restricted to the horizontal dimension. For example, of the 109 patient studies
in the Levy et al. (1993) review, the authors report only two vertical eye
movements in their results. Research by Previc (1982,1990), Previc and
Blume (1993), and Previc, Breitmeyer, and Weinstein (1995) has indicated that visual asymmetries, in relation to pattern recognition, functional specialisation, search, and discriminability, exists between the left, right, upper, and lower visual fields. This suggests that the ability to accurately record and analyse 2D eye movements may provide additional diagnostic information on pathologies, and further insights into the nature of brain functioning.

The limitations of the EOG do, however, present problems for humans and simple algorithms in scoring EOG signatures, hence the proposal to investigate artificial neural networks (ANNs) in respect of their pattern recognition abilities. ANNs may be able to overcome a lot of the limitations inherent in the EOG, as well as dealing with the problems introduced with 2D recordings such as different signal strength across channels and crosstalk (also a problem with IRIS®), thus allowing EOG to be extended to more 2D applications. As monitoring eye movements remains clinically important in the diagnosis of diseases of the central nervous system (Chen, Tsai, & Luo, 2000), application of ANNs to EOG should be beneficial to those researchers who continue to develop EOG and other systems.
CHAPTER 3

Artificial Neural Networks (ANNs)

The ANNs in this research were used as calibrating tools and no attempt was made to specialise or modify the networks other than parameter choices available in the Matlab® Neural Network Toolbox software. This chapter is designed to provide a basic understanding of ANNs. Chapter 4 will look at the application of ANNs to the calibration of 2D eye movement data. Rumelhart and McClelland (1986) provide a good introduction to the origins and basics of ANNs, and unless otherwise referenced, the material in this chapter is a summary of their relevant sections.

Origins of Parallel Distributed Processing

Jackson (1958) and Luria (1966), unique neurologists for their times, were two of the earliest pioneers. Jackson argued convincingly for multi-level conceptions of distributed processing and was a strong critic of late nineteenth century neurology's simplistic localisationist doctrines. Luria, a Russian neurologist and psychologist, proposed the notion of a dynamic functional system.

Hebb’s (1949) speculations about neural functioning captured the flavour of parallel distributed processing mechanisms, and this included the synaptic modification rule known as Hebb's rule. This stated that connection weights to units that were highly active should be strengthened and those inactive should be weakened. An important advance over the Hebb rule came from Rosenblatt (1959) with his perceptron convergence procedure. His vision
of a dynamic, interactive, and self-organising system of human information processing was at the core of the parallel distributed processing (PDP) approach.

Some more recent significant contributors are Feldman and Ballard (1982) who, under the name of connectionism, developed many of the computational principles or the PDP approach and also stressed the biological implausibility of most of the existing artificial intelligence computational models. Sutton and Barlo's (1981) insightful analysis of the delta rule for connection strength modification, provided crucial technical insights, and also accounted for some of the more subtle properties of classical conditioning, while Hopfield (1984) contributed the idea of network models seeking minima in energy landscapes.

**Background and Theory**

PDP models assume that the processing takes place through the interactions of a large number of units, each a simple processing element, which send either excitatory or inhibitory signals to each other (Figure 6). One reason for their appeal is their structural similarity to the brain ("physiological flavour"), which also has a large number of highly interconnected elements sending similar type signals to each other. They have changed the way of thinking about the time-course of processing, mechanisms of learning and the nature of representations.

Most other models store the pattern itself so that the representation in long-term memory and working memory are really the same. In PDP models, patterns are not stored directly but rather the connection strengths are modified and it is the encoding within the final weights that allows the pattern to be
Figure 6. A multi-layer network. The information contained in the input pattern is recoded into an internal representation by the hidden units. The internal representation generates the outputs, rather than the original pattern. Each input unit is connected to each unit in the hidden layer. Each unit in the hidden layer is connected to each unit in the output layer. There is a weighting associated with each connection.
recreated. This means that in processing, knowledge is not retrieved and brought to bear, but necessarily influences the course of processing, as it is actually part and parcel of the processing itself. There are also profound implications for learning. If knowledge is the strength of connections then learning is finding the right strengths. So it is then possible, as a result of tuning its connections, for the mechanism to learn the interdependencies between activations to which it is exposed during the course of processing.

Most computational approaches to learning assume that the goal is rule formulation, which captures powerful generalisations in a succinct manner. The PDP approach has connection strength modulation mechanisms, which work with information available locally at the connection. This allows the acquisition of a set of connection strengths that allow the network of simple units to act as though it knows the rules. There is also no reason to have the knowledge about a particular pattern stored in the connections of a single special unit reserved for that pattern. Rather, the pattern can be stored in a distributed fashion over a large number of processing units.

A General Framework for PDP Models

There are eight major aspects of a PDP model. These are: a set of processing units, a state of activation, an output function for each unit, a pattern of connectivity among units, a propagation rule for propagating patterns of activities, an activation rule for combining inputs on a unit with the current state to produce a new level of activation, a learning rule for using experience to modify the pattern of connectivity, and an environment within which the system operates.
Each unit (Figure 7) at a point in time has an activation value that is passed through a function to produce an output value that is then passed through a set of unidirectional connections to other units. Each connection has a weight or strength associated with it which determines how much the first unit effects the second. Some operator (usually addition) combines all inputs and another function (the activation rule) uses this combination and the current activation level to arrive at a new activation value. Furthermore, this activation is not a fixed function of the inputs but can undergo modification as a function of adjusting the connection weights.

*A Set of Processing Units*

In some models the units may represent particular conceptual objects; in others though, they are simply abstract elements over which meaningful patterns can be defined. There is no overseer; all the processing is carried out by the relatively simple units, whose job it is to receive input from its neighbours and to compute its output value to send to its neighbours. As many units can carry out their computations at the same time, the system is inherently parallel. Input units are those that receive input from sources external to the system, whereas output units send signals out of the system. Units whose inputs and outputs are within the system are the hidden units.

*The State of Activation*

This is specified by a vector representing the pattern of activation over the set of units at time t, and it is this pattern that captures what the system is
Figure 7. An individual node or local computational element from a network. The node forms a weighted sum of the inputs (rule of propagation) and passes the result through a transfer function (activation rule). A bias, b, may be added to the weighted sum ($\Sigma w_i x_i + b$), which shifts the transfer function to the left as shown in the final diagram. Five representative transfer functions are shown.
representing at any time. Each element of the vector stands for the activation of one unit. Activation values may be continuous or discrete (e.g. binary).

*Output of the Units*

Units interact by transmitting signals to each other. The strength of the units’ signal, and therefore the effect on their neighbours, depends on their degree of activation. Each unit has an output function that maps the current state of activation to an output signal. The output function may be the identify function but more often it is some sort of threshold function (see Figure 7) that only allows the unit to effect another unit if its activation exceeds some certain value.

*The Pattern of Connectivity*

The pattern of connectivity determines what the system knows and how it will react to an arbitrary input. Specifying a processing system and its encoded knowledge is a matter of specifying this pattern of connectivity. In many cases the total input to a unit is the weighted sum of the separate inputs and in this case specifying the weights for each of the connections in the system can represent the total pattern of connectivity. Positive weights are excitatory and negative weights are inhibitory. A weight matrix can represent all these. If more complex combination rules are required then there is usually a separate connectivity matrix for each type of connection. Whether a system is top-down or bottom-up or hierarchical are issues of the nature of the connectivity matrix. The fan-in and fan-out of a unit determines how much serial processing a network must perform and also how much information can be stored. The fan-in of a unit is the number of excitatory or inhibitory
elements that feed into it whereas the fan-out of the unit is the number of units that it directly affects.

*The Rule of Propagation*

The rule of propagation takes the outputs of the units connected to a particular unit and combines them with the connectivity matrix to produce a net input for that unit. It is usually the weighted sum of the inputs to the unit ($\sum w_{ij}x_i$). Sometimes a bias (see below) is also added to the weighted sum ($\sum w_{ij}x_i + b$).

*Activation Rule*

The activation rule combines the net inputs impinging on a particular unit with the current state of the unit to produce a new state of activation. This rule is sometimes a threshold function so that the net input must be of a certain value before it can contribute to the new state of activation. The most common types are the quasi-linear activation functions that are non-decreasing functions of a single type of input. They are called semi-linear if they are differentiable (a useful constraint). They may also be non-linear and differentiable, that is, sigmoid. The addition of a bias, $b$, to the weighted sum which forms the input to the activation rule, shifts the function to the left as seen in Figure 7. A bias may be used to selectively inhibit the activity of a certain neuron by shifting the ‘active range’ of the activation function to the right of its current input value (Schalkoff, 1997). A bias may also be used to shift the ‘active range’ of the activation function into the current range of neuron inputs.
Modifying Patterns of Connectivity as a Function of Experience (Learning Rules)

Learning rules can consist of developing new connections, losing old ones, or modifying the strengths of those already existing. The first two are really a special case of the third. All learning rules are really a variant of the Hebbian learning rule (Hebb, 1949). Hebb's basic idea being that if two units are both highly active then the weight between them should be strengthened. This idea has been modified and extended so that it can now be more generally stated that the change in connection between two units, \((u_i \text{ to } u_j)\), is the product of a function \(g(\ )\) of the activation of \(u_j\) (\(a_j\)) and its teaching input (\(t_j\)), and another function \(h(\ )\) of the output value of \(u_i\) (\(o_i\)), and the connection strength \(w_{ji}\).

\[
\Delta w_{ji} = g(a_j(t),t_j(t)) \cdot h(o_i(t),w_{ji}) \quad (3-1)
\]

In the simplest versions when there is no teacher, the functions \(g\) and \(h\) are simply proportional to their first arguments and so

\[
\Delta w_{ji} = \eta a_j o_i , \quad (3-2)
\]

where \(\eta\) is the constant of proportionally representing the learning rate.

Another variation is the Widrow-Hoff rule. This is also called the delta rule, as the amount of learning is proportional to the difference (or delta) between the actual activation and the target activation. For this case \(h(o_i(t),w_{ji}) = o_i(t)\) and \(g(a_j(t),t_j(t)) = \eta(t_j(t)-a_j(t))\) so the delta rule becomes

\[
\Delta w_{ji} = \eta(t_j(t)-a_j(t)) \cdot o_i(t) \quad (3-3)
\]
This is a generalisation of the perceptron learning rule for which the perceptron convergence theorem has been proved. The perceptron and the perceptron convergence theorem are discussed in more detail below in the section on important developments.

Representation of the Environment

The environment is represented as a time-varying stochastic function over the space of input patterns, which means that at any particular moment in time there is some probability that any one of the input patterns is impinging on the input units.

Some Important Developments

Simple Linear Models

In the simple linear models the activation values can be any real positive or negative numbers. The output function is equal to the activation level and there are only two types of units, input and output. There is no need for hidden units, since in a linear system there is no computation resulting from multiple steps that cannot be computed in a single step. Any unit in the input layer may connect to any unit in the output layer. The new activation value is simply the weighted sum of the inputs.

The linear model with the simple Hebbian rule is the linear associator. This system has two sources of inputs: the input patterns establish a pattern of activation on the input units, and the teaching units that establish a pattern of activation on the output units. When the learning rule has been applied after each set of inputs, the system is then tested by presenting an input pattern and seeing how close the output pattern matches the original teaching input. If the
input patterns are orthogonal there is no interference between them and the relevant associated patterns can be perfectly reproduced on the output layer. If the input patterns are not orthogonal, there will be interference among the input units. However if the inputs are not orthogonal but remain linearly independent, then with an error-correcting rule, it is possible to build up correct associations. This error-correcting rule is the delta rule. What is essentially learnt is how to minimise the difference between the desired response and the output actually attained. However, it may take many presentations of the input pattern set before the correct associations can be reproduced.

Non-linear Units and Multiple Layers

The use of nonlinearities can overcome some of the limitations of the pure linear systems. For example, the exclusive-or function (XOR) cannot be computed with a linear system as the inputs are not linearly independent (the desired output is a [1] if the input is either [0 1] or [1 0], or [0] otherwise). However this XOR function can be computed with two layers of linear threshold units. Linear threshold units are the simplest of the non-linear units (Figure 7). They are a binary unit with value 1 if the weighted sum is greater than some threshold, and 0 otherwise. Multi-layer systems of such units are very powerful and are capable of computing any Boolean function. Unfortunately, for this general case, there is no known learning algorithm to determine the necessary weights.

There is however a well understood learning algorithm for the special case of the perceptron. The perceptron is essentially a single layer network of linear threshold units. In fact, the perceptron convergence theorem guarantees
that if a pattern set is linearly separable, then the perceptron learning algorithm will find the correct set of weights in order to respond correctly to all input patterns. Unfortunately, the single layer perceptron for which there is a guaranteed learning rule, cannot learn any patterns that are not learnable by the linear associator. The limitations of what the perceptron could learn was what led to Minsky and Papert’s (1969) pessimistic evaluation of the perceptron, which, unfortunately and incorrectly, tainted the more interesting and powerful networks of non-linear units. Further developments however, have led to the generalised delta rule, which is capable of learning arbitrary mappings. It does not work for linear threshold units, but does work for semi-linear, i.e. differentiable activation functions.

**Backpropagation - Learning internal representations by error propagation**

Simple two-layer associative networks have only input and output units and so have no internal representations. What allows them to make reasonable generalisations and perform satisfactorily on unseen patterns is that they map similar input patterns to similar output patterns. A problem then arises when the representations provided by the outside world have very different structures for the input and output patterns. In this case, a network without hidden units will be unable to perform the mappings. This is the situation with the XOR problem, in which the patterns which overlap least, are the ones required to generate identical outputs. Minsky and Papert (1969), besides highlighting the limitations of perceptrons, also point out that if there is a layer of simple hidden units with non-linear activation functions, any required mapping of input to output units is possible. Information coming to the input units can be recoded into an internal representation by the hidden
units, which can then generate the required outputs. With the XOR problem this can be done with a single hidden unit.

The problem still remains that although there is a very simple guaranteed learning rule for problems solved without hidden units, there is no such rule for networks with hidden units. There have been three approaches to this basic problem. The first approach was to use simple unsupervised learning rules, which showed promise but had no external force to insure that the hidden units required for the appropriate mapping would be developed. The second approach assumed an internal representation, which on some a priori grounds seemed reasonable. This approach has been taken in the interactive activation model of word perception (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). The third approach was to attempt to develop a learning procedure that can perform the task. One such attempt, which is limited to symmetric networks, uses stochastic units and requires the network to reach equilibrium in two different phases (Boltzmann machine). The next section presents another alternative, which utilised deterministic units only involving local computations. This method is a clear generalisation of the delta rule and is called the generalised delta rule. Parker (1985) independently developed a similar rule that he called learning logic.

The Generalised Delta Rule

The learning procedure associated with the standard delta rule involves the presentation of a set of pairs of input and output patterns. If there is no difference between the output vector that the input pattern produces and the output pattern of the pair, then no learning takes place. Otherwise, the weights are changed such that they reduce the difference.
For linear units this learning rule minimises the squares of differences between actual and desired output values, summed over the output units and all pairs of input/outputs. This implements a gradient descent in sum-squared error. If weights are changed after each pair then there will be, to some extent, a departure from a true gradient descent. However, if the learning rate is kept sufficiently small, this departure will be negligible. With linear activation functions and no hidden units, the error surface is shaped like a bowl with only one minimum, so gradient descent is guaranteed to find the best set of weights. With non-linear activation functions and hidden units the error surface is not concave upwards and may contain local minima.

For layered feedforward networks which can have many hidden layers, but in which each unit must send its output only to higher layers and receive its input only from lower layers, the generalised delta rule works if the network is made up of units having semi-linear activation functions. It takes exactly the same form as the standard delta rule, with the weight change on each line being proportional to the product of an error signal, $\delta$, available to the unit receiving input on that line, and the output of the unit sending activation along that line (Figure 8).

$$\Delta_p w_{ji} = \eta \delta_{pj} o_{pi}$$  \hspace{1cm} (3-4)

where $\Delta_p w_{ji}$ is the change to be made to the weight from the $i^{th}$ to the $j^{th}$ unit following presentation of pattern $p$, and $o_{pi}$ is the $j^{th}$ element of the output pattern produced by the presentation of pattern $p$. 
Change in weight for the a-b line
\[ \Delta_p w_{ba} = \eta \delta_{pb} o_{pa} \]
where
\[ \delta_{pb} = f'_b (\text{net}_{pb}) (\delta_{pc} w_{cb} + \delta_{pd} w_{db} + \delta_{pe} w_{eb}) \]
and \( \text{net}_{pb} \) is the weighted input into unit b from the three input units.

Change in weight for the b-c line
\[ \Delta_p w_{cb} = \eta \delta_{pc} o_{pb} \]
where
\[ \delta_{pc} = (t_{pc} - o_{pc}) f'_j (\text{net}_{pc}) \]
and \( \text{net}_{pc} \) is the weighted input to unit c from the two hidden units.

**Figure 8.** A simple backpropagation network with three input units, three output units and two hidden units. An example of the application of the generalised delta rule is given for the connection weights on the line between output unit c and hidden unit b and for the line between hidden unit b and input unit a.
The determination of the error signal, $\delta$, is a recursive process that
starts with the output units. If the unit in question is an output unit, its error
signal involves the derivative of the semi-linear activation function, $f()$, that
maps the total input of the unit to its output i.e. $f_j'(\text{net}_{pj})$, where $\text{net}_{pj}$ is the net
input into unit $j$, and $t_{pj}$ is the target input for the $j^{\text{th}}$ component of the output
pattern from pattern $p$.

$$\delta_{pj} = (t_{pj} - o_{pj}) f_j'(\text{net}_{pj}). \quad (3-5)$$

For non-output units (hidden units where there is no specified target) the error
signal is determined recursively in terms of the error signals related to the
units to which it is connected, and the weights of those connections.

$$\delta_{pj} = f_j'(\text{net}_{pj}) \sum_k \delta_{pk} w_{kj} \quad (3-6)$$

Applying the generalised delta rule requires two steps. The first is
presenting and propagating the input forward through the network to compute
the output value for each unit, $o_{pj}$. The second step is a backward pass through
the network during which the error signal is passed to each unit and the
necessary weight changes are made. This involves calculating the error signal
for each output unit by multiplying the difference between the actual and the
expected outputs by the derivative of the activation function. For example, in
Figure 8 the error signal for output unit $c$ ($\delta_{pc}$). will be the product of, the
difference between the actual and expected outputs of unit $c$ ($t_{pc} - o_{pc}$), and the
value of the derivative of c’s activation function acting on its net input (weighted sum) from both hidden units. The weight changes for all connections feeding into the final layer can then be calculated. In Figure 8 the weight change for the connection between unit b and unit c will be the product of the learning factor (\(\eta\)), the error term coming down the line from c (\(\delta_{pc}\)), and the output coming up the line from unit b (\(o_{pb}\)).

The error terms for units in the penultimate layer can now be computed and so propagate the errors back one layer. In Figure 8, as hidden unit b does not have any specified output target, its error signal is the product of the derivative of its activation function acting on its net weighted input from the three input units, and the sum of the products of the \(\delta\)'s from output units d and c, and the original weights along the lines to those units. The process can then be repeated for any number of layers. This is gradient descent for feedforward networks with semi-linear units.

**The Sigmoid Function and Improving Learning**

As the linear threshold function, on which the perceptron is based, is discontinuous, it is not suitable for the generalised delta rule. A more useful activation function is the sigmoid function

\[
O_{pj} = \frac{1}{1 + e^{-\left(\sum_i W_{pi} O_{pi} + \theta_j\right)}}
\]  

(3-7)

Its error signal for an output unit is

\[
\delta_{pj} = (t_{pj} - o_{pj}) o_{pj}(1-o_{pj})
\]  

(3-8)
and for a hidden unit

\[ \delta_{pj} = o_{pj}(1-o_{pj}) \sum_{k} \delta_{pk}w_{kj} \]  

(3-9)

The derivative, \( o_{pj}(1-o_{pj}) \), reaches its maximum for \( o_{pj} = 0.5 \) and, since \( 0 \leq o_{pj} \leq 1 \) (Figure 7), approaches its minimum as \( o_{pj} \) approaches zero or one. This means that weight changes will be largest for those units that are near their midrange of activation, and as such, not really committed to being either on or off. This feature probably contributes to the learning stability of the system. As the system cannot really reach its extreme values of 1 or 0 without infinitely large weights, values of 0.9 and 0.1 are typically used as target output values.

Although true gradient descent requires infinitesimal steps, for practical purposes, a learning rate that is as large as possible without causing oscillation, should be used. Oscillation occurs when the system jumps from one side of an error surface valley to the other, without descending to the minimum lying below. One way to increase the learning rate without leading to oscillation is to modify the learning rule to include a momentum term.

\[ \Delta w_{ji}(n+1) = \eta(\delta_{pj}o_{pi}) + \alpha \Delta w_{ji}(n) \]  

(3-10)

where \( n \) indexes the presentation number, \( \eta \) represents the learning rate, and \( \alpha \) the momentum parameter. This parameter is a constant which determines the effect of past weight changes on the current direction of movement in weight space. This additional term in the weight change rule
simulates a momentum effect in weight space, which effectively filters out high-frequency variations of the error surface. It is possible to get similar solutions with $\alpha=0$ and reducing the size of the learning rate, however the system learns much faster with larger values of $\alpha$ and $\eta$. If a solution requires unequal weights but the system starts with them equal, it can never learn, as error is propagated back in proportion to the values of the weights. Starting with random weights can counteract this problem.

Application Performance

Rumelhart and McClelland (1986) describe some simulation results for several common problems, including the XOR problem previously described, and the parity problem, where the output required is 1 if the input pattern contains an odd number of 1's, or 0 otherwise. The encoding problem: where the task is to map a set of orthogonal input patterns to a set of orthogonal output patterns through a small set of hidden units. The symmetry problem: which involves classifying a set of input patterns as to whether they are symmetric about their centre. The addition problem: solving simple binary addition. The negation problem: which has $n+1$ inputs and $n$ outputs, with one input a negation bit indicating that the complement should be output if the negation bit is on. The T-C problem: which involves discriminating between a ‘T’ and a ‘C’ regardless of the letter’s translation or rotation. In all these applications it was found that local minima were very rare and that the system learnt in a reasonable length of time.

Conclusion

Minsky and Papert's (1969) challenge to find a learning mechanism for the multi-layer networks, appears to have been answered by finding an error
propagation scheme that, although it does not guarantee it, leads to solutions in a wide range of problems.
CHAPTER 4

Application of ANNs to the Measurement of 2D Saccadic EOG.

As stated in the introduction, EOG is usually restricted to horizontal recordings, as simultaneous recording of both horizontal and vertical EOG introduces other difficulties that create problems in relating the 2D EOG to actual rotations of the eyes. In Chapter 3 it was mentioned that MLPs with non-linear units are capable of learning any mapping, and hence may provide a solution to the problems encountered in mapping 2D EOG signatures into 2D eye positions.

Translations Inherent in Recording 2D EOG

To determine point of gaze from EOG recordings, voltages recorded in the horizontal and vertical channels of the EOG need to accurately reflect the amount of eye rotation necessary to fixate a target of given horizontal and vertical spatial coordinates. In addition to the physiological sources of noise inherent in the nature of the EOG signal (muscle activity, brain potentials, and skin potential response), there are several geometrical sources of distortion that can occur in the transformation between the screen coordinates of the target and the recorded voltages in the horizontal and vertical channels of the EOG. This is also true of all other methods relying on fixed head recordings with subsequent resolution into horizontal and vertical components.

Individuals have different corneo-retinal potentials, facial structure, and skin conductivity characteristics. These factors will give rise to a scaling problem across subjects, that is, the same change in eye position for two
different subjects will result in different recorded voltages. A further difference in the gains will result if the electrode placements are such that the recording axis remains in the correct position (Figure 9a) but the electrodes are displaced further from the eye (horizontal axis in Figure 9b). As a result of electrode placement the recording axis may also have undergone a translation and no longer pass through the rotational centre of the eye (vertical axis in Figure 9b). Any combination of these two effects is also possible and will result in the electrodes recording a smaller potential as they are further removed from the source of the electric field. Since physiological noise will be essentially constant, this will result in a reduced signal to noise ratio. Thus, individual anatomical differences and the placement of electrodes will affect the signal and result in different gains for each channel both within and across subjects.

Other transformations may occur as a result of the recording channels being non-orthogonal to each other, or being misaligned to the true vertical and horizontal with respect to the primary eye position, or some combination of both. These rotation transformations can be understood by considering the angle through which each channel has independently been rotated away from its true alignment (see Figure 9c). These rotations will produce crosstalk within the recorded channels. A pure vertical eye movement, for example, could result in voltages being registered in both vertical and horizontal channels. Placement of the electrodes to provide correctly aligned, orthogonal recording channels is nearly impossible. Even great care in electrode placement will not necessarily guarantee correctly aligned, orthogonal recording channels, as there can exist a lack of homogeneity in facial tissue
Figure 9. Electrode placement and recording channels. (a), electrode placement is optimal, resulting in horizontal and vertical recording channels that are both orthogonal and aligned with the true horizontal and vertical directions. (b), the recording channels are still orthogonal and in true alignment, however the placement of the vertical electrodes produces a translation of the vertical recording axis resulting in a reduced signal strength. The placement of the horizontal electrodes also produces a reduction in signal strength due to their non-proximity to the eye. (c), the channels are neither orthogonal nor in true alignment. The angles \( \alpha \) and \( \beta \) represent the effective clockwise rotations that the horizontal and vertical channels receive as a result of the electrode placements shown. This results in reduction of signal strength as well as crosstalk between the recording channels. (d), \( \hat{\mathbf{h}} \) and \( \hat{\mathbf{v}} \), represent the unit veridical recording coordinate system, and \( \mathbf{h}^\prime \) and \( \mathbf{v}^\prime \) represent the actual recording coordinate system. \( O_H \) and \( O_V \) represent the horizontal and vertical displacements, \( a \) and \( b \) the horizontal and vertical gain factors, and \( -\alpha \) and \( \beta \) the horizontal and vertical rotations imposed on the veridical coordinate system as a result of the use of electrodes.
and thus in the distribution of the potential field from the eyeballs (Shackel, 1967). Accurate methods for calibrating 2D eye movements using EOG will necessarily need to be able to rectify these geometrical transformations.

Figure 9d represents the change in the signal amplitude in each of two components (horizontal and vertical) due to the transformations discussed above. An eye movement resolved into \( \mathbf{h} \) and \( \mathbf{v} \), the orthogonal unit direction vectors of the veridical coordinate system, represents the true horizontal and vertical rotation of the eye from its primary position. With the head held in a natural erect position, the eyes are in their primary position when the lines of sight are horizontal and perpendicular to a line joining the approximate centre of rotation of each eye (Carpenter, 1988). The displacements, \( O_H \) and \( O_V \), represent a change in the horizontal and vertical components of the signal, respectively, due to electrode placement, variations in skin or tissue conductance, or some combination of both. The actual coordinate system used is represented by \( \mathbf{h}' \) and \( \mathbf{v}' \), with rotations \( \alpha \) and \( \beta \) due to electrode misalignment. As well, this coordinate system does not possess the same ‘gain’ in each direction, and this in turn differs from the unity gain in the veridical coordinate system.

The following equation describes the transformations that may occur as a result of non-optimum placement of electrodes.

\[
\begin{pmatrix}
\mathbf{h}' \\
\mathbf{v}'
\end{pmatrix} = \begin{pmatrix}
\cos(\alpha) & \sin(\alpha) \\
-\sin(\beta) & \cos(\beta)
\end{pmatrix} \begin{pmatrix}
(a.(h + O_H)) \\
(b.(v + O_V))
\end{pmatrix}
\]

(4-1)
where $h$ and $v$ are the horizontal and vertical signals that would be recorded by the veridical recording axes, $\hat{h}$ and $\hat{v}$, which intersect at the centre of the eyeball and allow accurate measurement of the true rotation of the eye in both dimensions. The differences in gain in each direction are represented by $a$ and $b$, for the horizontal and vertical directions respectively. The transformed signals, $h'$ and $v'$, represent the actual signals that are recorded at the horizontal and vertical electrodes respectively.

There are then a number of sources of error, including crosstalk, that complicate signal interpretation. The goal of the ANNs was to find a function that would overcome these confounding influences by finding regularities in their occurrence.

*Calibration of 2D EOG*

The usual method of determining eye position from EOG based on amplitude differences of the EOG voltage before and after a saccade (Shackel, 1967; Oster & Stern, 1980), already difficult with the physiological noise contained in the EOG, become even more a challenge when applied to two dimensions. The extraction of the saccadic displacement becomes problematic when dealing with vertical EOG due to the presence of rider or eyelid artefact (Oster & Stern, 1980). This artefact is schematised in Figure 4c. Blinks (Figure 4b) also have to be eliminated from the EOG before calibration can be attempted.

In order to match the EOG to the correct direction and displacement of the eye movement, it would first be necessary to extract the voltages that represent the true horizontal and vertical eye movements from the $h'$ and $v'$ recordings of the EOG. Thus, an inverse transformation is required to undo the
effects of the rotation and gain difference transformations that are contained in the two EOG channels (Equation 1). Given that these transformations of the EOG signal will be generally problematic, some signal processing method for computing and applying the inverse transformation automatically is required.

*Artificial Neural Networks (ANNs)*

Feedforward ANNs trained with back propagation (BP) are recognised for their pattern recognition capabilities and their computational powers. Lee (1994) reports that the multi-layer perceptron (MLP) with only one hidden layer using arbitrary sigmoid activation functions are universal approximators, and with sufficient hidden units are capable of arbitrarily accurate approximation to arbitrary mappings between given input/output sets. That is, they are usually capable of mapping a given set of inputs onto a given set of outputs. Linear perceptrons (LPs) can also be trained with BP but have no hidden layers and linear activation functions and hence are limited to performing only linear mappings. Thus, feedforward networks trained with BP, especially the MLP, should have the requisite computational ability to extract eye position from 2D EOG recordings. Comparing the performance of the LPs with the MLPs is also of interest to investigate the complexities of the distortions present in the EOG recordings. Any non-linear transformations should be dealt with more effectively by the MLPs.

Following most calibration tasks, the 2D EOG waveforms would correspond to saccades made from a central fixation point to targets of known coordinates. The ANN is trained on a portion of these waveforms to learn a mapping from the EOG signatures to eye position. Generalisation, the ability of the network to perform correctly on inputs not in the training set (Schalkoff,
1997), is tested on the remaining waveforms. Figure 10a displays typical horizontal EOG waveforms of saccades to three different horizontal targets. Figure 10b represents the general architecture of a MLP designed to train on these EOG waveforms. The number of input nodes will depend on the number of data points chosen to represent the saccadic waveforms. This could be the entire set of raw data points sampled during the saccade or perhaps some subset of these points. The two output nodes of the network provide the horizontal and vertical 2D spatial coordinates of the target on a particular training or test trial. For a given viewing distance, provided that the subject accurately fixated the target, these coordinates represent the position of the eye in its orbit. The training vector for a particular saccade would consist of the actual horizontal and vertical spatial coordinates of the target for that trial.

In designing a MLP an important consideration is the number of hidden nodes in that too many and the training time increases due to additional computations with the risk of overfitting when the network trains on noise within individual trials (Prechelt, 1998). Overfitting improves the performance on the training set, however it decreases the network’s performance with new data. With too few hidden nodes, the network may either not have the computational power necessary to perform the mapping, or if the power is available, it may take excessive training time to converge on a solution (Golden, 1996). Experiment 1 was designed to provide an empirical estimate for a suitable number of hidden nodes to use with the backpropagation MLPs.
Figure 10. Figure (a) shows schematic characteristic waveforms resulting from saccades to horizontal targets at three different locations to the right of the fixation point, which the multi-layer perceptron (MLP) shown in (b) should be able to distinguish after training on the target coordinates. Saccade 1 was closest to the fixation point. W(i, j) represents the network weight from input neuron j to hidden neuron i, and W(k, i) represents the network weight from hidden neuron i to output neuron k.
CHAPTER 5

Experiment 1

*Determination of the MLP Hidden Layer*

The aim of Experiment 1 was to use subject data to determine an appropriate number of hidden nodes to be used with the backpropagation MLPs for calibrating the 2D saccadic eye movements. Subject data were gathered using a NeuroScan® EEG system. This system had its own display software to present the task, and 2D EOG was recorded using some of the available EEG channels.

**Method**

*Participants*

The five subjects were psychology students enrolled in a first year statistics course at Griffith University. They all had normal, uncorrected vision and received course credit for their participation. There were 2 males and 3 females (age $M = 18.6$ years, $SD = 1.4$ years).

*Materials*

*Electrodes and Hardware*

Disposable silver-silver chloride electrodes were used to record the vertical and horizontal EOG. The saccadic task was presented on a Labtam monitor located in a Faraday cage within the Human Psychology Laboratory at Griffith University. A programme written in GENTASK presented the stimuli for the task on a 486 PC. GENTASK is DOS-based
software produced by Neurosoft Inc., dedicated to displaying stimuli used in connection with their NeuroScan® EEG recording equipment.

_Stimuli_

The stimuli were a series of dots that appeared individually on a monitor screen 45 cm from the subject. At that distance a stimulus dot (target) subtended 0.4 degrees of visual angle. A fixation cross that subtended 0.6 degrees of visual angle along each arm of the cross was located at the centre of the screen. Targets appeared for a duration of 0.5 sec with a stimulus onset asynchrony (time between consecutive stimulus onsets) of 2.25s allowing fixation on the target followed by fixation on the central cross before the next target appeared. Since horizontal EOG is reported to have, under ideal conditions, a resolution of between 1 to 2 degrees of visual angle (Young & Sheena, 1975; Oster & Stern, 1980), 2 degree increments in the visual angle was considered a practical minimum difference to have between adjacent stimuli. At a viewing distance of 45 cm, a target near the top-centre of the screen was approximately 12 degrees of visual angle away from the centre of a fixation cross at the centre of the screen. This range of eye movements is common in tasks such as reading and viewing a computer monitor.

A curved screen would be necessary in order to ensure that 2 degrees of visual angle would transverse the same distance on the screen regardless of whether it occurs in the centre or periphery of the screen. With a monitor screen at 45 cm viewing distance, the relationship between distance on the screen (screen coordinates) and increasing visual angle is non-linear and can be seen in Figure 11. Although the non-linearity appears trivial up to 12 degrees of visual angle, screen coordinates of target stimuli were calculated
based on the non-linearly to ensure that the stimuli were at 2 degrees of visual angle increments along each of the meridians.

![Graph showing distance versus visual angle](image)

*Figure 11.* Distance on a flat screen as a function of increasing visual angle.

Stimuli were restricted to the upper right quadrant of the monitor providing a sufficient number of stimuli at the required resolution and avoided problems due to possible perceptual asymmetries in the visual field (Previc & Bloom, 1993; Previc, Breitmeyer, & Weinstein, 1995). Restricting stimuli to the upper right quadrant also minimised the non-linearity between degrees of visual angle and screen coordinates of target stimuli. It was important at this
stage of the research not to introduce additional variables into the calibration problem. Incorporating other quadrants using the same target density would have also made it impossible for a subject to complete the task in one session without excessive fatigue. The task took approximately 40 minutes to complete.

Clement (1991) reports that the kinematics of saccadic eye movements are best described as the eyes moving around fixed axes intersecting the centre of the eye. For this reason targets were arranged along meridians extending out from the central fixation cross. From the fixation point located centrally on the monitor (primary position of the eyes) the targets appeared along one of four possible meridians at 0°, 30°, 60°, or 90° above the horizontal (see Figure 12). Along each meridian were six possible positions, corresponding to between 2 degrees of visual angle and 12 degrees of visual angle from the central position at 2 degrees of visual angle increments. A different fixed axis of rotation enabled the eyes to make saccades along each of the chosen meridians. The direction of the saccade having been determined by the orientation of the axis of rotation, the displacement was then determined by the amount of rotation about this axis, and this would be reflected in the voltage of the EOG.

Thus, based on recommendations by Clement (1991), there were a total of 24 different locations for the targets. The 25th stimulus was the word “BLINK” located centrally on the screen. Each of the 25 stimuli was repeated 20 times at random, providing 500 trials for each subject to perform. Twenty repetitions were considered a reasonable compromise between the need for
repetitions required for training and testing the ANNs, and the consequent amount of time required for the subject to sit still and control their blinking and eye movements without considerable fatigue.

Figure 12. Approximate location of dot stimuli for the saccadic eye movement task. Subjects fixate on the fixation cross, located in the centre of the monitor and at the primary position for the eyes. Targets in the form of dots appear randomly, one at a time, at one of the possible locations shown. The locations are in 2 degrees of visual angle increments along each of the meridians.

Artificial Neural Networks

The MLPs used in the experiment were designed and run using the Matlab® Neural Network Toolbox version 2.0 (The MathWorks, Inc., 1994). Details of the MLPs are in Appendix A.
**Procedure**

*Subject Preparation*

Electrodes were placed at the outer canthi of both eyes, above and below the left eye, and on both ears. The ear electrodes were connected together and to the earth of the amplifiers to reduce the occurrence of DC drift. Horizontal EOG was a bi-polar recording from the electrodes placed on the outer canthi of both eyes. Vertical EOG was a bi-polar recording from the electrodes placed above and below the left eye. Subjects were seated in a comfortable lounge chair within the Faraday cage, with their heads resting against the back of the chair. No head restraint was employed. The Faraday cage was well lit hence no dark adaptation period was necessary.

Once subjects had the electrodes in place, the EOG signals were checked using the NeuroScan hardware to ensure impedances (coupling to the skin) were satisfactory. Subjects were also allowed at this time to practise blinking on command. Some subjects when asked to blink shut their eyes so tightly that the EOG signal was lost in the EMG signals generated by the eye and facial muscles. It was possible at this stage to monitor the subjects’ natural blinks, and provide feedback, so that their deliberate blinks were similar to their natural ones.

*Saccadic Task*

Subjects were instructed to fixate on the fixation cross in the middle of the screen and to make saccades to the stimulus dots, as they appeared, and then back to the fixation point. Subjects were asked to sit as still as possible, maintain their head position, and try not to blink in the time between fixating on the cross and making a saccade to a target. The task paused automatically
after each 125 trials to allow a brief rest. The subject used a mouse-button press to restart the task after they felt they were rested sufficiently. On presentation of the “Blink” stimulus, subjects were instructed to blink as naturally as possible.

Data Recording

The voltage differences at the electrode pairs were digitally sampled at a frequency of 500 Hz for a period of 50 ms prior to target presentation and for 950 ms from the onset of the stimulus, and saved to a binary file. At the end of the test period, the data file consisted of a matrix of numbers with each row representing the data for a particular trial (presentation of a stimulus). The first entry in the data file contained the accept byte which determined whether or not that trial was accepted (the Neuro Scan software allows the waveforms to be viewed and accepted or rejected). The second entry contained the trial type (a number designating the position on the screen where the stimulus appeared), and the remaining data contained the values resulting from the digitised sampling. One second of sampling at 500 Hz, less those points sampled before the onset of the stimulus, produced 474 data points.

Data Screening

The data were screened using a computer algorithm (clean_M.m). This was a Matlab® program that rejected those trials that contained EMG or other excessive noise, blinks and other unsolicited eye movements, or those trials in which the subject failed to initially fixate on the cross before making a saccade. As there was no interest in the return saccade, only the first 600 ms (300 data points of clean range) were required to be free of blinks or other interference to the saccadic signal. Visual inspection of pilot data for the
largest saccades indicated that their duration was less than 100 ms, and allowing 200 ms for the onset of slow regular saccades (Fischer & Weber, 1993) left 300 ms for fixation of the target. For details of the algorithm see Appendix B.

ANN Architecture and Training

Several subjects’ data were then previewed using the Neuro Scan software. The aim was to simplify the overall input to the network. The number of data points used to represent each trial was reduced to 200 from each channel; data points 41–241. These 200 data points were chosen, after viewing averaged waveforms, as best representing the saccadic eye movement across all saccades (see Figure 13).

![Figure 13](image)

*Figure 13.* The window indicates the approximate range of data points chosen to represent the saccadic eye movements. Data points 41–241 (approximately 80 - 480 ms after the stimulus onset). The example waveform is from the vertical channel and shows some eyelid artefact.

This was a somewhat arbitrary selection as it was only based on a visual impression of what was the best compromise interval to use, as the ideal...
interval varied from trial to trial (no computer algorithm was available at the time to extract the saccadic onsets). For each accepted trial, the 200 data points from each channel were concatenated to produce the input matrix. The architecture of the MLPs (see Figure 6) thus consisted of 400 inputs (200 to represent the saccade from each recording channel), 5, 10, 15, 20, or 25 hidden nodes, and 2 output nodes. The training matrix consisted of the 2D spatial coordinates of the targets corresponding to the accepted trials.

Feedforward MLPs containing 5, 10, 15, 20, 25, or 30 nodes in the hidden layer were systematically trained using backpropagation to determine a suitable number of hidden nodes for calibration of 2D saccadic EOG. Each training session consisted of 10 repetitions. Each repetition used training and testing sets obtained from a 50-50 random division of the data matrices, with each set containing equal numbers of each target. For each repetition there was random assignment of network weights and biases. Network performance was checked with the testing set after every 100 iterations of training. Each iteration consisted of batch training in which the weights and biases were only updated after all of the inputs had been presented. A weight configuration that resulted in a performance improvement $\geq 1\%$ was saved to disk. The trial number of the last update was used to assess the number of iterations to convergence (training time).

The closest-target estimate of calibration accuracy.

To determine whether the ANNs were successful in their calibration of the eye movements, the distances between the point defined by the computed output coordinates and each of the possible target coordinates were calculated. If the actual target was nearest, then the output was scored as correct.
Results

The effects of different numbers of hidden nodes on the calibration accuracy of the MLPs are shown in Table 1. The accuracies reported for each subject, and number of hidden nodes, are the mean accuracies of each of the 10 training repetitions. The mean accuracies across subjects suggest that 20 hidden nodes produce the best calibration of the 2D saccadic EOG. The mean of training iterations across subjects varied from a minimum of 2210 iterations for 25 hidden nodes to a maximum of 4256 iterations for 10 hidden nodes.

Table 1

*Means and Standard Error of the Means (n = 10) for Calibration Accuracy (percent correct) Using Different Number of Hidden Nodes in the Multi-layer Perceptrons.*

<table>
<thead>
<tr>
<th>Number of Hidden Nodes</th>
<th>30</th>
<th>25</th>
<th>20</th>
<th>15</th>
<th>10</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>35.4</td>
<td>2.0</td>
<td>25.0</td>
<td>4.3</td>
<td>30.4</td>
<td>3.0</td>
</tr>
<tr>
<td>S2</td>
<td>44.5</td>
<td>6.5</td>
<td>36.2</td>
<td>6.8</td>
<td>36.3</td>
<td>7.1</td>
</tr>
<tr>
<td>S3</td>
<td>20.2</td>
<td>6.3</td>
<td>19.9</td>
<td>5.4</td>
<td>36.8</td>
<td>3.8</td>
</tr>
<tr>
<td>S4</td>
<td>18.4</td>
<td>6.6</td>
<td>30.2</td>
<td>6.5</td>
<td>31.1</td>
<td>6.9</td>
</tr>
<tr>
<td>S5</td>
<td>29.8</td>
<td>5.6</td>
<td>16.8</td>
<td>5.6</td>
<td>29.9</td>
<td>5.6</td>
</tr>
<tr>
<td><strong>means</strong></td>
<td><strong>29.6</strong></td>
<td><strong>4.8</strong></td>
<td><strong>25.6</strong></td>
<td><strong>3.5</strong></td>
<td><strong>32.9</strong></td>
<td><strong>1.5</strong></td>
</tr>
</tbody>
</table>
Discussion

The results in Table 1 suggest that 20 nodes in the hidden layer of the MLPs provides the best calibration of the 24 different target types present in the saccadic eye movement task. Increasing the number of hidden nodes offers no improvement in calibration accuracy but does increase the complexity of the networks architecture and the possibility of overfitting. Although the complexity and risk of overfitting could be reduced with fewer hidden nodes, the calibration accuracy is also reduced.

The best mean calibration accuracy of 33% suggests that either MLPs may not be a very useful tool to calibrate 2D EOG, or that the data obtained contains too much noise to allow the best ANN performance to be realised. This experiment was designed primarily to determine a suitable number of hidden nodes for the MLPs. In subsequent experiments several design aspects were improved which were intended to decrease possible sources of noise or invalid eye movements. For example, the use of a head restraint when subjects are performing the eye movement task should produce a more valid EOG signal. The use of a computer algorithm to detect the saccadic onset within each trial’s data should produce a marked improvement in performance over the selection of data points based on average waveforms. Replacing the closest target estimate as the method for measuring calibration accuracy with a distance metric should also provide a better estimate of network performance. Closest target estimate pinpoints the closest target but fails to quantify the margin of error. This may be important as a relatively small error for an inner target may put the output closer to a neighbouring target, and hence be recorded as an incorrect calibration trial, whereas this may not be the case for
the same error made to an outer target (see Figure 12). A better performance measure would be to express network performance as an error distance (in degrees of visual angle) between the target location and the output coordinates from the calibrator.

If the data are at fault, for example, the selection of data points to represent the saccades is grossly in error, then this would explain the poor performance of the MLPs. On the other hand, if the data are satisfactory, then the results raise doubts as to the computational capabilities of the MLPs to perform the necessary inverse transformations required to calibrate the 2D eye movements from recorded 2D EOG. Two further experiments were conducted to address these issues.

Experiment 2 was a feasibility experiment using simulated 2D EOG signals. As these signals would be free of eyelid artefact and individual differences arising from human subjects they could be used to directly compare the ability of two ANNs and a simpler classifier to perform the inverse transformations required to calibrate simulated 2D data. In Experiment 3, the calibrators were then supplied with real saccadic EOG and IRIS® data. Any discrepancies in performance between the real and simulated EOG data should be a result of some combination of artefact or other errors introduced by individual subjects, rather than a failure of the ANNs to compute the necessary mapping between the eye movement signal and the target coordinates.
CHAPTER 6

Experiment 2

Simulation

EOG recordings of saccadic eye movements in 2D are affected by a number of factors. These may be generally grouped into physiological noise, variability in the point of fixation at both beginning and end of a saccade, channel gain differences, and crosstalk between recording channels. The crosstalk may be the result of non-orthogonal recording axes or the recording axes not being truly horizontal and vertical, or a combination of both. The purpose of this experiment is to address these factors using previously defined saccadic eye movement models to generate a series of data points simulating the signals that result from EOG recordings of 2D saccadic eye movements. Several simulated data sets were constructed containing simulated physiological noise and fixation errors and with varying amounts of gain differences and crosstalk between channels. This was done to directly compare the ability of two ANNs and a simpler classifier to calibrate such simulated data. The simple classifier was based on an algorithm developed to reflect the kind of manual scoring that a human operator might undertake in interpreting the data through visual inspection of the EOG signals.

Method

The ANNs were tested on data generated from mathematical models of saccadic eye movements. Sources of noise or error that have been documented in the literature were also simulated. These were the basic physiological noise
of the EOG and fixation errors. Transformation of the data as a result of electrode misalignment was also simulated. Parameters related to the differences in gain between the two recording channels and the amplitude of the physiological noise present in the EOG were based on the experimental data from Experiment 3. A saccadic simulation engine (SSE) produced horizontal and vertical components of saccades that varied, in 2 degrees of visual angle increments, from 2 degrees of visual angle to 12 degrees of visual angle, and occurred along the 0°, 30°, 60° or 90° meridians in the upper right quadrant of the visual field (see Figure 12 for a representation of target locations). This range of simulated 2D saccadic eye movements duplicated those necessary to perform the saccadic eye movement task in Experiment 1.

Saccadic Simulation Engine (SSE)

The parameters most commonly used to describe a saccade are the saccadic duration, displacement and peak velocity attained by the eye. These parameters are usually expressed in the form of the ‘main sequence’ relationship of the saccade (Bahill, Clark, & Stark, 1975). This relationship describes a function between saccadic displacement and saccadic duration, and between saccadic displacement and peak velocity. More recently, the skewness of the velocity profile has also been included as an important saccadic parameter. Skewness becomes more apparent with increasing displacement (Van Gisbergen, Van Opstal, & Ottes, 1984). For the purposes of the present simulation of saccadic eye movement data, the ‘main sequence’ relationships reported by Baloh, Sills, Kumley, and Honrubia (1975) gleaned from 25 normal subjects were used to generate the duration and peak velocity parameters for the range of saccadic displacements of interest (2 degrees of
visual angle to 12 degrees of visual angle). The parameter values of the displacement-skewness relation reported by Smit, Van Gisbergen, and Cools (1987) were used to compute the skewness of the velocity profile.

Having computed the duration, peak velocity, and skewness parameters, a velocity profile was generated for each saccade using the ‘gamma’ function adopted by Van Opstal and Van Gisbergen (1987). A saccadic onset was identified when eye velocity exceeded 30 degrees of visual angle /sec and was considered to end when eye velocity fell below this value (Van Opstal & Van Gisbergen, 1987). Hence, the velocity profiles were generated so that their duration, when measured at the 30 degrees of visual angle /sec criterion line, matched the value obtained from the ‘main sequence’ relationship. Numerical integration of the velocity profiles provided the position profiles.

Saccadic displacement was determined after consideration of the fixation error possible in making the saccade. The following equations and parameters from Baloh et al., (1975) were then used to determine the corresponding duration $d$ (msec) and maximum velocity $V_{\text{max}}$ (degrees of visual angle /sec) for a saccade with the given displacement $p$ (degrees of visual angle).

$$d = 2.7p + 37$$ \hspace{1cm} (6-1)

$$V_{\text{max}} = 551 \left[1 - \exp \left(-\frac{p}{14}\right)\right]$$ \hspace{1cm} (6-2)
The skewness value \( s \) of the saccade’s velocity profile was then calculated.

\[
s = 0.135p + 0.64	ag{6-3}
\]

The parameters in this equation were based on the mean of the parameters for the displacement versus skewness relationship reported by Smit et al. (1987) for subject JVG (Table 3 p.1757) when executing visible target saccades (V-saccades). Smit et al. also provide an equation describing the relationship between the skewness value and a shape parameter \( \gamma \) indicating the degree of asymmetry. Small \( \gamma \) values imply asymmetrical velocity profiles, with the velocity function assuming a symmetrical (Gaussian) shape as \( \gamma \) goes to infinity.

\[
s = \frac{2}{\sqrt[5]{\gamma}}	ag{6-4}
\]

Having calculated the shape parameter \( \gamma \), the velocity profile was obtained using the gamma function reported by Van Opstal and Van Gisbergen (1987).

\[
v(t) = \phi \left( \frac{t}{\theta} \right)^{\gamma-1} \exp\left( -\frac{t}{\theta} \right) \quad t \geq 0; \ \theta > 0; \ \gamma \geq 1 \tag{6-5}
\]
where \( v(t) \) is the saccadic velocity at time \( t \), and \( \phi \) and \( \theta \) are scaling constants for velocity and duration respectively. The velocity profile was first generated with these scaling constants equal to one. The velocity parameter \( \phi \) was then changed such that the new velocity profile had the maximum velocity given by Equation 6-2. The duration parameter, \( \theta \), was then adjusted so that after numerical integration of the velocity profile, the position profile had precisely the required displacement, and the duration using the velocity criterion of 30 degrees of visual angle /sec was within 10% of the value from Equation 6-1. The position signal was sampled at 500 Hz, similar to Experiment 1. The first 100 points of the position profile were selected to represent the saccade.

**Noise Added to Simulated data**

*Fixation error.*

To simulate actual EOG necessitated the addition of typical fixation error as well as typical physiological noise inherent in the EOG signal. When a subject is fixating on a point, Ditchburn and Ginsborg (1953) assert that the image of the fixation point is confined within a foveal area of about 0.33 degrees of visual angle in diameter. A foveal area of this size corresponds to the high receptor density area found near the centre of the foveal floor (Steinman, Cushman, & Martin, 1982). St. Cyr and Fender (1969) also support this estimate in reporting the foveal area to be about 0.25 degrees of visual angle for one of their subjects and about 0.5 degrees of visual angle for another. This foveal *dead zone* suggests that a subject would not be aware of this amount of variation in their point of fixation. So a subject who may have been fixating at the end point of the arm of a fixation cross, may in fact have been fixating some distance past that point at the instant the saccade to a target
was instigated. This means that the range of possible saccadic displacements (the fixation error) for a saccade made between a fixation cross and a target is increased as a result of this foveal dead zone.

For the horizontal target situation shown in Figure 14, centring the foveal dead zone on the extreme fixation points, and assuming the worst-case scenario of 0.5 degrees of visual angle diameter for the size of the zone, then 0.25 degrees of visual angle is added to either end of the longest possible displacement. Similarly 0.25 degrees of visual angle is subtracted from either end of the shortest possible displacement. So the foveal dead zone introduces 1 degree of visual angle to the possible range of displacements for a particular saccade. When the fixation and target sizes are also taken into account (see Figure 14), there could be a maximum displacement error of 2 degrees of visual angle for a horizontal saccade that starts with fixation of the central cross and finishes with fixation of a horizontal target. A saccade made along the diagonal line shown in Figure 14, from the bottom of the dotted circle around the fixation cross to the top of the dotted circle around the target, would introduce the maximum angular fixation error. This maximum angular error would depend on the displacement of the saccade. For the innermost horizontal target (2 degrees of visual angle) the maximum angular fixation error would be $27^\circ$ ($\theta = \tan^{-1} 1/2$), and for the outermost horizontal target the maximum angular fixation error would be $4.8^\circ$ ($\theta = \tan^{-1} 1/12$). This angular fixation error would introduce crosstalk between the channels, independent of that produced due to misalignment of the recording electrodes.

Similar arguments can be applied to saccades made to targets along the vertical and oblique meridians. It is a reasonable assumption that for two
dimensions the distribution of fixation points on the cross (and target) takes the form of a 2D circularly symmetric Gaussian probability distribution. Thus, any particular saccade has an equal chance of starting at any angle measured

![Diagram showing possible fixation errors](image)

**Figure 14.** Possible fixation errors (degrees of visual angle) for a saccade made to a horizontal target. The foveal dead zone areas are centred on the extreme edges of either the fixation cross or the target. Errors of fixation were sampled from a Gaussian distribution. The dotted circles around the fixation cross and target are the three standard deviations boundaries for the Gaussian distributions, and enclose the areas where the saccade may start or terminate, respectively. The difference between the minimum and maximum displacements is 2 degrees of visual angle (1.1 + 0.9). A saccade made along the diagonal line shown would produce the maximum fixation angular error of θ° above the horizontal.
with respect to the cross, but a higher probability of starting close to the centre. The same situation applies at the target for the location of the end of the saccade. The areas enclosed by the dotted circles in Figure 14 represent three standard deviations from the centre of the cross (target). Hence the standard deviation of the Gaussian distribution used at the fixation cross was 0.18 degrees of visual angle, and the standard deviation used at the target was 0.15 degrees of visual angle. Once the beginning and end points of the saccade were generated, the distance between the points was calculated and this was the value for the displacement of the saccade that was provided to the SSE. The standard error of the difference between two independent means is given by the following equation (Minium, King, & Bear, 1993).

\[
\sigma_{X-Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}
\]  

(6-6)

As the sample sizes are the same for each distribution, the standard deviation of the difference between random values from the two distributions would be

\[
\sigma_{X-Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}
\]  

(6-7)

Thus the standard deviation for the difference between points from each distribution was 0.23 degrees of visual angle. This represents the standard deviation of the error in saccadic amplitude for the task due to fixation error.

Gain differences.

In real signals, the gain in vertical EOG recordings is usually less than that of the horizontal recording, so gain differences between the channels were
simulated by simply multiplying the vertical component of the SSE output by a constant factor (parameter \( b \) in Equation 4-1, \( b < 1 \)) and keeping the horizontal component at a nominal gain of one (parameter \( a \) in Equation 1). The range of values used for the gain transformation parameter \( b \) was estimated from empirical data from the five subjects in Experiment 3.

**Rotation and translation of recording axes.**

Geometric transformations were simulated by multiplying the outputs from the SSE by a rotation matrix containing the parameters \( \alpha \) and \( \beta \) (see Equation 4-1 & Figure 9). The range of values used for the rotation transformation parameters were also estimated from empirical data from the five subjects in Experiment 3. Parameters \( O_H \) and \( O_V \) were eliminated from the transformation equation (Equation 4-2) by a simple baseline correction using the set of points just prior to the onset.

**Physiological noise.**

Adding a pseudo-randomised value from a Gaussian distribution to each sampled data point of the generated saccadic waveform simulated physiological noise in the EOG. This Gaussian distribution had a standard deviation of 0.33 degrees of visual angle, which produced a maximum range of physiological noise around the simulated signal of about 2 degrees of visual angle (± 3 standard deviations). This value was also estimated from empirical data from the five subjects in Experiment 3. Figure 15 shows a histogram of the deviations from the mean signal strength for the last 20 data points of all accepted trials in the horizontal channel across all five subjects. To express the deviations in degrees of visual angle the signals were calibrated separately for
each subject based on the largest accepted saccades (12 degrees of visual angle) in each channel.

*Figure 15.* Histogram of the deviations from the mean signal strength for the last 20 data points of all accepted trials in the horizontal channel across all five subjects.

This physiological noise was added after all other transformations had been applied. This ensured that applying a lower gain transformation to the
vertical channel resulted in a reduction in the signal/noise ratio for that channel, similar to an actual EOG recording. Gain change scaling was applied before axes rotation transformations, as gain differences between channels are present in recorded data, regardless of whether any rotational transformations are present.

**Simulated Waveforms**

The SSE provided vectors of 100 data points for each of the horizontal and vertical components of a simulated saccade of given displacement and direction. These waveforms ended in a plateau corresponding to fixation of the target (similar to those in Figure 4a). Vectors consisting of the concatenation of these 200 data points made up the input matrix provided to the ANNs (Figure 10b). Figure 16 displays some examples of the SSE waveforms as horizontal and vertical components. Figure 16a shows the waveforms representing a saccadic eye movement of 12 degrees of visual angle displacement made along the 30° meridian in the upper right quadrant of the visual field. These waveforms contain simulated fixation error and physiological noise but no other distortions. Because no gain difference or rotational transformation has been applied, and the simulated data contains no drift, the final signal strengths from each channel (representing fixation) provide a reasonably accurate estimate of the eye position at the end of the saccade. Figure 16b shows the same waveforms after the application of a gain scaling of 0.4 to the vertical channel. Figure 16c shows the original waveforms after a rotation of the vertical recording channel with $\beta = 12^\circ$ (see Figure 9).
Figure 16. Examples of simulated and transformed waveforms. The simulation was for a saccade of 12 degrees of visual angle displacement along the 30° meridian in an upper quadrant of the visual field. The waveform is composed of 100 data points computed by the saccadic simulation engine (SSE) for both horizontal and vertical components. The untransformed waveforms are shown in (a). The waveforms in (b) have had a gain transformation applied that reduced the gain of the vertical component by a factor of 0.4. The waveforms in (c) have had a rotation transformation applied which produced the equivalent effect of having the vertical recording axis rotated 12°. This resulted in the vertical channel receiving negative crosstalk from the horizontal channel, and hence reduced the amplitude of the vertical waveform. The waveforms in (d) have had both the gain and the rotation transformations applied in that order. This combination has reduced the amplitude of the vertical waveform virtually to zero, and produced the effect that the saccade was in a horizontal direction only. The simulated waveforms used in Experiment 2 also had rotation transformations applied to the horizontal recording axes.
This resulted in the vertical channel receiving negative crosstalk from the horizontal channel, and hence reduced the amplitude of the vertical waveform. Figure 16d shows the same waveforms after the application of both the gain scaling then the rotation. The combined transformations produce a marked difference in the appearance of the EOG signals. With a vertical component no longer evident, visual inspection of Figure 16d would suggest that the saccade producing these signals was purely horizontal in nature.

**Artificial Neural Networks**

The ANNs used in the experiment were MLPs and LPs from the Matlab® Neural Network Toolbox version 2.0 (The MathWorks, Inc., 1994). The LPs are only capable of linear transformations on the data and their performance would reflect the extent to which calibration of the simulated signals is a linear problem. Better results from the MLPs would indicate a non-linear component to the calibration process.

The MLPs were the same as in Experiment 1. The architecture of the MLPs (see Figure 6) consisted of 200 inputs (100 to represent the saccade from each recording channel), 20 hidden nodes, and 2 output nodes. Production of the training and testing sets, initialisation, and training of the MLPs was also the same as in Experiment 1. Training curves (calibration error vs. number of training iterations) were used to establish the necessary training times for convergence. Weight configurations that resulted in a performance improvement $\geq 1\%$ were saved to disk. The trial number of the last update was used to assess the number of iterations to convergence (training time).

The LPs were also trained using the MLP training function from Experiment 1 (trainbpx.m). For LPs, the optimum network design for
calibrating the training set could be calculated mathematically with a Matlab® function (solvelin.m), however this approach may result in some overfitting (Prechelt, 1998), with reduced performance on the testing set. The LPs were thus subjected to training, but because they were slow in their learning using the linear training functions (requiring hundreds of thousands of training iterations to reach their optimum performance), they were also initialised using the MLP initialisation function (initff.m). This was achieved by declaring zero hidden nodes and linear activation functions for the output nodes. The LPs could now be trained using the MLP training function (trainbpx.m), which allowed the use of momentum and adaptive learning rate to speed up the training procedure and reduce the number of training iterations required for optimum performance. To avoid premature saturation (Schalkoff, 1997), momentum and increased learning rate (mom = 0.95, lr_inc = 1.02) were only introduced into the learning algorithm after 300 training iterations had been completed. Even so, training the LPs on the simulated data required 100,000 training iterations for convergence. The configuration of the LPs consisted of 200 inputs (100 to represent the saccade from each recording channel), no hidden nodes, and 2 output nodes.

Thus both the LPs and the MLPs were subjected to backpropagation training. Each training session consisted of 10 repetitions. Each repetition for the MLPs represented a random division into training and testing sets and a random assignment of network weights and biases. As the architecture of the LPs was different to that of the MLPs, each repetition of the LPs also received a random assignment of network weights and biases, but the same training and testing sets as the MLPs. The performance figures for the LPs and the simpler
classifier (see below) were based on the same 10 training and testing repetitions that were randomly generated for the MLPs.

*The Simple Classifier (SC)*

To assess the effectiveness of the ANNs, calibration of the eye movements was also examined using a simple algorithm that represented how a naive human operator might hand-score the EOG waveforms for calibration by simply using the fixation signal amplitudes from each channel. As with the ANNs, the input matrices, containing the concatenated horizontal and vertical saccadic components, were re-scaled to ±1. The SC used the medians of the last 20 data points from each directional component to represent the fixation voltages. These two median values constituted the output from the SC. These values could then be directly compared to the targets’ screen coordinates (training vectors) as they had also been rescaled to ±1.

*Simulated Data Sets*

Using the SSE, twenty trials of each target type were generated to provide a simulated data matrix for the ANNs. This matrix contained a random fixation error for each simulated saccade but no distortions or physiological noise. This particular data matrix (m100998.mat) was used to benchmark the effects of different transformations of the data on the ability of the ANNs and the SC to correctly calibrate the eye movements, that is, it was the base matrix to which all further transformations were applied.

Different rotations of the recording axes and gain differences between the recording channels were systematically applied to the data matrix to assess the effect of different gains in the vertical channel and combinations of gain changes and oblique rotations of both axes on the performance of the
calibrators. The transformed matrix contained a random addition of physiological noise for each repetition and was supplied to each of the calibrators (SC, LP, and MLP) in turn. Non-orthogonality between recording axes, in combination with gain differences between the recording channels, represents the conditions most likely to be found when recording real data.

**Calibration Accuracy**

For a particular trial, the performance of a calibrator was determined by calculating the linear distance between the point defined by the output coordinates of the calibrator and the intended target coordinates. This linear distance was an absolute value obtained using the formula for the linear distance between two points of coordinates \((x_1, y_1)\) and \((x_2, y_2)\) respectively.

\[
\text{linear distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}
\]  

(6-1)

The mean of this distance for all trials was then calculated and expressed in degrees of visual angle.

**Results**

Table 2 summarises the effects on the accuracy of the three calibrators for different gain changes within the vertical channel. Also shown are the effects of applying oblique rotations to the recording axes in addition to a substantial gain change in the vertical channel. The mean calibration errors (degrees of visual angle) in this table were based on calibration performance averaged over 10 random training-testing splits of the data. Overall, the performance of the SC was very inferior to the performance of both the ANNs. The performance of the MLPs exceeded that of the LPs. Indeed, for untransformed data, the MLPs showed approximately 53% improvement on the error left by the LPs. The SC was excluded from the statistical analyses.
Table 2

Means and Standard Error of the Means (n = 10) for Calibration Errors
(degrees of visual angle) in Simulated EOG Data with Different Gains in the
Vertical Channel and Oblique Rotations of the Recording Axes.

<table>
<thead>
<tr>
<th>Transformations</th>
<th>Performance</th>
<th>Artificial Neural Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Simple Classifier</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error M SE</td>
</tr>
<tr>
<td>Gain</td>
<td>Rotation</td>
<td>Error M SE</td>
</tr>
<tr>
<td>ver</td>
<td>hor ver</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>00 00</td>
<td>0.514   .002</td>
</tr>
<tr>
<td>0.8</td>
<td>00 00</td>
<td>0.966   .002</td>
</tr>
<tr>
<td>0.6</td>
<td>00 00</td>
<td>1.611   .002</td>
</tr>
<tr>
<td>0.4</td>
<td>00 00</td>
<td>2.382   .002</td>
</tr>
</tbody>
</table>

Gain changes in vertical channel

<table>
<thead>
<tr>
<th>Gain</th>
<th>Rotation</th>
<th>Error M SE</th>
<th>Error M SE</th>
<th>Error M SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-03 03</td>
<td>0.738   .002</td>
<td>.544   .012</td>
<td>.275   .003</td>
</tr>
<tr>
<td>1.0</td>
<td>00 06</td>
<td>1.061   .002</td>
<td>.535   .012</td>
<td>.278   .006</td>
</tr>
<tr>
<td>1.0</td>
<td>-06 06</td>
<td>1.073   .002</td>
<td>.517   .010</td>
<td>.275   .002</td>
</tr>
<tr>
<td>1.0</td>
<td>00 12</td>
<td>1.725   .002</td>
<td>.534   .009</td>
<td>.279   .004</td>
</tr>
<tr>
<td>0.4</td>
<td>-03 03</td>
<td>2.383   .001</td>
<td>.631   .022</td>
<td>.378   .011</td>
</tr>
<tr>
<td>0.4</td>
<td>00 06</td>
<td>2.455   .001</td>
<td>.638   .027</td>
<td>.380   .007</td>
</tr>
<tr>
<td>0.4</td>
<td>-06 06</td>
<td>2.411   .002</td>
<td>.592   .018</td>
<td>.380   .015</td>
</tr>
<tr>
<td>0.4</td>
<td>00 12</td>
<td>2.736   .002</td>
<td>.570   .018</td>
<td>.382   .008</td>
</tr>
</tbody>
</table>
due to its obvious failure to match the performance of the ANNs once any transformations had been applied to the data.

As several ANOVAs were required to analyse the data in Table 2 an alpha level of .01 was adopted to reduce the Type I error rate. A $2 \times 4$ (Calibration Method $\times$ Gain Change) mixed ANOVA, with calibration method serving as the repeated measure, was used to examine the effect of gain changes in the vertical channel on the accuracy of the two ANNs. There was a significant main effect for calibration method, $F(1, 36) = 1262.1, p < .001$, with the MLPs ($M = 0.307, SE = 0.003$) outperforming the LPs ($M = 0.600, SE = 0.010$). There was a significant main effect for gain change, $F(3, 36) = 23.2, p < .001$, but no significant interaction. HSD post hoc tests (experiment-wise alpha level = .01) revealed that, for both calibrators, a gain of 0.4 produced significantly poorer accuracy than all other gains. For the MLPs, a gain of 0.6 also produced significantly poorer accuracy than a gain of 0.8 or 1.0.

The second section of Table 2 displays the results from calibrations of data matrices, which had a vertical gain of either 1.0 or 0.4 applied, prior to oblique rotations of the recording axes. The rotations resulted in either 96° or 102° of separation between the axes (see Figure 9). This was achieved by rotating only the vertical axis +6° or +12°, or by rotating the vertical axis by half these amounts and rotating the horizontal axis -3° or -6° respectively.

A $2 \times 2 \times 2 \times 2$ (Calibration Method $\times$ Gain Change $\times$ Number of Axes Rotated $\times$ Rotation Angle) mixed ANOVA was conducted with calibration method serving as the repeated measure. This analysis was done to determine if no gain change versus a gain change of 0.4, oblique rotation of only one versus both recording axes, or a 6° versus 12° increase in axes separation, had
a direct or interactive effect on the performance of the ANN calibrators. There was a significant main effect for calibration method, $F(1, 72) = 1408.2, p < .001$, with the MLPs ($M = 0.328, SE = 0.003$) outperforming the LPs ($M = 0.570, SE = 0.006$), and a significant main effect for gain change, $F(1, 72) = 167.0, p < .001$, with the 0.4 gain change producing a poorer performance ($M = 0.494, SE = 0.005$) than the 1.0 category ($M = 0.405, SE = 0.005$).

A plot of mean accuracy vs. the number of training iterations for the training and testing data (see Figure 17) indicated that the LPs required approximately 20 times the number of training iterations than did the MLPs (100000 vs. 5000). Figure 17 displays the median training curves across all training-testing sets in the second section of Table 2. Although the LPs converge well with the training data they do not generalise to the test data as well, or as quickly, as the MLPs.

Figure 18 displays the horizontal and vertical calibration errors for the ANNs in relation to each individual target for transformations in the second part of Table 2. The calibration errors shown are for test data only. For each target, the mean horizontal and vertical errors (averaged over all training-testing runs) have been added to the target coordinates and plotted as a point which represents the mean output of the ANN for that particular target. Unlike the absolute calibration error distances in Table 2, the errors displayed in Figure 18 are directional and incorporate undershoot and overshoot in both horizontal and vertical directions. The error bars displayed for each point represent 1 standard deviation from the mean. The mean computed location lies at the intersection of the error bars. The points on the plot marked by asterisks represent the true location of the targets.
**Figure 17.** Plots of mean calibration error vs. the number of training iterations for training and testing simulated data for multi-layer perceptrons (MLPs) and linear perceptrons (LPs). The plots represent the median waveforms of all training-testing sets in the second part of Table 2.
Simulation Data - LP

Simulation Data - MLP
**Figure 18.** Horizontal and vertical calibration errors for the artificial neural networks (ANNs) for simulation data in relation to each individual target. An asterisk marks the true location of each target point. The intersection of the error bars represents the mean output of the ANN for that particular target averaged over all training-testing runs and all subjects. The error bars represent ±1 standard deviation from the mean. Figure 18a displays the errors from calibration of the simulation test data using the linear perceptrons (LPs). Figure 18b displays the errors from calibration of the simulation test data using the multi-layer perceptrons (MLPs).
Discussion

Results indicate that both ANNs are capable of performing the necessary computations required to calibrate simulated 2D saccadic eye movement EOG. Overall the performance of the SC was very poor in comparison to the ANNs. Simply using the fixation signal amplitudes directly from each channel for calibration (i.e., the SC) was as effective as the LPs, but only when no transformations had been applied to the simulated data. This would indicate that visual inspection of 2D eye movement records is acceptable provided no sources of distortion are introduced and this is not a realistic condition in most EOG recordings. With distortion present, in many situations, depending on the task requirements, an error of up to nearly 3 degrees of visual angle (see Table 2) would be unacceptable. This result emphasises the need for an automated procedure for interpreting 2D saccadic eye movement EOG.

Both ANNs overcame gain changes in the vertical channel of the data without appreciable loss of performance. However, the MLPs significantly outperformed the LPs in calibrating simulated EOG containing this type of transformation. The mean accuracy of the MLPs was always less than 0.40° of visual angle, whereas the LPs’ mean accuracy fell to 0.75° of visual angle for simulated EOG containing a vertical gain of 0.4. Reduced accuracy for this category was expected as a lower gain in the vertical channel effectively decreases the signal/noise ratio making the calibration of the smaller displacement saccades more problematic.

The data presented in the second section of Table 2 is indicative of the performance of the calibrators on data containing transformations likely to be
found in real EOG data, that is, combinations of gain changes and oblique rotations of one or both recording axes. The significant main effects on accuracy for calibration method and gain change, indicate that the MLP is the more accurate calibrator, however both ANNs experience difficulties calibrating data with a 0.4 gain in the vertical channel due to its poorer signal to noise ratio. Both ANNs appear to be unaffected when the oblique rotation is of the higher magnitude, or results from rotation of both recording axes simultaneously. However, it is important to note that the mean values from the second half of Table 2 indicate that the MLPs provide a 42% improvement over the calibration accuracy of the LPs (0.570 vs. 0.328 degrees of visual angle).

All the transformations introduced into the simulated saccadic signals were linear transformations, so theoretically the LPs should be just as capable of mapping their outputs to the target screen coordinates, as were the MLPs. The only limitation on the LPs should be the minimal non-linearity in the relationship between degree of visual angle and target screen coordinates (see Figure 11) as the non-linear screen coordinates of the targets was still used for the training vector. The median training curves plotted in Figure 17 do suggest that the LPs were capable of performing the necessary mapping as indicated by their performance on the training data set. In training, the LPs performed well indicating that the necessary mapping was predominately linear. The MLPs though, were superior in dealing with the unseen test set data in both calibration accuracy and training time. The MLPs had mostly converged after 5000 training iterations with an accuracy of about 0.3 degrees of visual angle, whereas the LPs had only achieved an accuracy of about 0.6 degrees of visual angle.
angle after 100,000 training iterations. The network training algorithm programme required an improvement in performance of at least 1% before new network weights were saved to disk and only the occasional network from the 10 training runs was still improving at the respective training limits allowed for each type of ANN. By the time the MLPs had completed their training (5000 iterations), the LPs had only achieved an accuracy of about 0.5 degrees of visual angle on the training data. The training curves also suggest that considerable training time, well in excess of an additional 100,000 iterations, would be required for the LPs to significantly improve their accuracy on the test data.

The standard deviations displayed in Figure 18 support the absolute error measurements shown in Table 2 in that the MLPs have considerably smaller standard deviations. For the MLPs, there appears to be larger standard deviations for targets along the oblique meridians compared to those along the horizontal and vertical axes. For the LPs, the standard deviations appear to be similar across all targets. As the MLPs have calibration errors close to the theoretical fixation limit of 0.23 degrees of visual angle (see Equation 6-7), this suggests that the MLPs have been successful in dealing with the physiological noise added to the simulated waveforms. If the variance from the physiological noise is combined with that due to the fixation error in the manner of Equation 6-7, the resulting standard deviation is 0.40 degrees of visual angle. The calibration errors for the LPs are all above 0.50 degrees of visual angle, suggesting that the performance of the LPs is affected by the physiological noise added to the simulated waveforms. The ability of the MLPs to perform non-linear mappings may allow them to better map the input
waveforms containing the physiological noise and result in smaller standard deviations in their calibration errors. The oblique targets may be slightly more difficult to calibrate, as the signal in both channels is more critical to correct calibration than is the case with either purely horizontal or vertical targets.

With the same number of hidden nodes the MLPs have performed much better on the simulated data than they did on subject data in Experiment 1. This stresses the importance of minimising sources of variance such as head movement. Searching for the saccadic onset and using this to minimise the amount of data presented to the MLP also reduces the amount of noise in the input signals. The use of degrees of visual angle, instead of percent of outputs closest to the correct target as the scoring criterion, would also be expected to be a more generally useful indicator of calibration performance.

**Summary**

MLPs are more effective than the LPs at calibrating data containing combined transformations of gain changes and rotations. They were much faster to train, more robust through the various transformations most likely to be encountered in real 2D EOG, and significantly more accurate in terms of absolute calibration error. MLPs appear to be a suitable choice of calibrating tool for simulated 2D EOG.
CHAPTER 7

Experiment 3

Subject Data

In Experiment 3, the aim was to compare the three saccadic eye movement calibrators (SC, LPs, and MLPs) on their performance with real 2D saccadic EOG and with simultaneously recorded infrared limbus tracking data. A dedicated eye movement laboratory was established for the study of eye movements, so the saccadic task was presented, and 2D EOG recorded, on equipment specifically developed for this purpose, rather than the EEG equipment used in Experiment 1. The saccadic eye movement task involved the same eye movements as in Experiment 1 and simulated in Experiment 2. The measure of calibration accuracy was also the same as that used in the simulation experiment. Simultaneous 2D infrared limbus tracking recordings were also obtained to provide a benchmark for the performance of the ANN-EOG recording system.

Method

Participants

Participants were colleagues and associates of the researchers. All had uncorrected vision, although acuity was not checked. There were five subjects, three females and two males. The mean age was 27.4 years ($SD = 8.9$).
Materials

Electrodes and Hardware

Disposable silver-silver chloride electrodes were used to record vertical and horizontal EOG, using shielded leads connected to CONTACT® Precision Instruments, High Sensitivity Bioamplifiers. The infrared limbus tracker was SKALAR Medical's infrared IRIS® model 6500. The saccadic task was presented on a monitor located in a small quiet well-lit room. A programme written in Visual Basic® on an IBM compatible 486 PC presented the stimuli for the task. The PC had a ComputerBoards Inc.® I/O card installed that allowed 12 bit digital sampling and storage of the ±3 volt output of the bioamplifiers.

Head Restraint

A custom made head restraint was utilised to stabilise the subject’s head during the task. The head restraint incorporated a chinrest and two slightly convex, polished alloy cheek rests (see Figure 19). Extension tubes containing two forehead rests, similar to the two cheek rests, were available for the two vertical tubes but were not employed for this task because they would interfere with the IRIS® headband (see Figure 5). The choice of this type of head restraint, rather than a bite bar, which may have been more effective in stabilising the head, was used to avoid contaminating the EOG with electromyographic activity (EMG) from the jaw muscles (Iacono & Lykken, 1981).

Stimuli.

The stimuli were the same as those used in Experiment 1. The stimuli were a series of dots presented individually on a monitor for a duration of 0.5
sec. The screen was 45 cm from the subject. At that distance a stimulus dot (target) subtended 0.4 degrees of visual angle, and the fixation cross subtended 0.6 degrees of visual angle along each arm of the cross (Figure 14). There were a total of 24 different locations for the targets (see Figure 12). Each of the 24 target stimuli was repeated 20 times at random, providing 500 trials (including the ‘Blink’ trials) for a session.

Figure 19. The custom made head restraint. The chin and cheek rests were made of polished alloy and the vertical tubes brass. The cheek rests were adjusted so that they offered firm support just under the cheekbones.
Procedure

EOG

Horizontal EOG was recorded using a bipolar montage with electrodes placed at the outer canthi of each eye. Vertical EOG was also a bipolar recording from electrodes placed above and below the right eye. Any skin area which would come in contact with an electrode or any of its adhesive surface was thoroughly rubbed and cleaned with alcohol to remove dead cells, dirt, and skin oils. The placement of the horizontal electrodes was done by visual inspection. A transparent sheet, on which was drawn a set of orthogonal axes, was then used as an aid to locate the vertical electrodes. An additional electrode was also placed behind the right ear and connected to the amplifier earth to reduce the occurrence of DC drift. The task was performed in a well-lit room, so no dark adaptation period was necessary.

A test screen displayed zero-centred gauges, which enabled the researcher to observe the EOG signal as subjects shifted their gaze from left to right and from up to down. Observation of these gauges provided an estimate of the magnitude of the EOG signal in each of the recording channels, and also an estimate of the degree of crosstalk present. No adjustment of the EOG system was really possible other than actually moving (i.e., replacing) one of the electrodes. This would only be necessary if feedback from the test screen indicated that the original placement of the electrodes was grossly in error, or if an electrode’s skin contact had too high an impedance to be functional. The gain on the bio-amplifiers for both EOG channels was set such that full-scale deflection (fsd) was 800 µV. Pilot work had indicated that if the gain on the vertical channel was increased to compensate for its smaller signal, then
invariably the data from some subjects’ trials would be clipped (outside the recording range). Having both channels with equal amplification also allowed for easier visual comparison of signal strength, physiological noise levels, and degree of crosstalk.

*IRIS*® *Infrared Limbus Tracker*

Subjects were also fitted with the IRIS® infrared limbus tracker. The IRIS® head-mounted sensor device shown in Figure 5, was placed on the subjects head and the straps going over and around the head adjusted so that the fit was firm but still comfortable. The infrared sensors were then initially positioned in front of the eyes by visual inspection. As with EOG, vertical data was recorded from the right eye, whereas horizontal IRIS® data was recorded only from the left eye. The infrared system contained its own LED arrays indicating signal strength in each direction for each channel. Optimal alignment of the sensors was then obtained using the test screen and the iterative process described in the manufacturer-supplied manual. This ensured that the gain of the signal was adequate, and that there was no asymmetry between left and right in the horizontal channel, or between up and down in the vertical channel.

*Saccadic Task*

This was the same task used in Experiment 1 presented by a programme written in Visual Basic®. The only difference to the task was that the subjects could pace the task using a mouse press to advance to the next trial. Subjects were asked to rest their chin on the chin rest of the head restraint. The cheek rests were then adjusted so that they offered firm support just under the cheekbones. The subject was seated comfortably with their head
naturally erect and supported by the head restraint. In this final position, the subject’s eyes were in the primary position and looking straight ahead at the centre of the fixation cross. Subjects were instructed to fixate on the centre of the fixation cross, and when ready press the mouse button. The button press presented the next target. Subjects were instructed to fixate on the target and then look back to the centre of the cross.

*Data Recording*

The voltage differences at the electrode pairs were digitally sampled at a frequency of 500 Hz for a period of one second from the onset of the target stimulus. At this stage the data consisted of 500 points from each recording channel.

*Data Pre-processing*

The data were pre-processed using a computer algorithm (auto_cross2.m). This was a Matlab® program that rejected trials due to the presence of excess noise or blinks (as in Experiment 1), detected the saccadic onsets in each recording channel, then extracted and concatenated the waveforms representing the saccades for satisfactory trials. Only trials with satisfactory waveforms in both recording channels were accepted. Details of the programme can be found in Appendix B.

*Effectiveness of Automatic Trial Rejection*

The effectiveness of the automatic trial rejection was checked by a comparison to four sets of subject data that were viewed manually to determine if their EOG waveforms were acceptable for inclusion in the calibration trials. Individual trials were rejected if the waveforms contained excessive noise, blinks, or double saccades in either recording channel. To
view the data manually the data from each recording channel for each trial were viewed separately on a computer monitor using a programme written in Matlab®. If the data appeared to be satisfactory, then each channel’s data was enlarged to full screen size so that the saccadic onset could be selected with the mouse pointer. When viewing either purely horizontal or vertical target trials, crosstalk, if present, was evident when initially viewing both channels side by side. Crosstalk, however, was not evident when viewing both channels data from oblique target trials. Of course, crosstalk being present was not evident when viewing data from one channel during the process of selecting the onset. This made it very difficult at times for the viewer to judge the overall suitability of the trial’s data and the correct saccadic onsets from the different channels due to the distortion of the signal from the crosstalk. This included at times, through cancellation, a delaying effect of the onset in one channel. Nevertheless there were a total of 2000 trials visually inspected. As with the automatic pre-processing, trials corresponding to the blink trials types were omitted. This left 1920 trials remaining for the comparison.

*Training the ANNs*

Vectors consisting of the concatenation of the 100 data points selected from the unsmoothed raw data from each channel, made up the input matrix provided to the calibrators. The calibrators, and training methods and architecture of the MLPs and the LPs were the same as in Experiment 2. The training matrix consisted of the screen coordinates of the corresponding targets.
Results

Table 3 displays the results of the comparison between the automatic pre-processing algorithm and visual inspection of subject data. The percentage agreement between the automatic pre-processing algorithm and visual inspection is 91%. When running on a PC with a Pentium II, 400 MHz processor, the mean time for the automatic pre-processing of 500 saccadic trials was just under 10 minutes.

Table 3

<table>
<thead>
<tr>
<th>Method</th>
<th>Accepted</th>
<th>Rejected</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Inspection</td>
<td>1783</td>
<td>137</td>
<td>1920</td>
</tr>
<tr>
<td>Auto Pre-processing</td>
<td>1740</td>
<td>180</td>
<td>1920</td>
</tr>
<tr>
<td>Both</td>
<td>1673</td>
<td>70</td>
<td>1743</td>
</tr>
</tbody>
</table>

Table 4 displays the number of accepted trials in the data set (automatic algorithm) and the mean calibration errors for the three calibrators on both the EOG and the IRIS® data. Mean calibration errors were averaged over 10 random training-testing splits of the data. Inspection of these data indicate, that as with the simulated EOG data in Experiment 2, the SC’s performance on real EOG (and IRIS®) data was greatly inferior to the performance of both the ANNs. The accuracy of both ANN calibrators was
Table 4

Means and Standard Error of the Means (n = 10) for Calibration Errors (degrees of visual angle) in Subject Data.

<table>
<thead>
<tr>
<th>Subject Information</th>
<th>Performance</th>
<th>Artificial Neural Networks</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Simple Classifier</td>
<td>Linear Perceptrons</td>
<td>Multi-Layer Perceptrons</td>
<td></td>
</tr>
<tr>
<td>Subject No of Trials</td>
<td>Error</td>
<td>M</td>
<td>SE</td>
<td>Error</td>
<td>M</td>
</tr>
<tr>
<td>S1 456</td>
<td>3.50</td>
<td>0.01</td>
<td>1.53</td>
<td>0.02</td>
<td>0.95</td>
</tr>
<tr>
<td>S2 417</td>
<td>2.89</td>
<td>0.01</td>
<td>1.76</td>
<td>0.02</td>
<td>1.21</td>
</tr>
<tr>
<td>S3 455</td>
<td>3.03</td>
<td>0.01</td>
<td>1.37</td>
<td>0.02</td>
<td>1.01</td>
</tr>
<tr>
<td>S4 412</td>
<td>3.37</td>
<td>0.01</td>
<td>1.45</td>
<td>0.02</td>
<td>1.03</td>
</tr>
<tr>
<td>S5 402</td>
<td>3.26</td>
<td>0.01</td>
<td>1.53</td>
<td>0.02</td>
<td>1.13</td>
</tr>
<tr>
<td>means 428</td>
<td>3.21</td>
<td>0.03</td>
<td>1.53</td>
<td>0.02</td>
<td>1.07</td>
</tr>
<tr>
<td>EOG Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1 413</td>
<td>3.00</td>
<td>0.01</td>
<td>0.83</td>
<td>0.01</td>
<td>0.78</td>
</tr>
<tr>
<td>S2 336</td>
<td>2.07</td>
<td>0.01</td>
<td>1.01</td>
<td>0.02</td>
<td>0.90</td>
</tr>
<tr>
<td>S3 429</td>
<td>3.06</td>
<td>0.01</td>
<td>1.14</td>
<td>0.02</td>
<td>1.06</td>
</tr>
<tr>
<td>S4 418</td>
<td>4.09</td>
<td>0.02</td>
<td>0.95</td>
<td>0.02</td>
<td>0.76</td>
</tr>
<tr>
<td>S5 325</td>
<td>3.39</td>
<td>0.01</td>
<td>1.28</td>
<td>0.01</td>
<td>1.23</td>
</tr>
<tr>
<td>means 384</td>
<td>3.12</td>
<td>0.09</td>
<td>1.044</td>
<td>0.02</td>
<td>0.95</td>
</tr>
</tbody>
</table>
better when using the IRIS® data, though not significantly better when using MLPs. For both data sets, the MLPs outperformed the LPs.

A 2 × 2 (Calibration Method × Data Type) mixed ANOVA, with calibration method serving as the repeated measure, was used to examine the effect of calibration method and data type (EOG data vs. IRIS® data) on the mean calibration error. An alpha level of .05 was adopted and the SC was not included in the analyses. There was a significant main effect for calibration method, $F(1, 8) = 121.6, p < .001$, a significant main effect for data type, $F(1, 8) = 9.6, p < .05$, and a significant two-way (Calibration Method × Data Type) interaction, $F(1, 8) = 52.3, p < .001$. The interaction between calibration method and data type can be seen in Figure 20, which displays the marginal mean calibration error (degrees of visual angle) for the LP and MLP calibrators, for both EOG and IRIS® data types. Figure 20 also shows the 99% confidence interval theoretical limit of resolution imposed by the fixation error associated with the nature of the task (3 standard deviations of 0.23 degrees of visual angle – see Experiment 2). This is the dashed line at the bottom of the figure.

Figure 21a and Figure 21b show plots of mean accuracy vs. the number of training iterations for the median performance curves from all training-testing sets for subject EOG data. These exhibit similar shapes to those obtained with the simulation data (see Figure 17). The notable difference being that the LPs now only required approximately double (rather than 20 times) the number of training iterations for convergence than did the MLPs (10000 vs. 5000). Figure 21c and Figure 21d show plots of mean accuracy vs. the number of training iterations for the median performance curves from all
Figure 20. Marginal mean calibration error (degrees of visual angle) for the linear perceptron (LP) and multi-layer perceptron (MLP) calibrators, for both EOG and IRIS® data types. The dashed line at the bottom of the figure represents the theoretical limit of resolution (99% confidence interval) imposed by the fixation error associated with the nature of the task.
Subject EOG Data, LP - Median Waveform

(a)

Subject EOG Data, MLP - Median Waveform

(b)
Subject IRIS Data, LP - Median Waveform

(c)

Subject IRIS Data, MLP - Median Waveform

(d)
Figure 21. Plots of mean calibration error vs. the number of training iterations for training and testing subject data with the artificial neural networks. The plots represent the median waveforms of the 10 training-testing sets of data. Figure 21a represents the training of linear perceptrons (LPs) with EOG subject data. Figure 21b represents the training of multi-layer perceptrons (MLPs) with EOG subject data. Figure 21c represents the training of LPs with IRIS® subject data. Figure 21d represents the training of MLPs with IRIS® subject data.
training-testing sets for subject IRIS® data. There is very little difference in the shape of the MLP plots between IRIS® and EOG data, however, the LP plots reveal the better accuracy and quicker convergence obtained with the IRIS® data.

The accuracy data reported in Table 4 reflects the overall performance of the calibrators in terms of the mean absolute 2D error distance between calibrator output and target coordinates. Figure 22 displays the horizontal and vertical calibration errors for the MLPs in relation to each individual target for both subject EOG and IRIS® data. For each target, the mean horizontal and vertical errors (averaged over all training-testing runs and all subjects) have been added to the target coordinates and plotted as a point which represents the average output of the MLP for that particular target. Unlike the absolute calibration error distances in Table 4, the errors displayed in Figure 22 are directional and incorporate saccadic undershoot and overshoot in both horizontal and vertical directions. The error bars displayed for each point represent 1 standard deviation from the mean. The points on the plot marked by asterisks represent the true location of the targets. Figure 22a and Figure 22b display the errors from calibration of the EOG data. Figure 22c and Figure 22d display the errors from calibration of the IRIS® data.

Discussion

The results of Experiment 3 have shown that for saccadic eye movements, EOG calibrated using a MLP, can provide a resolution similar to an infrared eye tracker. Specifically, for a saccadic movement task with targets in the upper right quadrant, both EOG and IRIS® data provide eye position resolution to about 1 degree of visual angle.
Subject EOG Data - LP
(a)

Subject EOG Data - MLP
(b)
(c) Subject IRIS Data - LP

(d) Subject IRIS Data - MLP
Figure 22. Horizontal and vertical calibration errors for the artificial neural networks (ANNs) in relation to each individual target. The true location of each target point is marked by an asterisk, and the intersection of the error bars represent the average output of the ANN for that particular target (averaged over all training -testing runs and all subjects). The error bars represent ±1 standard deviation from the mean. Figure 22a displays the errors from calibration of the EOG data using the linear perceptrons (LPs). Figure 22b displays the errors from calibration of the EOG data using the multi-layer perceptrons (MLPs). Figure 22c displays the errors from calibration of the IRIS® data using the LPs. Figure 22d displays the errors from calibration of the IRIS® data using the MLPs.
The theoretical limit of 0.7 degrees of visual angle imposed by the fixation error (see Figure 20) indicates that the 1 degree of visual angle resolution is very good for this task. As the EOG contains physiological noise with a range of 2 degrees of visual angle and the IRIS® data are free of this type of noise, the calibration of both sets of data to the same resolution indicates that the MLPs were successful in dealing with the physiological noise of the EOG. For each trial, the EOG provides up to 30 data points during fixation of the target. These points contain physiological noise with a Gaussian like distribution. As the standard error of the mean is inversely proportional to the square root of the number of samples, the error in the fixation produced by the physiological noise is reduced considerably. It is feasible that the nodes in the hidden layer are performing this kind of averaging. The fixation errors associated with the saccadic task and the lack of a bite bar restricting the capabilities of the infrared system, results in no significant differences in the accuracy between EOG and IRIS® data when calibrating 2D saccadic eye movements using MLPs.

The data displayed in Figure 22 show that calibration errors were largest within the vertical component of the calibrators' output for those target locations that include a vertical component. This is less evident in the IRIS® data than in the EOG data. This is probably because the EOG vertical data has the poorest signal to noise ratio. For both eye trackers, the average output for the highest displacement saccades consistently undershoots the target locations regardless of the meridian on which they occur. As target screen coordinates were calculated using the non-linear relationship between visual angle and distance on a flat screen, this undershooting is unlikely to be due to this non-
linear relationship, and may suggest that 12 degrees of visual angle is approaching the limit for saccadic eye movements without head movement. These results could be obtained from subjects either undershooting the target on the larger saccades in an effort to keep their head still or making a small head movement in conjunction with the larger saccades.

Given sufficient training (100,000 iterations), the LPs appeared capable of mapping the combinations of linear transformations introduced into the simulated EOG data in Experiment 2. However, on real data, their calibration performance was only about half as accurate, although with fewer iterations (10,000). This reduction in training times may indicate the presence of some non-linearity in the 2D EOG saccadic eye movement recordings limiting the maximal performance of LPs. LPs are not capable of non-linear mappings, hence may converge more quickly with reduced effectiveness.

The training curves (Figure 21) and statistical analyses for both calibrators and data types indicate that LPs provided significantly more accurate calibration of saccadic eye movements when using IRIS® versus EOG data. This is as expected as the IRIS® utilises an external signal source free of physiological noise and produces no differences in the gains between the two recording channels.

The plots of mean calibration error vs. the number of training iterations (Figure 21) suggest that MLPs perform the same for both EOG and IRIS® data with both accuracy of calibration and convergence time being very similar. The LPs performed more accurately and reached convergence more quickly on IRIS® data with the accuracy being similar to that obtained with MLPs. This suggests that IRIS® subject data, unlike EOG subject data, may not contain any
non-linear transformations of the eye movement signals. This is also
suggested by the interaction shown in Figure 20. Unlike the calibration
accuracy for EOG data, the calibration accuracy of the IRIS® data did not
improve significantly with the use of MLPs, which can deal with non-linear
transformations.

The automatic pre-processing algorithm for screening 2D saccadic eye
movement data was successful and provides considerable time saving over
manual inspection (less than 10 minutes for 500 trials). Due to the presence of
crosstalk, the manual screening of subject data was very time consuming,
subjective, and highlighted the need for an automatic processing algorithm.
The two innermost oblique targets required the subject to perform a 2 degree
of visual angle displacement saccade along the 30° or 60° meridian, which in
some cases, produced horizontal and vertical components of less than 1 degree
of visual angle in displacement. With physiological noise present in the EOG
signal with a range of 2 degrees of visual angle, and saccades in some
channels being less than 1 degree of visual angle in displacement, it become
very difficult to manually detect the onset of these saccades. The automatic
pre-processing algorithm accepted 138 (69%) of the 200 innermost oblique
trials from the 5 subjects. As can be seen in Figure 22, these trials were then as
successfully calibrated by the ANNs as were other target trials.
CHAPTER 8

Conclusions

The most interesting result is that using MLPs for calibration appears to be able to overcome some of the disadvantages of the EOG and for a saccadic eye movement task provide results with similar accuracy to the infrared system. On the basis of the inherent fixation error in the task, the best possible result would be an error of about 0.7 degrees of visual angle. So a mean error of just over 1 degree of visual angle for 2D EOG recordings in the first quadrant suggests that the resolution of the MLP-EOG system is approaching the apparent limitation imposed by the task and is not being limited by the physiological noise or artefacts within the 2D EOG signal. In general terms, the results are very promising for building a MLP/EOG system to provide an economical alternative for clinicians or researchers requiring a 2D saccadic eye tracker within this range of resolution.

Simulated Data

All of the calibrators were more accurate on simulated EOG data than on real 2D saccadic eye movement data. The MLPs were able to calibrate the simulated EOG signals with a mean error of 0.33 degrees of visual angle. The LPs’ calibrations produced nearly double this error and this is possibly partially due to the non-linear relationship between visual angle and distance on a flat screen.

Referring to Figure 14, the average error of fixation can be estimated as approximately 0.23 degrees of visual angle (see Equation 6-6). The top
entry in Table 1 shows the MLPs, when applied to the SSE data, had an average error of 0.275 degrees of visual angle. Thus, the achieved accuracy for the SSE data is what was mathematically anticipated for data with no crosstalk or gain changes introduced. This validates the SSE modelling of fixation error and subsequent calibration potential of the MLPs.

**EOG and IRIS® Data**

In calibrating real eye movement data, the MLPs again produced the best result, with a mean error of just less than 1 degree of visual angle for the IRIS® data. The LPs calibrations of real data were similar to the MLPs for the IRIS® data but 50% less accurate for the EOG data.

The calibration of the IRIS® data is similar for both LPs and MLPs. This suggests that the pattern space is linearly separable, that is, no non-linear transformations of the data have occurred. Hence, the lower performance of the LPs on the EOG data, although influenced by the presence of physiological noise and eyelid artefact in the vertical channel, is possibly due to presence of some non-linear transformations of the signals due to the biological nature of the EOG signal source.

The results for MLPs on simulated EOG data suggests that MLPs are capable of performing the necessary mappings for calibration, so the overall trend to less accurate calibrations for real data, both EOG and IRIS®, is most likely due to additional individual information within the data. One possible source of additional information would be head movement. However, due to base line correction of the data for each trial during pre-processing, any changes in potential due to small head movements between trials are not expected to be a source of calibration error. Without the use of a more
restrictive head brace or bite-board, some small head movements occurring during the execution of a saccade are possible, but were not evident when observing subjects undertaking the task. The use of a more restrictive head restraint, such as a bite bar, is not advisable using EOG (due to EMG interference), and perhaps not even desirable within clinical settings. Another possible source of error could be the eyelid artefact in the vertical channel, which was not introduced into the simulated data.

**Automatic Pre-processing**

The development of an automatic algorithm for processing saccadic EOG subject data provides a useful tool for researchers using EOG in this way. The ability to automatically reject blinks, excessively noisy trials and those in which the subject fails to comply with the task saves considerable researcher time in processing data. The algorithm also finds the saccadic onsets and terminations and can extract a nominated set of data points around these critical occurrences to represent the saccades. The algorithm is already effective on infrared tracker saccadic data and development work is continuing to allow it to deal with smooth pursuit eye movement data.

**Limitations**

The lack of a more restrictive head restraint and fixation errors inherent in the task were probably the reasons why the IRIS® did not perform at its full potential. The simulation experiment verified the ability of the MLPs to do the necessary mappings for calibration and provided a resolution consistent with fixation errors modelled by the SSE (i.e. 0.3 degrees of visual angle). This modelling of the fixation errors inherent in the human visual system assumes a consistent, genuine effort by subjects to fixate on the targets.
as best they can. Figure 22 shows that both systems indicate that subjects consistently undershot the targets on the larger saccades. If subjects fail to correctly fixate the targets, or keep their heads still to the best of their ability, due to tiredness or lack of interest, then a resolution of 1 degree of visual angle from the more accurate IRIS® tracker is probably the best result that could be expected. With MLPs as calibrators, both methods provided this result. This also suggests that the MLPs were capable of dealing with the eyelid artefact in the EOG vertical channel.

One of the current restrictions on the present use of the system is the large number of trials that subjects are required to perform. Half of these trials are for training the MLPs, which may account for some of the reason why the resolution achieved with EOG (and IRIS® data) appears to be restricted by the fixation error of the task rather than the physiological noise within the EOG. The standard error of the mean is inversely proportional to the square root of the number of samples, and as this effect reduces the error in establishing a particular fixation because of the number of points available in defining that fixation, so to in a similar way, might the large numbers of training trials have helped the MLPs to deal with the physiological noise in the EOG and provide the same resolution as the IRIS® data. The use of such a large number of trials, and to a lesser extent, the re-fixation of a central point before each trial, restricts the research useability of the EOG-MLP system in its present stage of development.

Future Directions

A future experiment is planned that will extend the saccadic task to all four quadrants and also incorporate a 9 point calibration grid. Subjects will
perform the calibration task, similar to that outlined by Stampe (1993), and
then be required to make fixations on a series of targets (of known
coordinates) appearing randomly in all quadrants, without re-fixating on a
central fixation cross before each target appears. This experiment will test the
ability of the EOG-MLP saccadic calibration system to be applicable to a
much wider range of eye movement research. The MLPs will be trained using
the calibration data obtained pre- and post-trial. Assuming such a small
number of trials are adequate for training is an empirical question that needs
testing. If so, training should be a lot quicker. Attempts will also be made to
further speed up training of the MLPs by selecting an initial set of network
weights that are not randomly selected but based on final weights from prior
training on subject data. The experiment will allow for the use of the bi-
quadratic calibration method and provide a direct comparison between
contemporary use of infrared trackers and the EOG-MLP system. It will also
provide the opportunity to compare the effectiveness of MLPs versus the bi-
quadratic method for the calibration of IRIS® data.

The ability of feedforward backpropagation MLPs to deal with non-
linear transformations should enable them to learn to recognise eye blinks and
double movement saccadic data, and hence provide a means to further
automate the calibration of 2D EOG. The algorithms described in the present
studies can also be modified to provide a functionality that adapts the accuracy
of calibration by displaying more targets and recording data at those locations
in the visual field where accuracy is unacceptable. This may allow the EOG-
MLP system to be also useful in research that does not provide specific visual
targets but attempts to identify where subjects are looking as they scan the
visual field. Another possibility for the MLPs is to extend the use of the EOG to much larger visual angles than the linear range (±30°) most commonly used. A small range of visual angle is a limitation of many eye measurement techniques (Oster & Stern, 1980), especially the optical methods, and continues to this day. Another important advantage of the MLP-EOG system is that the necessary standard equipment is relatively inexpensive and is available in laboratories and hospitals where basic physiological responses are routinely recorded.

MLPs should also be effective in other applications requiring calibration of 2D signals where the two signals cannot be guaranteed to have the same gain or to have been obtained from orthogonal, correctly aligned recording axes. The application of MLPs to calibrating smooth pursuit eye tracking data is one possibility. A simplified infrared eye tracker built into a pair of glasses, with little concern for complete orthogonality of the sensors is a practical example, but the application need not be restricted to measurement of eye movements.
REFERENCES


Vision Research, 31(11), 2029-2032.


Kobayashi, S., Hara, T., & Goi, Y. (1996). Predicting changes in drivers’


Psychiatry and Clinical Neurosciences, 54(3), 276-277.


APPENDIX A

Details of Multi-Layer Perceptrons (MLPs)
The input matrices, containing the concatenated horizontal and vertical saccadic components, were re-scaled to ± 1 to make training the ANNs more efficient (Demuth & Beale, 1994). The training matrices were also re-scaled to be just within the range of ± 1. This ensured that no actual target coordinates lay on the saturation limits of the output activation functions. This process of re-scaling also helped prevent premature saturation of the output and hidden neurons (Schalkoff, 1997). Each training session started with a random allocation of weights and biases between the input nodes and the hidden nodes, and between the hidden nodes and the output nodes. The Matlab® function used to initialise the MLPs was initff.m. (for details on Matlab® supplied functions see Demuth & Beale, 1994).

The MLPs were trained using a version of the Widrow-Hoff learning rule generalised to multi-layer networks and non-linear differentiable transfer functions (generalised delta rule). The following information concerning this technique comes from the Neural Network Toolbox User’s Guide (Demuth & Beale, 1994). The Widrow-Hoff learning rule is used for linear perceptrons that have no hidden layers and linear transfer functions for the output neurons. The rule uses the derivative of the sum-squared error ($sse$) with respect to the weights and biases to find the direction in which calculated small changes in the weights and biases will decrease a node’s error. A learning rate ($lr$) acts as a gain on the weight and bias changes. The derivative of $sse$ with respect to a weight $W(i,j)$, from input $j$ to neuron $i$, for a single input vector $p$ and target vector $t$ is
where the network’s error $e$ is the difference between the networks output and target $t$. The rule is implemented by absorbing the constant 2 in the equation, into the learning rate $lr$, and making changes to the weight in the direction opposite to the direction that error is increasing.

$$\Delta W(i, j) = lr \cdot e(j)p(j)$$

which in matrix form (across all training vectors) becomes

$$\Delta W = lr \cdot E_p$$

As biases are simply weights that have a constant input of one, the bias update expression in matrix form becomes

$$\Delta b = lr \cdot E$$

The generalised version of this rule when applied to multi-layer networks and non-linear differentiable transfer functions uses identical expressions for the changes to be made to a layer’s weights and biases, except that the error matrix $E$ is replaced with a matrix of derivatives $D$. Derivatives
of error (delta vectors) are calculated for the output layer using the error vector, then backpropagated through the network allowing delta vectors to be calculated for the hidden layers. The delta vector for a hidden layer is calculated from the following layer’s delta vector and weights. The Matlab\textsuperscript{®} function used to train the MLPs was trainbpx.m.

Momentum and adaptive learning rate were available to speed learning. Momentum allows the network to respond to recent trends in the error surface as well as to the local gradient. An adaptive learning rate attempts to keep learning stable while keeping the learning step size as large as possible. Momentum (mom) was not employed in this instance as it led to premature saturation problems. For similar reasons the option to allow the learning rate to increase (lr\_inc) was also not employed, although the learning rate was allowed to decrease (lr\_dec). If the new error exceeded the old error by more than a predefined ratio (err\_ratio), the new weights, biases, and output were discarded and the learning rate decreased by a certain ratio (lr\_dec). Training continued until either the sum-squared error goal (err\_goal) was reached or until the maximum number of training epochs (max\_epoch) occurred. The maximum number of training epochs required for convergence was estimated by viewing training curves of calibration performance vs. number of training iterations. To avoid premature saturation problems, some trial and error experimentation was necessary to arrive at a workable lr parameter, whereas the other parameters used the Matlab\textsuperscript{®} default values. The value of the parameters used were as follows

\[ \text{lr} = 0.003 \quad \text{mom} = 0.0 \quad \text{lr\_inc} = 1.0 \]
\[ \text{err\_ratio} = 1.04 \quad \text{lr\_dec} = 0.7 \quad \text{err\_goal} = 0.01 \]
max_epoch = 5000 (10000 for 5 & 10 hidden nodes).
APPENDIX B

Details of Automatic Pre-processing Algorithms
Experiment 1

The algorithm, clean_M.m, is summarised in the flowchart in Figure B1. The algorithm firstly removed the trials in which the word "blink" appeared instead of a target. These trials were included for future attempts at training the ANNs to automatically recognise blinks within the data. The algorithm then rejected trials that contained amplitudes beyond the range expected for the present task. Firstly, to eliminate d.c. drift, all trials were baseline corrected using the mean of the first 40 data points. Then the median maximum and minimum amplitudes were extracted from the clean range of the largest expected horizontal and vertical saccades (12 degrees of visual angle). Visual inspection of these largest saccades suggested that some acceptable trials had amplitudes 10 to 15% either side of the average amplitude, so any trial with a maximum value in its clean range 20% greater than the extracted median value for the relevant channel was rejected. This retained acceptable saccades with amplitudes larger than the expected maximum but rejected saccades containing EMG spikes, some other source of noise, or where subjects made too large an eye movement. A similar procedure was adopted to remove any trial that contained extreme negative values in either channel.

Although saccades were only made in one direction for both channels, legitimate negatives may occur due to negative crosstalk, and these negatives will have their greatest magnitude on the larger saccades (although in the opposite channel). The procedure was then repeated so that any initially rejected trials did not contribute to the median maximum or minimum values. As data was calibrated separately for each subject, the significant gender
Figure B1. The flowchart summarises the algorithm (clean_M.m) used to screen subject data in Experiment 1 prior to training with the MLPs.
differences in the EOG potential reported by Oster and Stern (1980) did not influence the screening process.

Trials that contained excess noise, or that were unstable during the initial fixation before making a saccade were rejected next. The standard deviation of the first 40 data points for each channel was calculated as an indication of the noise within that particular trial. This represented the 20 ms immediately following the subjects' button press indicating they were fixating on the central cross and ready for the next target. Hence the signal should be stable, and the calculated standard deviation should reflect the amount of physiological noise in the signal. Trials were rejected if this measure of their noise level exceeded 3 times the median standard deviation for the relevant channel calculated across the entire data set. These trials were potential outliers (Tabachnick & Fidell, 1989), and this indicated either that the trial contained excessive physiological noise, or that the subjects gaze was not stable prior to the target onset. This procedure was also repeated so that any initially rejected trials did not contribute to the median standard deviation of the data set.

**Experiment 3**

The data were screened using a computer algorithm (auto_cross2.m). Figure B2 shows a flowchart that represents the algorithm.

**Screening.**

The first procedure removed trials that contained blinks, excess noise, or those in which the subject failed to initially fixate on the cross before making a saccade. This was the same algorithm used in Experiment 1 (see Figure B1 & clean_M.m). Baseline correction of the data removed any
All trials

Remove blinks & trials with excess noise or amplitudes (clean_M.m)

Calibrate signals (calibrate.m) Estimate crosstalk (crosstalk.m)

Horizontal Channel

Smooth data (smooth.m) Determine best fit logsig fn (fmins.m, logsig_saccade.m) Search for onset of best fit logsig fn (onset_bf_v4.m)

Best fit logsig onset found?

Yes

Save onset data as zeros

No

Search for onset in smoothed data (onset_v4.m)

Smoothed data Onset found?

Yes

No

Save onset data as zeros

Offset found?

Yes

Search for offset (offset_v4.m)

Save onset index and sign

No

Reject trial

Vertical Channel

Smooth data (smooth.m) Determine best fit logsig fn (fmins.m, logsig_saccade.m) Search for onset of best fit logsig fn (onset_bf_v4.m)

Best fit logsig onset found?

Yes

Save onset data as zeros

No

Search for onset in smoothed data (onset_v4.m)

Smoothed data Onset found?

Yes

No

Save onset data as zeros

Offset found?

Yes

Search for offset (offset_v4.m)

Save onset index and sign

No

Reject trial

Does trial have any onsets?

Yes

No

Reject trial

Does trial have two onsets that are both negative?

Yes

No

Reject trial

If only one onset, is it too early or negative?

Yes

No

Reject trial

Was trial an oblique target?

Yes

No

(continued next page)
Figure B2. Flowchart summarising the automatic pre-processing algorithm (autocross2.m) for preparing subject data for presentation to the artificial neural networks.
translation effects of the recording axes due to electrode placement, any DC
drift of the EOG, and any changes in potential due to small head movements
between trials. This was done individually for each saccade and in both
channels.

Calibration and Crosstalk.

The amplitudes of the median wave forms for the 8, 10, and 12 degree
horizontal and vertical saccades were used to calibrate the signals to determine
eye velocity, the criterion used to determine saccadic onset. The amplitudes of
the median wave forms were the differences in the median values between the
initial 40 data points and the 50 points before the middle of the sampling
record (i.e. points 200 - 250).

Before rejecting a trial due to a lack of detected onsets, it is necessary
to determine the nature of any crosstalk present and how it may affect the
signals within each channel. The amplitudes of the median wave forms (in both
channels) of the largest horizontal and vertical saccades were used to estimate
the degree of crosstalk present in each channel. This was necessary so that the
effect of gain changes and crosstalk, as seen in Figure 16d, would not lead to
the automatic rejection of such a trial for having no onset in the vertical
channel. If no rotational transformations producing negative crosstalk were
present, an oblique target’s data would be expected to have saccadic onsets
detected in both recording channels.

Onset Detection.

After smoothing the data in both channels (applying a 25 point
Hamming Window filter), a best fit logsig function was used to locate the
vicinity of the eye movement onset. As the logsig function contained no noise,
its 'onset' was easier to find. The search for the saccadic onset in the real data was initiated at 10 data points before the best fit logsig onset. A slope criterion corresponding to an eye velocity over 2 data points of 30 degrees of visual angle per second (Van Opstal & Van Gisbergen, 1987) had to be satisfied 9 times in succession. Visual inspection showed this combination to be very effective in detecting the saccadic onset within the smoothed but still undulating signal. A similar algorithm was used to detect the termination of the saccade. This was followed by a search for a possible second onset occurring after the termination of the first saccade.

Trials were rejected if no saccadic terminations were found. A saccadic onset not followed by a saccadic termination most likely indicates the presence of a blink (which usually go off scale) or other non-task orientated gross eye movement. Detected onsets were required not to be too early or too late. Express saccades have a saccadic onset between 80 to 120 ms (Kingstone & Klein, 1993), so any saccades made earlier than this are likely to be anticipatory (i.e., erroneous) saccades. Onsets were considered to be too late if the target was likely to be switched off before the subject had time to fixate on it. The data indicated that the durations of the largest saccades, from onset to termination, were at least 60 ms. The target was on for 500 ms, so any saccade that started after 400 ms was considered to be too late to enable any reliable fixation on the target. This restriction was implemented by imposing an upper limit when searching for onsets. Trials were rejected if a second saccade was detected within 25 ms of the termination of the first saccade as this indicated that the subject performed a double saccade in order to fixate on the target.
Voltage deflections in the wrong direction in ideal data would represent negative saccades, that is, saccades to the left for the horizontal channel, or down for the vertical channel. In subject data these may be legitimate as they can arise from negative crosstalk. For example, if the horizontal channel was producing negative crosstalk in the vertical channel, then a 12 degree of visual angle horizontal saccade is likely to produce a signal in the vertical channel that resembles a smaller downward saccade.

Trials were rejected if no onsets in either channel were detected, both onsets were negative, or a single onset (i.e. an onset in one channel only) was too early or too late. A horizontal target trial was rejected if no horizontal onset was detected or it was negative or too early. A vertical target trial was rejected if no vertical onset was detected or it was negative or too early. Oblique target trials with both onsets and negative crosstalk to the vertical channel were rejected if both onsets were negative or the horizontal onset was too early. Oblique target trials with both onsets and negative crosstalk to the horizontal channel were rejected if both onsets were negative or the vertical onset was too early. The channel receiving the negative crosstalk may have its signal cancelled (see Figure 16d) or made negative, but if both onsets are negative then the trial is most unlikely to reflect a satisfactory execution of the task. Finally, oblique target trials with both onsets and no negative crosstalk were rejected if either onset were negative or the horizontal onset was too early. The horizontal onset was used to represent the timing of the saccade as the horizontal channel has the better signal to noise ratio and fewer artefacts.
**Saccadic Extraction.**

Using the automatically detected onsets, 100 data points were selected from the unsmoothed raw data from each channel to represent the saccade (eog100sac_autocross2.m). The first point of the selected saccade was 5 data points before the chosen onset. The onset detected in the vertical channel was used for vertical target trials. The onset detected in the horizontal channel was used for horizontal target trials. As the horizontal channel has the better quality signal, the horizontal onset was used for oblique target trials with onsets in both channels; otherwise the single onset was used.
APPENDIX C

Directory Structure for Attached CD
EXPERIMENT 1

Data

Preprocessing
expt1.m  
  reject.m  
  clean_M.m  
  m_blc_2048  
  reject_badtrials.m  
  noise_reject.m  
  selected_pointsM.m

Training Multi Layer Perceptrons
setup_subjects.m  
  sac_net_xydsetup.m  
    xysplit.m  
    xyrtassign.m  
    mjcinitff.m  
    xypc_dnet.m  
    analyse2.m  
    add_table_s.m

EXPERIMENT 2

Data

Generating Simulated Data
setup_dummy_r.m  
  gamma.generate.m  
    Polar_fix.m  
    amp_to_sac.m  
      main_seq.m  
    gamma.m  
    gammavel.m  
    dur_at_vc.m  
    amp_to_sac2.m
trans_sacdum_0_12scale.m

Scale0_12.m
bio_noise_after.m
scale_1.m

Training Multi Layer Perceptrons
sac_net_xyrsetup.m
xysplit.m
xyrtassign.m
mjcinitff.m
xy_dist_mlp.m
analyse2_r.m
add_table_r.m

Training Linear Perceptrons
setup_dummy_linff_r.m
trans_sacdum_0_12scale.m
bio_noise_after.m
scale_1.m
sac_net_xyrsetup_linff.m
xysplit.m
xyrtassign.m
mjcinitff_lin.m
xy_dist_lp.m
analyse2_r.m
add_table_r.m

Training Simple Classifier
compare_linff_r.m
trans_sacdum_0_12scale.m
bio_noise_after.m
ifxy.m
xy_dist_sc.m
xy_dist_mlp.m
xy_dist_lp.m
EXPERIMENT 3

Data

Preprocessing
auto_cross2.m
reject.m
clean_M.m
calibrate.m
typeext_accepted.m
crosstalk.m
smooth_row.m
logsig_saccade.m
logsac_fitH.m.m
logsac_fitV.m
onset_bf_v4.m
onset_v4.m
offset_v4.m
Onset2_v4.m
eog100sac_autocross2.m
combine_1scale.m

Training MLPs EOG Data
setup_subjects_r.m
sac_net_xyrsetup.m
xysplit.m
xyrtassign.m
mjcinitff.m
xy_dist_mlp.m
analyse2_r.m
add_table_s_r.m

Training MLPs IRIS Data
setup_subjects_r_iris.m
sac_net_xyrsetup_iris.m

xysplit.m
xyrtassign.m
mjcinitff.m
xy_dist_mlp.m
analyse2_r.m
add_table_s_r.m

Training LPs EOG & IRIS Data
setup_subjects_linff_r.m
sac_net_xyrsetup_linff.m

xysplit.m
xyrtassign.m
mjcinitff_lin.m
xy_dist_lp.m
analyse2_r.m
add_table_linff_s_r.m