GROUP EYE TRACKING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

OZAN DENİZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COGNITIVE SCIENCE

SEPTEMBER 2016

Group Eye Tracking

Submitted by Ozan Deniz in partial fulfillment of the requirements for the degree of **Master of Science in The Department of Cognitive Science Middle East Technical University** by,

Prof. Dr. Nazife Baykal
Director, **Graduate School of Informatics** ____________________

Assist. Prof. Dr. Cengiz Acartürk
Head of Department, **Cognitive Science** ____________________

Assist. Prof. Dr. Cengiz Acartürk
Supervisor, **Cognitive Science** ____________________

Examining Committee Members:

Assoc. Prof. Dr. Annette Hohenberger
Cognitive Science, Middle East Technical University ____________________

Assist. Prof. Dr. Cengiz Acartürk
Cognitive Science, Middle East Technical University ____________________

Assist. Prof. Dr. Murat Perit Çakır
Cognitive Science, Middle East Technical University ____________________

Assist. Prof. Dr. Erol Özçelik
Department of Psychology, Çankaya University ____________________

Assist. Prof. Dr. Umut Özge
Cognitive Science, Middle East Technical University ____________________

Date: 01.09.2016

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last Name : OZAN DENİZ

Signature :

ABSTRACT

Group Eye Tracking

Deniz, Ozan

MSc., Department of Cognitive Sciences

Supervisor: Assist. Prof. Dr. Cengiz Acartürk

September 2016, 68 pages

Social interaction relies on information obtained from the eyes to a large extent. However, most of the current eye movement research apply experiments where participants are recorded individually in separate rooms. Those experiments help explaining and model human visual system, individual saccade behavior and fixations under certain conditions. But they lack of explaining the social role of the eye movements. In this study, we design and develop a tool for analyzing the role of eye movements in social communication. In order to measure the role of eye movements in social cognition, we have developed a system called “Group Eye Tracking (GET)”. The software infrastructure of the GET is related to the network infrastructure of client machines, data communication among group members and data collection during the experiment. In this thesis, we describe the details about creating stimuli and designing group conditions running on the software infrastructure and report the analysis of group eye movement data.

Keywords: Eye Movements, Eye Tracking, Social Cognition, Human Computer Interaction

ÖZ

Grup Göz Takibi

Deniz, Ozan
Yüksek Lisans, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Yrd. Doç. Dr. Cengiz Acartürk

Eylül 2016, 68 sayfa

Sosyal etkileşim, büyük ölçüde göz hareketlerinin içerdiği bilgiye dayanır. Buna rağmen, günümüzde yapılan göz hareketi araştırmalarının çoğu deneklerin tek başına ve ayrı odalarda göz verisinin toplanması ile yapılmaktadır. Bu deneyler, insan gözünün optik sistemini, gözün sıçrama davranışlarını ve belirli koşullar altında gözün nasıl duraksadığını araştırmamıza olanak sağlamaktadır. Ama bu tür deneyler gözün sosyal etkileşimdeki yerini analiz etmemizde yetersiz kalmaktadır. Bu çalışmada, sosyal iletişimde gözün rolünü inceleyebilmek adına, bir ortam tasarlayıp geliştirdik. Geliştirdiğimiz ortam kısaca GET (Group Eye Tracking) ortamı, gözün grup içerisindeki davranışını modellemeyi amaçlamaktadır. GET ortamının yazılım altyapısı, istemci makineler arasında kurulan ağ altyapısı ve deney esnasında grup içerisindeki istemcilerin veri aktarımın modellenmesi üzerinedir. Bu tezde, grup göz verisinin ölçülmesi amacı ile tasarlanan uyaran ve grup koşullarının detayları anlatılacak, toplanan grup göz hareketi verisinin analiz sonuçları raporlanacaktır.

Anahtar Sözcükler: Göz Hareketleri, Göz İzleme, Sosyal Biliş, İnsan Bilgisayar Etkileşimi

To my Blue Pale Dot

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to my supervisor, Assist. Prof. Dr. Cengiz Acartürk, for his motivation, guidance, encouragement and support through my studies.

I would like to thank to my teammates in HAVELSAN. And also I would like to thank to Mehmetcan Fal and Maani Tajaldini for their assistance in COGS Lab.

I also thank to Semih Yağcıoğlu for kindly accepting to spend their valuable time to review and evaluate my thesis.

I express my gratitude to my father Üstüner, my mother Ülkü and Özge for always being motivating and supportive. None of this would have been possible without their love, selflessness and sacrifices they made on my behalf.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1 GROUP EYE TRACKING PARADIGM

1.1 Introduction

Group cognition emerges from the interaction of individuals in the group. In the recent state of the art, metrics have been developed to measure and model the group behavior. For instance, Moussaïd, Perozo, Garnier, Helbing, & Theraulaz (2010) used pedestrian walking patterns as a metric of group behavior. Gaze dynamic metrics can have a significant role in modeling gaze behavior in group cognition, in a similar way. Recent studies on group cognition reveal that gaze dynamics, especially fixations and saccades can model the individuals' behavior in different group conditions (Zheng, Hajari, & Atkins, 2016; Deniz, Fal, Bozkurt, & Acartürk, 2015; Pfeiffer, 2012).

There are several gaze dynamic metrics for detecting similarities among the different eye movement datasets. Most of these metrics can be applied to detect the similarity between the two different datasets, such as to scanpaths (Holmqvist, 2011). However, there may be more than two participants in a group (Tajfel, Billig, Bundy, & Flament, 1972). Therefore, group dynamic metrics are needed to model the individuals' gaze behavior in the group without depending on the member count.

In this thesis, we introduce the GET (Group Eye Tracking) environment and we report the experiments that used the GET environment. In this analysis, we use three eye movement metrics which can be applied to more than two participants' data. These metrics are convex hull area, circular area and jaccard index. In chapter 4 we analyze the convergence metrics in the current state of the art with these three metrics. Moreover, we make benchmarks and report our experimental results for these three metrics.

The GET environment has been developed to model the group gaze behavior with gaze dynamics. We designed a game using the GET environment which we call "balloon game" hereafter. In the experiment, participants played the "balloon game" with their eyes on the GET environment. We have collected and analyzed participants' eye movements while they played the game.

During the game experiment, the stimuli were displayed on different screens to each participant in the same room. Each group member was able to see other group members' eye movements on the screen by means of a visual cue. Group members' eye movements can be displayed as a circular cursor or any designed object on the screen overlaid onto the stimuli.

1.2 Minimal Group Member Paradigm

How many participants does it take to make a group? Minimal group member paradigm employs this question in social psychology. In the following, we address this issue and explain the minimal group member paradigm in detail and its relationship with the GET paradigm.

In social psychology, the group consisting of two people is called "**dyad**". In the "dyad" condition, no significant group effect is observed. Therefore in dyad condition, none of the group members focus significantly enough on the actions of the other participant. However, in a "**triad**", where a third participant is involved, prior work shows that one participant is influenced by the remaining two participants' actions in the group (Tajfel, et al., 1972). There is a significant difference between a dyad and triad. Interestingly, this group effect goes down in groups constituted by four participants. In four group member case, participants act as if there are two separate "dyad" in the group. Choosing odd number of members for small groups can overcome this situation (Menon & Phillips, 2011).

In what follows we describe the motivation behind this thesis and describe the details of using the GET paradigm, where we measure the gaze patterns of the participants in the group. Moreover, we describe the rationale behind the number of members in a group for our experiments in the "balloon game" experiment. The group conditions and other parameters for the experiments are discussed thoroughly in the chapter.

1.3 Motivation of the Thesis

In this thesis, we want to answer two questions. The first question, how can we model the group members' gaze behavior in different group conditions? The second question, is there any effect of gaze awareness in different group conditions? Throughout this thesis, we aim at answering these two questions by designing an environment and conducting several experiments to back our assumptions. Finally, we conclude this thesis with the findings we gather from those experiments.

In this study, we use the eye tracking methodology for modelling the individuals' gaze behavior in different group conditions. The major focus of the eye tracking methodology is to understand whether eye movement patterns exhibit a random sequence, or exhibit a structural pattern given the data in the external environment.

We collect participants' eye movements under a set of different group conditions. These group conditions are, "3 gaze necessary" (3GN), "single gaze necessary" (SGN) and "all enemies" (AE). The 3GN condition stands for cooperation. The other two conditions stand for the competition, as will be explained in the following chapters.

In the current state of the art, there are algorithms to detect the similarities between two eye movement datasets (e.g. scanpaths or gaze maps). These algorithms mainly use position dispersion metrics as an input (Holmqvist, 2011). However, there are also algorithms which use scanpaths of eye movements as an input (Holmqvist, 2011). In this thesis, we analyzed several eye movement similarity metrics. Subsequently, we benchmarked with three of those metrics to measure the competition and the cooperation among group members.

In chapter two, we analyze the historical background of eye movements and eye tracking and its usage in gaming and human computer interaction platforms. In chapter three, we explain the software infrastructure and hardware specifications of the GET. In chapter four, we analyze the popular gaze convergence metrics used in current eye movement research. In chapter five, we introduce the application designed for the GET and we also explain the group conditions of application designed for the GET. In chapter six, we display and analyze the results computed with the selected gaze convergence metrics. In chapter seven, we discuss our findings for the GET environment and explain our future work plan for the GET environment. In the following, we will make a literature survey about the evolution of the eye tracking and dual eye tracking and their usage in human computer interaction systems.

CHAPTER 2

2 LITERATURE SURVEY

2.1 Historical Background of Eye Movement Analysis in Social Psychology

Eye movements and gaze dynamics has been used in social psychology as a metric since late 1960's (Kleinke, 1986; Yarbus, 1967). Eye movements were used for finding out the relationship between the gaze and the attitude (Kleinke & Pohlen, 1971). Early eye tracking studies provided information about attentiveness, competence, social skills and mental health, credibility, dominance and communicating feelings (Kleinke, 1986).

One of the first eye movement experiments used in social psychology was a game called "Prisoner's Dilemma" (Kleinke & Pohlen, 1971). There are two aims for this study. The first one is to measure behavioral variables on reaction of gaze and non-gaze. The second one is to find a relationship between the gaze's role and emotional arousal. Results of this study showed that subjects rated generally positively when they interact gazing compared with non-gazing.

Collecting accurate eye movement data is important for analyzing the gaze patterns. In the beginning of the eye movement studies, researchers tried to collect eye movement data manually. The most common procedure for measuring gaze is to have two or more observers stand behind a one-way mirror. They press buttons which are connected to clocks when the participant directs his/her gaze toward the face of another. They also counter when the participant directs his/her gaze toward the face of another. The researchers also can record the experiment to observe the gaze change under different condition (Kleinke, 1986). There were diffuculties on collecting and measuring eye movement data with manual methods. However, several eye movement metrics developed for behavioral analysis. These metrics were gaze duration, gaze frequency and glance duration. Gaze duration was reffered to the length of time one person gazes at another and is the measure reported in most studies. Gaze frequency was referred to the number of glances made by one person toward another. Glance duration was computed by dividing gaze duration by gaze frequency.

Since 1960's, there have been many important advances in eye tracking technology. Computing the fixations and user gaze patterns are much more easier than the past. Moreover, the eye trackers have become more reliable and cheaper. These improvements give researchers the oportunity of designing more human intuitive and efficient eye tracking platforms. In this thesis, we have developed a multi-user eye tracking platform (Group Eye Tracking). However, there isn't a study focusing on group eye tracking (platforms which have more than two participants) in the current eye movement literature. In the following, we will analyze the closest studies to our thesis in the literature.

2.2 Using Eye Movements in Human Computer Interaction

With the evolution of the technology, the focus of the studies has been developing more human-intuitive input devices for better user experience in human computer interaction. In 2003, Sony produced the EyeToy. EyeToy is a camera which is connected to a PlayStation 2. EyeToy tracks the body gestures of the players. EyeToy allows the players to control the on-screen characters by moving their bodies (Austin, Mateo, Hansen, & Villanueva, 2009). In 2005, Nintendo produced the Wii Remote. Wii Remote is a novel gamepad for Wii. The Wii Remote includes an accelerometer and optical sensor technology. This allows games to be controlled by moving the pad in three-dimensional space (Austin, et al., 2009). In 2010, Microsoft produced the Kinect for XBOX 360. The aim of the Kinect is to play the games only with the body gestures without using any keyboard, mouse or joystick.

The quality of eye trackers has been increasing while the costs of the eye trackers have been decreasing. Therefore, eye tracking can be used for an input device in games and applications. The history of using eye trackers in games is older than other input devices which we talked above. One of the first study using eye trackers in human computer interaction was made by (Starker & Bold, 1990). In this study, they display planets to the participants as a stimulus for a while. After the pre-set time, they calculate each objects' interest-stamp. Interest-stamp of an object refers the interest level of participant to that object. Interest-stamp of an object stands for the participant's glance time to that object. After calculating each objects' interest-stamp, the system choses an object which has the most interest-stamp. After that, the system tells a story to the participant about selected object. This dynamic story-teller game may not be considered as a human computer interaction system which we are researching today. However, the underlying philosophy is similar. The participants explore the environment in the game with their eyes. With the information obtained from the environment, they change the game's state by using their gaze.

After Starker & Bold's (1990)'s story-teller game, gaze-controlled human computer interaction applications have been developed in numerous fields. Some of these applications developed for helping disabled or elderly people to make the daily life easier (Acartürk, Freitas, Fal, & Dias, 2015). In our thesis, we developed a **gaze-controlled game application**. In the following, we will analyze the benchmarking studies and design issues of gaze-controlled applications in the current literature.

There are many studies focusing on benchmarking the gaze input combined with other modalities (speech, haptic etc.) against mouse and keyboard. Isokoski & Martin (2006) developed a "first person shoout" game to benchmark the accuracy of eye movement control against mouse control. The results seem to be promising. Configuring the eye tracker may positively effect the player’s performence. Players configure the mouse and keyboard’s settings for better gaming experience in general. They can configure the eye trackers to get better results while playing the game as well.

Another study on benchmarking gaze and voice input modalities against keyboard and mouse was made by Donovan, Ward, Hodgins, & Sundstedt (2009). The name of the game is “Rabbit Run”. Players try to escape from a maze while they are tracking by evil rabbits. The game can be played with both <gaze, voice>(GV) pair and <keyboard, mouse>(KM) pair. They benchmark the usage of GV against KM. The find that in GV, there are collusions in responses. Because of that, KM is far better than GV with respect to user experience (UX). However, the study points that in GV, players do not need to use their hands. This will help disabled people to use such human-computer interaction environments by using GV.

Mostly, benchmarking is made between several input devices (eye trackers, mouse, keyboards etc.) in terms of accuracy. However, Djamasbi & Mortazavi (2015) made a study to benchmark the user experience of gaze input in different generations. The selected generations are “Generation Y” and “Baby Boomers”. Generation Y refers the people born between 1977 and 1994 in population. Baby Boomers refers the people born between 1946 and 1964. They compared the gaze usage in between two different generations. The game’s genre is “memory”. To play the game, user must remember and repeat the sequence that is played by the computer. They developed three different types of interaction. These interactions are; “gaze & gaze”, “gaze & blink” and “gaze & click”. The results of this study show that the younger generation had a better gaze interaction experience than the older generation. Younger players reported better scores for the experience of control, naturalness, and likeability of gaze as an interaction method. In the following, we will discuss implementation and design issues for gaze-controlled applications.

According to Isokoski , Joos, Spakov, & Martin (2009), there are four different ways of implementing eye control in games. First one is using eye movements as a mouse cursor. Users move the mouse cursor in the game by using their eyes. This implementation does not require any additional implementation. Because all of the modern games can be played with the mouse. And also, the modern eye trackers provide users to control the mouse cursor with their eyes. Therefore, there is no need for additional implementation for playing modern games by using the eye movements as a mouse cursor. The second solution is using additional software for controlling the game with the eye movements. This additional software need depends on the game condition. For regular usage in modern games, users use keyboards and mouse correspondingly. There are two different input device to control the game. However, for gaze usage, users can only use one input device. To overcome the lack of input device problem, we can use additional software between the eye tracker and the game. This additional software adapts the eye movements to the game conditions. The third approach is changing the game’s source code to adapt for the eye movements. The selected game’s source code may be available online. We can acquire the source code from online and modify the source code for eye movement usage. The fourth approach is developing the game from scratch. This approach is the most challenging and time-consuming one.

Istance , Vickers , & Hyrskykari (2009) made a study on playing gaze-based massively multiplayer online games (MMOG). In the study, novel input signal instead of classical usage of mouse – eye tracker connection. In classical eye tracker usage, eyes stand exactly for mouse cursor. However, in Istance et al. (2006)'s study, they create gaze patterns standing for certain functions in game. The results are promising and participants can use the patterns to play the game.

ISO 9241-9 is a standard for the requirements for non-keyboard input devices. The first study for evaluation of eye tracking for ISO 9241-9 is made by Zhang & MacKenzie (2007). There are three techniques developed for evaluation. These techniques are Eye Tracking Long (ETL), Eye Tracking Short (ETS), Eye + Spacebar (ESK). ETL stands for fixating on a target for 750 milliseconds to select it. ETS stands for fixating on a target for 500 milliseconds to select it. ESK stands for pressing spacebar just after fixating on a target. The evaluation metric is based on the user's mean movement time. Participants try to select an item during a preset time (2.5 seconds). If the dwell time is too short, participants accidentally may select the wrong target. If the dwell time is too long, participants became impatient while waiting for selection. The ETL technique has a lower score than the ETS technique. The ESK technique was the best among the three eye tracking interaction techniques. According to the study, participants generally liked the ESK technique. Actually, in ESK participants use additional input device for selecting item. This finding also validates the Donovan, et al. (2009)' study.

Today there are many game genres. The problem is adapting the different input modalities (gaze, speech, haptic etc.) into the different game genres. Isokoski et al. (2009) analyzed the eye tracker compability of different game genres. They list the game genres for gaze compability analysis. The genres are "board", "card", "shoot-em-up", "beat-em-up", "first person shooters", "flight simulators", "3rd person action and adventure", "turn-based strategy", "real-time strategy", "turn-based role playing", "real-time role playing" and "racing" (Isokoski et al.). Board and card games are turn-based games. In board and card games, users generally do not need to give quick reactions. They usually think a while and then make a move while they are playing the game (chess, go, checker etc.). Therefore, there is no need for additional implementation for gaze-control. The regular board and card game can be played with eye trackers without additional implementations. As stated above, modern eye trackers have the functionality to control the mouse cursors. In the following section, we address the design problems encountered in eye movement studies.

The first problem is that most of the modern games (especially "first person shoot" and "shoot-em-up" genres) require constant and frequent control in position. Players should look at the target to hit. However, players also should stay out of the target's shooting range if the target is defending itself (shooting to player). During the game, players generally tend to follow the target for running away from it. In gaze-control, players cannot shout the target while they are trying to run away from the target. There should be additional implementation for playing the game (for instance adding another input modality such as speech). We can say that there are restrictions for designing game with only gaze-control. Most of the modern games are designed for playing with mouse and keyboard. During the games, players should check the board's state and they also should make move with their eyes as well. This effects negatively the usage of the game. Because designing the game with only gaze-control, we assign one input (eye movement) to two functionalities (searching the environment and making move). The researchers should take into account this problem while designing the game.

The second problem may occur because of the size of the components used in gaze-controlled games. Wilcox, Evans, Pearce, & Sundstedt, (2008) developed a "puzzle" game. This "puzzle" game can be controlled by using eye and voice. In the game, they have used a smoothing filter for displaying the cursor on the screen for better accuracy. They found that the components (buttons, cursors etc.) used in the gaze-controlled games should be bigger than the components used in regular games for better user experience.

The third problem in gaze-control game design was pointed out by Donovan et al. (2009). The problem is called "Midas touch". Players normally send the commands to the game when they press the button on mouse or keyboard. However, in gaze-control games, when players look somewhere, they immediately send a command to the game. There should be a difference between looking somewhere in the game and sending command to the game. As stated above, in gaze-control games, gazing has two functionalities. Donovan et al. overcame this problem by adding voice recognition to the game.

In this thesis, we have developed a gaze-controlled game running on group eye tracking (GET) infrastructure. In the game participants try to pop the balloons with their eyes. We developed the game from scratch. We keep the game conditions as simple as possible to overcome the design and implementation issues which we addressed above. In our game, we have used the gaze as an only input modality. In the game, participants should follow only the balloon on the game board. However, in **gaze-cueing** game condition, participant can follow each other's eye movements. We will give further explanation about the game in section 5. In the following, we will analyze the dual eye tracking (DUET) studies which have the similar infrastructure to the group eye tracking (GET).

2.3 Dual Eye Tracking

Dual eye tracking(DUET) is a novel paradigm in eye tracking methodology. The aim of the DUET is analyzing the gaze patterns of participants while they are doing the tasks on shared application. In the DUET, there are two participants. The participants' eye movements are recorded while they are doing the task. The DUET is the simple version of the GET (Group Eye Tracking). In the DUET, maximum two participants can do the task. However, in the GET, there can be more than two participants. Below, there will be the studies conducted on the DUET infrastructure.

Brennan, Chen, Dickinson, Neider, & Zelinsky (2008) made a study to analyze gaze-speech effect on collobarative visual search. They used different modalities for measuring the effect. These modalities are shared-voice, shared-gaze and shared-gaze-and-voice. In shared-voice, participants can only hear each other. In shared-gaze, participants can see each other's eye movements as cursor. In shared-gaze-and-voice, participant can hear and see each other's eye movements. They found that adding voice modality to gaze effect negatively the people. Because coordination by speaking takes time. They also found that people can coordinate each other better with only shared-gaze modality.

Another study made for analyzing peer-programming by Bednarik, Shipilov, & Pietinen (2011). In extreme programming paradigm, peer-programming helps programmers to be more productive. Programmers create two member groups. While one programmer are coding, the other programmer checks the errors in the code. They change the roles while they are programming. In the study, an expert explain two algorithms to a novice programmer. There are two conditions. In first condition, there is a gaze-animation. In second condition there is no gaze-animation (novice programmer cannot see the expert's eye movements). The results show that there is a positive effect of gaze-animation. In gaze-animation condition, novice programmer's gaze pattern has less variation when compared with non-gaze-animation. Gaze animation guides the novice programmer and increases the collobaration.

Meijering, van Rijn, Taatgen, & Verbrugge (2012) have developed a "Marble Drop" strategy game for analyzing the theory of mind. "Marble Drop" game can be played with two people. In this study, participants cannot see each others' eye movemements (no gaze-cueing). They collected the eye movement data after the experiment. They compare the participants eye movement and their moves in the game. Later they found a relationship between the eye movement and theory of mind.

Cherubini, Nüssli, & Dillenbourg (2008) have analyzed the correspondence of partners' eye movements during problem solving. In this study, partners sit in two different rooms and cannot see each others eye movements (no gaze-cuing). Cherubini et al. found that at similar timestamps pairs look at the same things above chance level.

Olsen, Ringenberg, Aleven, & Rummel (2015) analyzed the gaze pattern as a metric of joint attention. They developed an "Intelligent Tutoring System" applcation. In the application, participants try to solve the problems on their shared screen. The participants cannot see each others eye movements. There are three features in tutoring system. In first feature, only one participant can give the answer and the other participant assists. In second feature, participants has the different piece of information about the task and they should share the information with his/her partner. In third feature, participants answer the question individually before seeing his/her partner's answer and then the system asks to each participant to guess the consensus answer (cognitive group awareness). During these tasks, Olsen et al. analyzed the joint attention. They aimed at answering three questions. First question is "Is there any effect of talking (communicating during the task) to joint attention?". Second question is "Is there any effect of task's feature (features in tutoring system which we described above) to joint attention?". And third question is "Is there any effect of learning procedural and conceptual knowledge to joint attention?". They didn' find an effect of success of the task on joint attention. However they find an effect of talk on joint attention. When the participants talk to each other, the joint attention increases. Another founding is the type of the task has an effect on joint attention.

Patrick, Nüssli, & Weifeng (2010) developed a tetris game to measure the expert and novice players' collobaration. They used DUET (Dual Eye Tracking) paradigm. The players control the pieces in tetris with their eyes. Pieces have different colors for each player. There are two players in the game. One player is novice and the other is expert. Piece color is just an identifier for the player. There is a colloboration among the players because the players have the same score on the board. Patric et al. found that players adapt their behaviours to the social context of interaction.

We have analyzed the joint attention and dual eye tracking paradigm in the current literature. In our study, we extend the dual eye tracking paradigm. Dual eye tracking (DUET) is not generalizable to group eye tracking in terms of the gaze convergence measures. We have developed a Group Eye Tracking (GET) infrastructure which enables the multiplayer human computer interaction platform with gaze modality. In this study, our aim is developing and benchmarking the measures for joint attention in group members.

In addition to this, we developed a multiplayer game for the GET. Whereas the single player games, Swalwell (2006) has observed that players find multiplayer games more attractive. In chapter 5 we will explain the details of the multiplayer game and in the following chapter we will explain the infrastructure of GET platform.

CHAPTER 3

3 THE INFRASTRUCTURE OF THE GET

3.1 Software Infrastructure

The group eye tracking software environment (GET) mainly consists of two parts which is depicted in Figure 1. The first part is the client and the second part is the server.

In the server side, we distribute raw eye movement data over multiple clients synchronously. We have developed an application for the data distribution and this application listens to a certain port in the server for the incoming eye movement data and sends data back to the connected clients. There should be a correlation among the client machines with respect to their timestamps. Therefore, the application on the server provides the timestamp information for eye movement data.

Each client in the local network sends its data to the server and listens the server for the incoming data from other clients. The data consists of raw eye movement data which is being collected from the participant and timestamped with the system clock. The eye movement cursors can have different color or solid color according to experiment design.

The client application has been developed with C# language and .Net 4.0 framework. We have used C#'s build in TCP socket communication library to connect to the server. We have developed the client application on Visual Studio (VS) 2013 Integrated Development Environment (IDE).

The TeamViewer client application is running on client. We use TeamViewer application to receive server's screen. The game is running and visualizing on the server machine. We display the game screen on client machine by using TeamViewer application.

There are three applications running during experiment. First one is C# client application which takes the raw eye movement data, smoothens it then send it to server. Second one is C# server application which gets smoothed eye movement data from clients and visualize the game board based on coming smoothed eye movement data. Third one is TeamViewer application. TeamViewer application runs on both server and client side. We are using local area network for distributing the eye movement data to the clients. Therefore, there is no significant delay in data transmission.

The structure of collected data is as follows [x coordinate, y coordinate, timestamp, group id]. The x and y coordinates represent the raw eye movement coordinates on the screen. After data collection, the fixation and saccade information might be extracted from raw eye movement data, should it be necessary.

We have used Eye Tribe's C# Software Development Kit (SDK) to connect to the eye tracker's driver. We have implemented a listener for incoming eye movement data. The data coming from Eye Tribe's SDK is raw data. We have processed the data for smoothing. There are spikes in raw eye movement data. Visualizing these spikes can affect the experiment design. To overcome this issue, in client machines, we smooth the eye movement data coming from the eye tracker. For smoothing operation, we are taking the mean of the raw eye movements. In the first step, we define small time windows (around 200 milliseconds). An average fixation duration is 200-250 milliseconds (Serano & Rayner, 2003). Therefore, we are using 200 milliseconds time windows. Then we take the mean of the data (x and y coordinates) in these time windows for visualizing. We display the smoothed data "online" to the participant.

The server application has been developed with C# language and .Net 4.0 framework. The server listens the coming smoothed eye movement data from clients. The game application is running on server side. According to the coming eye movement data, server change the state of game. The TeamViewer application is running on the server for sharing its screen to the clients.

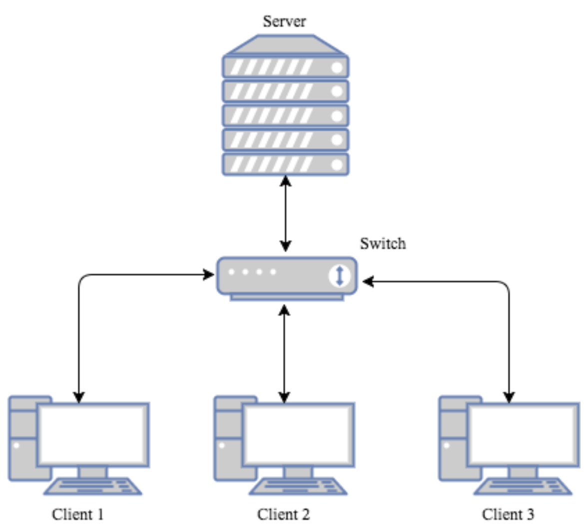


Figure 1: Group Eye Tracking Infrastructure

3.2 Hardware Infrastructure

The eye movement data is recorded by a 30-60 Hz Eye Tribe eye tracker, with a typical accuracy of 0.5 degrees and a typical spatial resolution of 0.1 degrees. The raw eye movement data set is the output of Eye Tribe API, which is the software provided by the manufacturer.

The group member data is distributed on a Windows based Intel(R) Xeon(R) E3 3.20GHz computer on 32 GB ram size.

CHAPTER 4

4 MEASURING GAZE CONVERGENCE IN GET

4.1 Metrics

4.1.1 Basic Position Metrics

There are two groups of basic position metrics. These metrics are position, and landing position in area of interest (Holmqvist, 2011)**.**

Position metric is represented by raw eye movement coordinate or extracted fixation x and y coordinate list or dwell set. In eye movement data analysis, most reliable information is raw eye movement data. Raw data can be in Cartesian (x, y) or Polar coordinate form (r, θ). Data consists of x and y coordinates. However, binocular raw data has one more dimension. It stands for the relative distance between eyes. This dimension is extracted from raw data. We know that the raw x and y coordinates may contain noise because of several reasons (eye tracker calibration, environmental conditions etc.). We compute the third dimension with x and y coordinates. Therefore this extraction process might lead to noise (Holmqvist, 2011). Therefore, in eye movement data analysis, we use two dimensional position data. Cartesian coordinate system represents eye movement’s position on screen. In Figure 2, sample eye movement data is scattered in Cartesian coordinate system and circles stand for the raw eye movements. Radial coordinate is calculated with participant’s distance to screen. Generally, in eye tracking systems, the origin of the location is at the top-left corner of the screen.

Landing position in area of interest is generally used in reading researches. The stimulus screen is divided into multiple areas. (AOI – Area of Interests) Landing position information can be letter pixel position in word or percentage of landing position in area of interest (Holmqvist, 2011)**.**

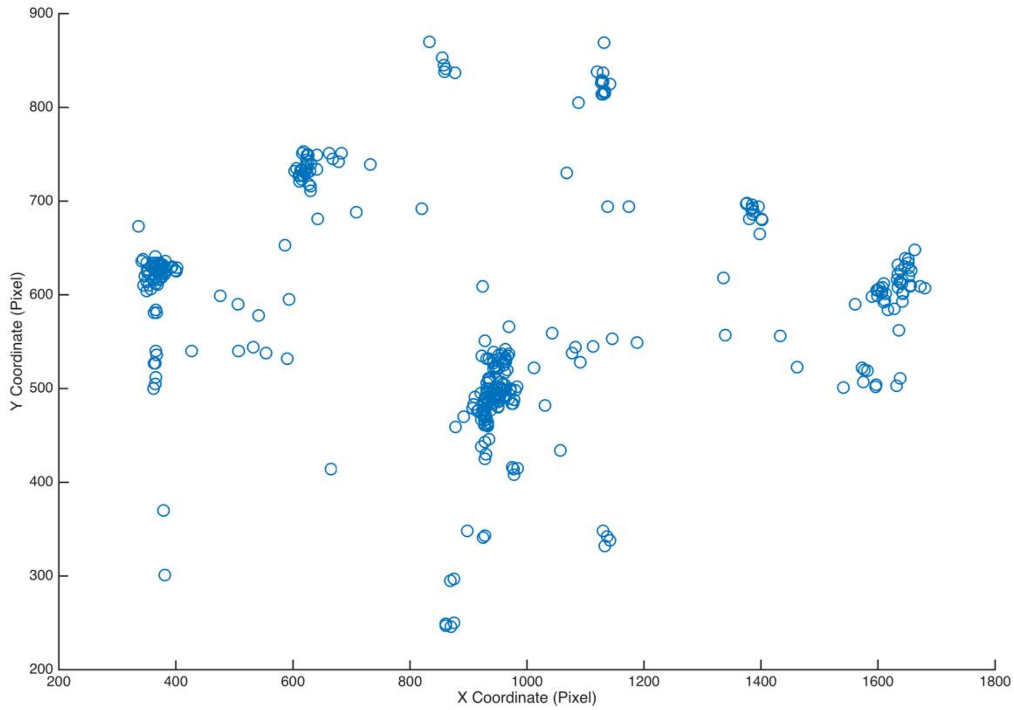


Figure 2: Sample raw eye movement data in Cartesian coordinate system

For more representative information, we can extract fixation information from raw eye movement data by using the selected fixation detection algorithms or we can produce dwell information by using Area of Interests (AOI).

Fixations denote maintaining of the visual gaze on a single location. Value of dwell position is related with selected area of interests. Dwell information is x and y coordinate set in selected area of interest.

4.1.2 Position Dispersion Metrics

In group eye tracking data, the researcher needs to compare multiple eye movement datasets to find effect of group condition. There can be two or more raw eye movement datasets based on the experiment condition. The dataset count depends on the number of group members. In current state of art, there are several dispersion metrics used to find dispersion between two datasets. Our assumption is the dispersion information is highly correlated with joint attention. For instance, in gaze cueing, if the dispersion between two or more data set is low or high compared with the single dataset under same condition and stimuli, we can conclude that members in the group are effected from each other's gaze.

Similarity metrics can be calculated from both raw eye movement data and fixation data. These metrics are standard deviation, variance, range, convex hull area, BCEA (Bivariate contour ellipse area) and Kullback-Leibler Distance (Holmqvist, 2011). Some of the similarity metrics are hard to implement and observe the divergence among different eye movement dataset such as BCEA, whereas some of the similarity metrics can be very intuitive to the researcher. In the following we describe the advantages and disadvantages of each similarity metric.

4.1.2.1 Standard Deviation

Standard deviation quantifies amount of variation or dispersion in data. It is commonly used to measure confidence in statistical conclusions. In eye movement analysis, standard deviation represents the dispersion in raw eye movement positions or fixations for a single set. Standard deviation is defined as;

$$\sigma = \sqrt{\frac{1}{N} * \sum_{i=1}^{N} (x_i - \overline{x})^2}$$

Where N stands for sample size in data set, $\overline{x}$ stands for the mean of the data set.

The problem of standard deviation is that it is very sensitive to outliers and it doesn't detect any cluster formation in raw eye movement data. To overcome this problem, we can first cluster the raw eye movement data or fixations then measure the variation or standard deviation within that cluster.

4.1.2.2 Range

Range represents smallest border points which cover the raw data samples fixations or saccades in eye movement dataset. In a sample eye movement dataset, range method calculates the triangle area which is constructed with the minimum and maximum x and y coordinates. Range is defined as;

$$R_h = max(x) – min(x)$$

$$R_v = max(y) – min(y)$$

R_h and R_v defines a rectangular shaped area. This area is the minimum area which covers all the raw eye movement data in given set. Range is used mostly in human factors research especially in measuring saccadic extent parameter. Range is also used in I-DT fixation extraction algorithm to detect fixations in raw eye movement data set.

4.1.2.3 Convex Hull Area

Convex Hull Area is similar to the Range metric. The difference between range and convex hull area is the shape of convex area which the algorithms compute. In Range, there is a rectangular shape area which defines the minimum border of eye movements in set. In Convex Hull Area, there is a minimal convex area that spans all points in dataset which is depicted in Figure 3.

Figure 3: Gaze Dynamic Metric - Convex Hull Area

Convex Hull Area is an operational definition of dispersion as the smallest convex polygon containing set of data samples. The convex hull area was adapted to eye tracking analysis by Goldberg & Kotval (1999). Some researchers still use convex hulls to analyze the spatial distribution of eye movement data.

4.1.2.4 Circular Area

We have developed a metric to analyze the group effect. The metric depends on the coverage area of subjects' eye movements in the group. Most important thing is defining the geometric shape of the area. In our study, we have 3 subjects in group. We can define a triangle to calculate the group effect which is depicted in Figure 3. The area of triangle can give information about closeness of subjects in group. However, there can be misleading information because of the nature of triangle. As depicted in Figure 3, area of triangle doesn't give correct result always.

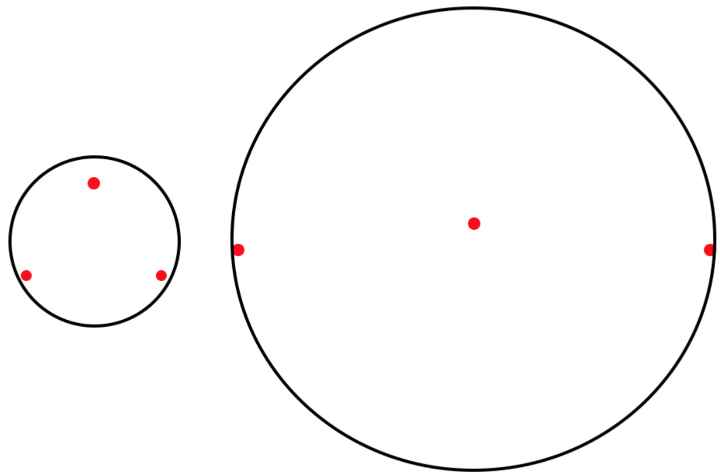


Figure 4: Gaze Dynamic Metric - Circular Area

As an example, the eye movement positions in Figure 3 and Figure 4 are identical. The divergence in first group is lower than the divergence in second group.

If we get the area of triangle in Figure 3 as a gaze dynamic metric, the result will be inefficient. It can be seen that the first group's triangle area is close to second group's triangle area. However, for divergence of eye movements, the first group's divergence is much lower than the second group's divergence.

If we get the area of circle in Figure 5 as a gaze dynamic metric, the result will be more intuitive and correct. The divergence is getting higher or lower with the area of circle. On the other hand, we can observe the groups which have more than 3 members. –This will be impossible for triangle calculation–

4.1.2.5 Bivariate Contour Ellipse Area (BCEA)

BCEA is a dispersion measure which stands for density values of each raw eye movement data sample. In this study to calculate BCEA, multivariate kernel density estimation has been used. Multivariate KDE is a nonparametric estimation of probability density functions of a random variable. Multivariate KDE is generalized form of histograms. Multivariate KDE (Kernel Density Estimation) algorithm is used to detect density values, which uses second order Gaussian kernel (Z. I. Botev, 2010). The goal of the density estimation is to take a finite sample of data and to make inferences about the underlying probability density function in all locations, including the locations where no data points are observed. In Kernel Density Estimation, the contribution of each data point is smoothed out from a single point into a region of space surrounding it. A major issue in using the Multivariate KDE is the need for adjusting the bandwidth value. If the bandwidth value is too high, the data will be over smoothed, if the bandwidth value is too low the data will be spiky. In our algorithm, the bandwidth adjustment is made depending on data characteristics. A method of calculating the bandwidth value is to use the MISE (Mean Integrated Square Error) value (Wand & Jones, 1995). A less complicated means of calculation is provided by the AMISE (Asymptotic Mean Integrated Square Error) value which was developed by Wand & Jones (1995), which we employ in our study.

BCEA has been used to measure fixation dispersion in clinical applications. Moreover, BCEA is used for quantifying inter participant dispersion. We also use BCEA as an input in Kullback-Leibler divergence to measure position dispersion among group members.

As an example, multivariate KDE applied to sample raw eye movement data which is depicted in Figure 5. Multivariate KDE generates a heat map from the raw eye movement data. The dense areas (dense areas have lighter color in this heat map) can be observed from the heat map. Participants look more to the dense areas in heat map.

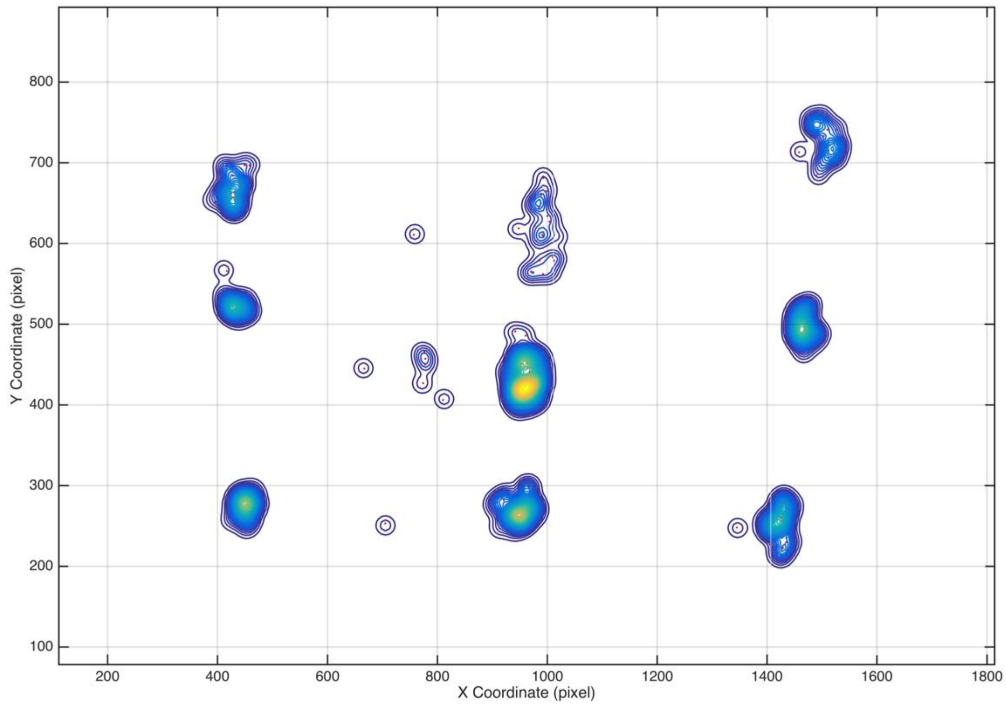


Figure 5: Multivariate KDE of sample raw eye movement data

4.1.3 Scanpath

There are three main eye movement metrics used in eye movement researches. These metrics are saccades, fixations and scanpaths. Scanpath refers the series of fixations and saccades. Scanpaths are mainly used in clinical research and human computer interaction and website usability studies (Soegaard & Friis, 2013).

In eye movement data analysis, there are several ways to define and analyze scanpaths in raw eye movement data. Scanpaths are very useful tool for analyzing eye movement data sequentially. However, the order information can be lost in fixation or saccade data. Because of the nature of the scanpaths, order of the events (saccades and fixations) will be preserved.

One way of quantifying scanpaths is using string sequences. To create string sequences, we first divide the stimulus screen into small pieces (area of interests). After that, we entitle each area of interest a letter in alphabet. Close letters –for instance A and C– should represent close area of interests in coordinate system (Holmqvist, 2011). After naming operation, we can represent the fixation and saccade list in a string sequence like

"AACCCCBBBBB". There are several studies using this technique to compare two eye movement dataset. In this study, we also use this metric to compare group members eye movement events. Another way to define scanpaths is sing vectors. In mathematics, vector structure is in <u, v> form. Where u is the beginning position of vector and v is the ending position. In eye movement data analysis, u is used as beginning of eye movement event and v is used as ending of eye movement event (Holmqvist, 2011).

Another way of quantifying scanpaths is calculating repeating eye movement metrics (fixations and saccades). Regression means that repetitive fixation movements on the same location. During analysis, higher number of fixation regressions on target which has been fixated before indicate that it lacks meaningfulness or visibility (Allport, 1968).

4.2 Methods

4.2.1 Mannan Similarity Index

Mannan similarity index is calculated by subtracting scanpaths in one eye movement dataset from another one. Fixations or raw x and y coordinates can be used for subtraction. Scanpaths give spatial information about the eye movement data. The subtraction is made with close scanpaths (fixation sequence or raw eye movement sequence) in eye movement data.

Mannan similarity metric/index is developed by Mannan, Ruddock, & Wooding, 1995 to measure the distance between each fixation in one eye movement data and its nearest one in the other.

The average linear-distance is defined as D^2, where

$$D^2 = \frac{\sum_{j=1}^{n_2} d_{2j}^2 + \sum_{j=1}^{n_2} d_{2j}^2}{2 * n_1 * n_2 * (a^2 + b^2)}$$

and where n_1 and n_2 are the number of fixations in each scanpath and a and b are the dimensions of the image. d_{1i} is the distance between the i^{th} fixation in the first set and its nearest fixation in the second set, and d_{2j} is the same distance for the j^{th} fixation in the second set. There can be scanpaths having different number of fixations. Therefore, to make the metric more robust, the size of stimulus added to the function. ($a^2 + b^2$). Where "a" stands for the horizontal screen size and "b" stands for the vertical screen size.

Mannan et al. (1995) finds a similarity index to produce an estimate of the absolute degree of similarity. The similarity index is defined as I_s , where,

$$I_s = \left(1 - \frac{D}{D_r}\right) * 100$$

D_r is difference between randomly generated scanpath and average linear distance between two scanpaths (D). D is average linear distance between two scanpaths. I_s gives a value between zero (chance similarity) and 100 (identity). There can be negative values. The negative values indicate that scanpaths are more different than expected. We define expected (D_r) distance by using randomly generated scanpath. The distance between randomly generated scanpaths (D_r) produces the normally distributed similarity that would be expected from chance or uniform scanning. It is observed that for a constant display size $(a^2 + b^2)$, the average random distance gets smaller (D_r) as more fixations are added to the scanpath.

Firstly, the measurement does not take into account the temporal sequence of the scanpath. Fixation locations are compared to whichever fixation is closest, regardless of when it occurred. Despite the fact that in one scanpath the observer starts at the bottom left and works upwards whilst in the other they do the opposite. One way to avoid this problem might be to compute a "serial position" version, where the distance is computed between each fixation and that fixation which occurred in the same serial position in the other scanpath. However, this would be skewed by any small deviations.

4.2.2 Kullback-Leibler Distance

KLD is first used with eye tracking data by (Rajashekar, Cormack, & Bovik, 2004; Nystörm, Novak, & Holmqvist, 2004; Tatler, Baddeley, & Gilchrist, 2005). Rajashekar et al. (2004) used the symmetric KLD to quantify the distance between fixation predictions depend on the stimuli and recorded fixations from humans.

Kullback-Leibler (KL) distance, is the difference between two probability distributions (P and Q). To calculate probability distribution of eye movement data, we have used Multivariate Kernel Density Estimation.

The distribution of P and Q is not symmetric. P stands for true distribution of data or observation and Q represents model or approximation of P. “P||Q” stands for how much distribution of Q similar to distribution of P. In eye movement analysis to make the calculation more symmetrical, we will compute both P||Q and Q ||P.

$$D_{KL}(P \mid\mid Q) = \sum_{i} P(i) log \frac{P(i)}{Q(i)}$$

$$D_{KL}(Q \mid\mid P) = \sum_{i} Q(i) log \frac{Q(i)}{P(i)}$$

After calculating each distance, in this study we take the harmonic distance (Johnson & Sinanovic, 2000). To calculate a resemblance score based on KL Distance, we take the mean of harmonic distance between two probability distribution;

$$Score = mean(\frac{1}{\frac{1}{D_{KL}(P \mid Q)} + \frac{1}{D_{KL}(Q \mid P)}})$$

4.2.3 Scanmatch

Scanmatch is an algorithm to measure the eye movement similarities between two datasets (Cristino, Mathôt, Theeuwes, & Gilchrist, 2010). Algorithm uses scanpath metric as an input. Scanpath is a representation of area of interests in stimuli. The stimuli screen is divided into small grids. Then each grid in stimuli is named using alphabetical letters. (A, B, C …) Each x and y coordinate tuple, stands fora grid letter on stimuli screen. After naming operation, based on gaze data and its grid, a string sequence is created.

In scanmatch algorithm, string sequences produced with eye movement datasets are compared by Needleman-Wunsch algorithm (Needleman & Wunsch, 1970). In bioinformatics, Needleman-Wunsch algorithm is used to compare the DNA sequences. However, it is novel approach in eye movement data analysis. In Needleman-Wunsh algorighm, string sequences (produced with scanpaths), are subtracted from each other to find the distance to each other. According to the study of Cristino et al. (2010), scanmatch method is strong compared with other comparision methods when the order of the fixations is significant. However, if the order is not important, this method may not be as useful.

4.2.4 Jaccard Index

Jaccard index or jaccard similarity coefficient measures the similarity or divergence of two datasets. It is computed with the size of intersection of datasets dividing by the size of union of datasets (Levandowsky & Winter, 1971). Jaccard index is based on set theory where;

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

$J(A,B)$ stands for the jaccard index. If dataset "A" and "B" are empty, the jaccard index or coefficient is computed as one. On the other hand, the jaccard distance is calculated by subtracting one from the jaccard index.

Jaccard distance is complementary to the jaccard index. It is calculated by subtracting 1 from Jaccard index where;

$$J_D = 1 - J(A,B) = 1 - \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$$

J_D stands for jaccard distance. In our analysis, we have used a special form of jaccard distance. (Generalized jaccard index) In GET paradigm, there are 3 different gaze points. However, in standard jaccard distance calculation, there should be two different datasets. To calculate jaccard distance for multiple dataset, (In our analysis, the dataset count is 3) a formula is computed described below;

$$J_G = \frac{\sum_i \min(x_i, y_i)}{\sum_i \max(x_i, y_i)}$$

J_G stands for the generalized form of jaccard index. x_i and y_i are data coordinates in each eye movement dataset. In our analysis we have used generalized form of jaccard index.

4.2.5 Jensen-Shannon Divergence

Jensen-Shannon divergence is a measure for similarity between two probability distributions. Jensen-Shannon divergence is generally used in bioinformatics and social sciences. The divergence is based on Kullback-Leibler distance. Kullback-Leibler distance has asymmetrical feature. This means that the distance between A and B dataset is not the same as the distance between B and A. However, the Jensen-Shannon Divergence is symmetrical.

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M)$$

$$M = \frac{1}{2}(P + Q)$$

In KLD, we can compute merely $D(P||Q)$ or $D(Q||P)$. But in Jensen-Shannon divergence, we can compute a single value which stands for the distance from P to Q and Q to P. To make the Jensen-Shannon symmetrical, first we take the mean of the KLD distances of each distribution (P and Q). Then we take the mean distribution called.

In our thesis, there are three group members. There are three eye movement dataset for the group. We should calculate the divergence for three eye movement dataset. Therefore, there will be three eye movement data points (x and y coordinates) for each line or timestamp. We apply three of the measures to the group eye movement data. These three measures are convex hull, circular area and jaccard index. The other gaze divergence algorithms can be applied only two different eye movement dataset. However, convex hull, circular area and jaccard index algorithms can be applied to multiple (more than one) eye movement dataset. Therefore, we have chosen these three algorithms for the GET analysis.

CHAPTER 5

5 APPLICATION OF THE THREE MEASURES TO BALLOON GAME DATASET

5.1 Participant Procedure

We developed a multiple eye tracking infrastructure (GET – Group Eye Tracking) and a multiplayer game “balloon game” running on this infrastructure to analyze the joint attention in group members. In our thesis, there are 66 participants. 37 of them are male and 29 of them are female. The mean of the participant’s age is 22,7. We have defined group member count as 3. Therefore, there are 22 groups in the experiment.

The experiments were conducted at Cogs Lab METU. We have used EyeTribe eye trackers to collect the eye movement data. The resolution of the client screen is 1920 x 780. As we stated earlier, we have developed an infrastructure to connect the participant machines to each other. With this infrastructure, the participants can see each other’s eye movements on their screen (gaze awareness).

There are two **game conditions**. The game conditions are gaze cueing and non-gaze cueing. In gaze cueing game condition, participants can see eye movements of each other while they are playing the game. In non-gaze cueing game condition, participants cannot see each other’s eye movements. We aim at finding a difference between gaze cueing and non-gaze cueing in terms of the group conditions. There are three **group conditions.** These group conditions are all enemies(AE), single gaze necessary(SGN) and three gaze necessary (3GN). The group descriptions will be explained in the next section.

Before we start the experiment, we trained the participants for all the group conditons in the experiment. In the experiment, participants aim at popping the balloons on the screen to get a score. According to the group descriptions, we have three training set for each group condition (AE, SGN, 3GN). In each training set, there are six targets (balloons). For AE’s training, each participant has a unique color. Therefore, there are six balloons with three different colors (there are two balloons for each color). For SGN and 3GN, there are six balloons(targets) having the same color. Because in SGN and 3GN conditions, there is not a balloon which has different color than others. After training the participants, we display an instruction text to the participants and then we start the experiment. In the following, I will explain the application which is running on the GET (Group Eye Tracking) infrastructure.

5.2 A Sample GET Application: Balloon Game

We have developed two applications for the GET paradigm. These applications are balloon game and lottery game[1]. In this thesis, we report the analysis of the balloon game eye movement dataset.

In balloon game, participants try to pop the balloon with needles. The genre of the balloon game is whac a mole. In the game, there are balloons instead of moles and needle instead of hammer. Participants control the needles with their eyes and try to pop the balloons on the screen with these needles which is depicted in Figure 2. Participants play the balloon game either in teams or individually.

There are three different group conditions. In 3 gaze necessary group condition (3GN) all the participants are ally and try to pop the balloons together. To pop a balloon all three group members must hit the same balloon at the same time. They have same needle and balloon color. In all enemies group condition (AE), each participant has its own color and tries to blow up the balloon whose color is related to their color. When participant blows up the balloon, he/she gets score. In single gaze necessary group condition (SGN), all participants have same color and tries to blow up the balloons independently.

We set balloons' appearance duration randomly. If a participant hit a balloon, the balloon disappears immediately. However, if none of the participants hit the balloon, balloon will be displayed on the screen for the duration which we set randomly.

We set balloons' appearance locations pseudo-randomly. Pseudo-randomly means that before the experiment, we produce random locations for balloons. After starting the experiment, we display balloons according to the randomly generated locations. We produce random positions before the experiment to prevent appearance of the balloons next to the participants' eye cursors at the beginning of the experiment.

We use term "session" for the duration between appearance and disappearance of the balloon. For analyzing the eye movement dataset, we divide the dataset into hit sessions and miss sessions. In hit session, participants should hit the target according to the group condition. In miss session, participants do not hit any balloon.

---

[1] The application is a lottery game with ten paired lottery choices played by individuals or three person groups in which they have to choose between safe (low-risk) and risky (high-risk) options. In half of the questions in the group setting, each person can see where other members of the group are looking at in the real time.

Domain of application is gambling with monetary outcomes, where the participants make risky choices under various experimental conditions.

Currently, we are still collecting the eye movement data for this application. We put this application as an example of the GET.

In AE group condition, we give each participant a color. In this group condition, participants have colorful needles and colorful balloon targets. Each participant tries to pop his/her own balloon according to the its color. However, in other conditions, we give each participant same color of balloon and needle (grey color).

There are also 2 game conditions. First one is gaze cueing, second one is non-gaze cueing. In gaze cueing, participants can see the eye movements of each other on their screen. In non-gaze cueing, participants cannot see the eye movements of each other. In both condition, participants sit and do experiments together as a group in the same room.

Game is playing on server side and we are using “TeamViewer” software to share the game environment scene to the clients. Participants must connect to the server with TeamViewer. Group members take into account the target location on the screen. The same targets appear in the same time and location in each participants’ screen. The group members know the target location and the group condition. Also in gaze cueing, the group members can see eye movements of each other on their screen. In gaze cueing, the group members aware of the other group members’ awareness.

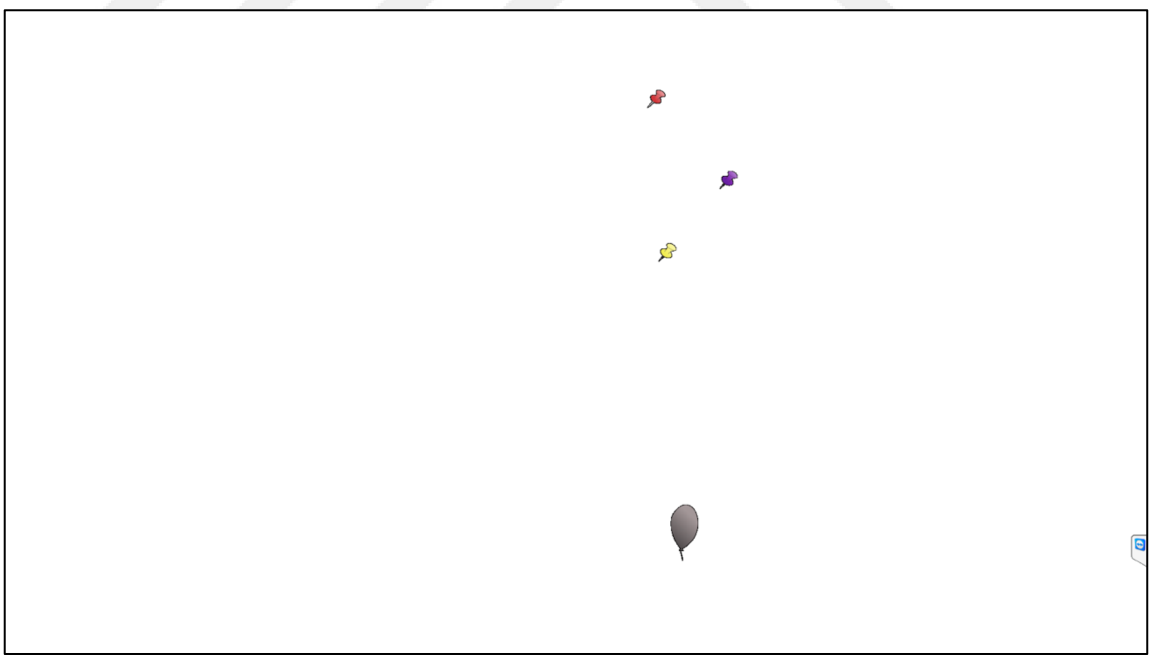


Figure 6: Balloon Game Environment

5.3 Balloon Game Dataset Analysis

For position similarity measure, we will use raw x and y coordinates. The extracted fixation information or dwell time can be used for analyzing data. We use three different scoring algorithms to measure the group effect in balloon game experiment dataset. The scoring algorithms are; circular area, convex hull area and jaccard index. In this study, we benchmark these 3 scoring algorithms. In the following figures, we depicted the flow charts of the three scoring algorithms' mechanism.

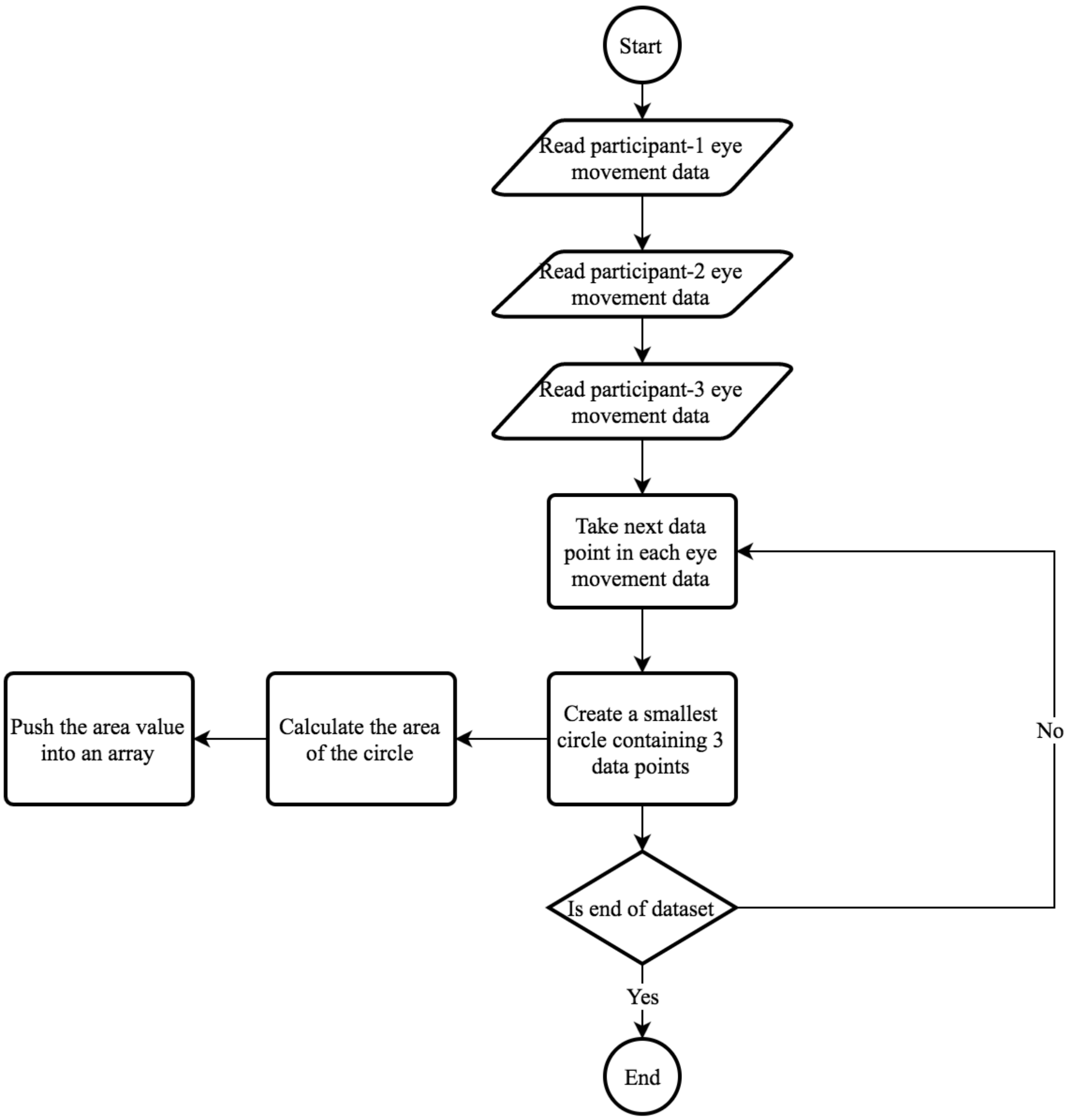


Figure 7: Circular Area Score Calculation

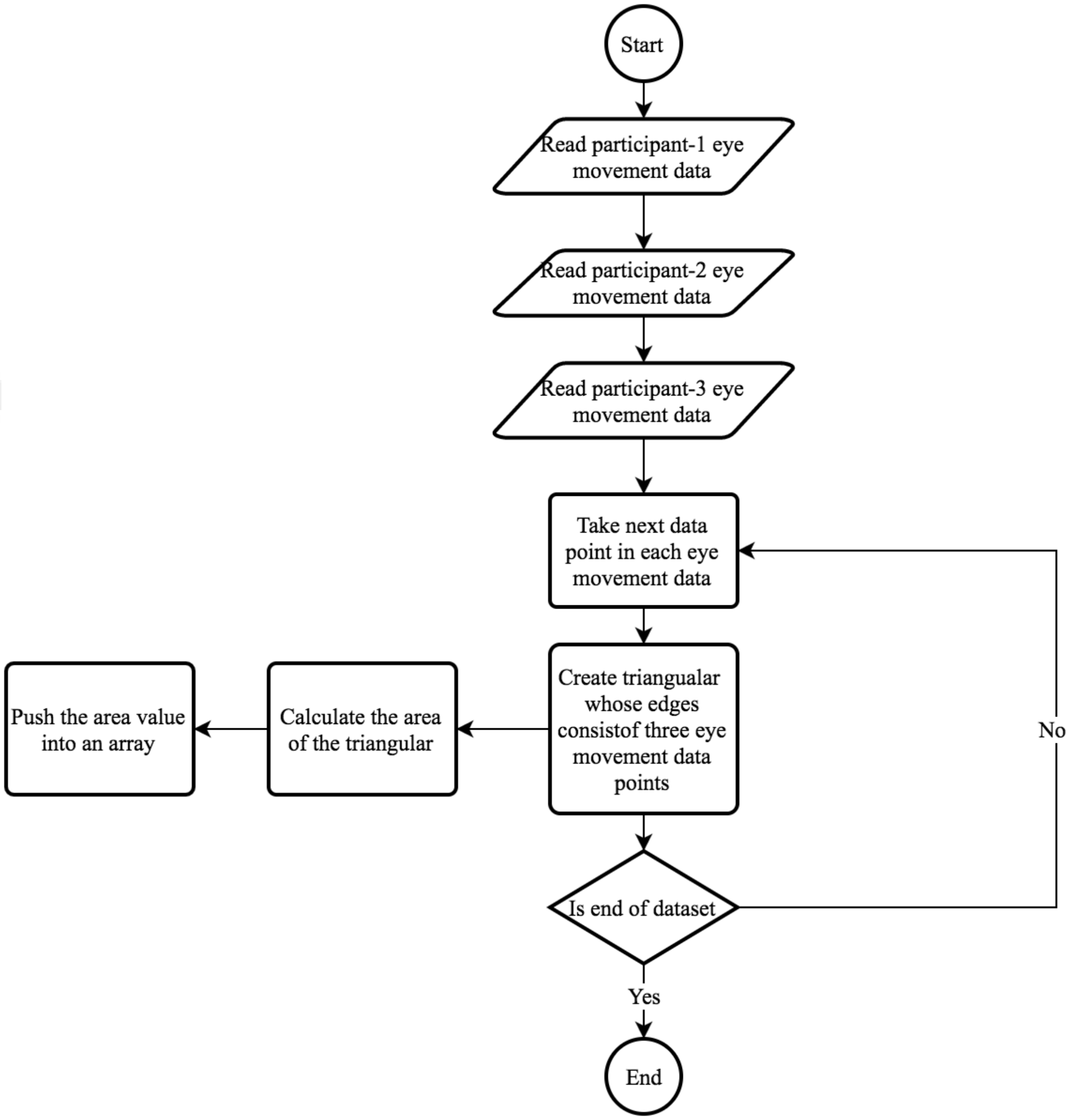


Figure 8: Convex Hull Area Score Calculation

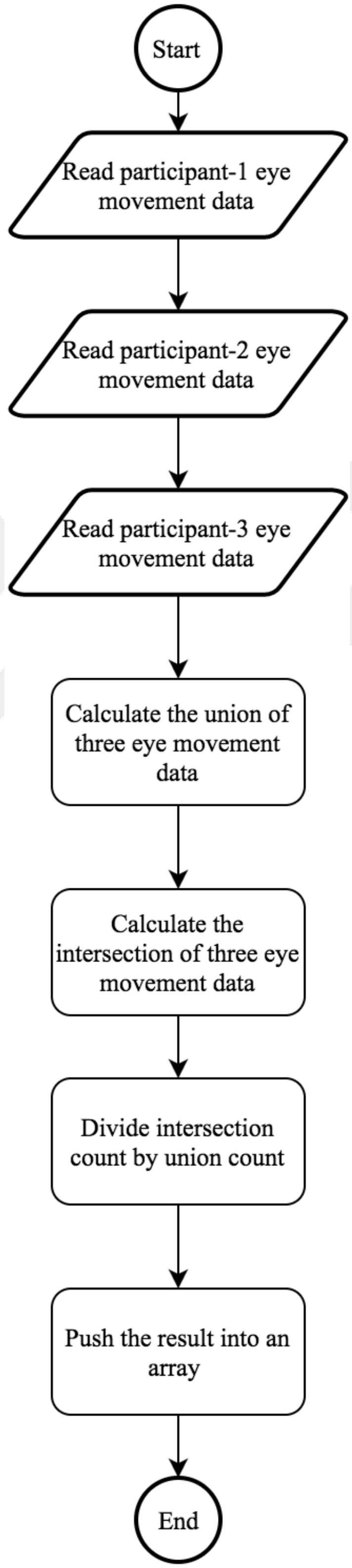


Figure 9: Jaccard Index Calculation

In first step of analysis, we divide the eye movement dataset in three part. These parts represent the game conditions. The gaming conditions are 3 gaze necessary (3GN), single gaze necessary (SGN) and all enemies (AE). The goal of the game is the same across all the game conditions. The goal is popping the balloon on the screen. In 3GN group condition, to pop the balloon three gaze must locate at the same balloon and at the same time. In SGN group condition, to pop the balloon single participant's gaze is necessary to pop the balloon. In AE group condition, the participants are enemy to each other. Each participant has own balloon color. If participant pops a balloon having different color, he/she gets penalty score.

Moreover, we took into account the hit and the miss sessions in the experiment. In the experiment a session starts from the time where the balloon appears on the screen and ends when the balloon disappears. After the session ends, another balloon appears on the screen in a different location. Hitting and missing conditions depend on the group conditions. For instance, in 3 gaze necessary (3GN) group condition, to hit a balloon, all the three participant must hit the same balloon in the same time. In all enemies (AE) and single gaze necessary (SGN) group conditions, one participant's hit is enough for popping the balloon. In all of the group conditions, if the balloon is not popped, we name its session as miss. In second step of analysis, we divide the three gaming condition's dataset into two as hit and miss sessions. After dividing the dataset into six, we calculated the scores for each part. (group condition and hit/miss condition)

The scores are calculated row by row. This means that we produce the scores for each row in dataset. We will plot the scores and try to see the trends and patterns. After plotting, we make additional statistical calculation which will be explained in further sections.

CHAPTER 6

6 RESULTS

There are 3 group conditions and 3 scoring algorithms used for benchmarking. In this thesis, we have two research questions. First one is how **different group conditions** (3Gaze necessary, single gaze necessary, all enemies) effect the gaze pattern of the people. Second one is how **gaze animation** effect the gaze pattern. In group eye tracking data analysis, we will use the term “gaze cueing” for gaze animation and “non-gaze cueing” term for non-gaze animation. We conducted the experiments with gaze cueing and non-gaze cueing. In gaze cueing and non-gaze cueing, the data is collected under 3 gaze necessary(3G), single gaze necessary(SG) and all enemies(AE) group conditions.

In this thesis, we have used three scores for group eye tracking data analysis. These scores are convex hull area, circular area and jaccard index. Scores stand for the divergence of the participants’ gaze location. The divergence is calculated for each data point in eye movement data. Group consists of three members. Therefore, we calculate the divergence of the three eye movement data points.

In circular area score, we calculate the area of the minimum circle which covers three participants’ eye movement data points. In convex hull area score, we calculate the area of a triangular whose edges consist of three participants’ eye movement data points. However, in jaccard index, we divide the intersection count of the three participants’ eye movement dataset by union count of them. Jaccard index calculates the divergence with respect to the set theory in mathematics.

The question still remains after calculating the divergence of the gaze. What does the gaze divergence represent in group eye tracking data? Each score stands for the closeness of the group members. During the experiment, the group members are getting closer or moving away from each other in time. We calculate these three scores to observe the divergence and convergence of the eye movements in time. In the end, we aim at finding an effect of the group conditions (3G, SGN or AE) and gaze cuing to the gaze patterns of group members.

As we stated above, we have 22 groups and each group has three members. In the following section, we report the analysis of the overall response time of the group members in different group conditions. Then we will look at the convergence scores of the group members.

6.1 Response Time

We measure the average response time of the participants in different group conditions which is depicted in Table 1. In balloon game, the response time is the time between the balloon appearance and its popping.

Table 1: Response Times of the Participants

	Gaze Cueing	Non-Gaze Cueing
3 Gaze Necessary (3GN)	340.54 ms	286.53 ms
Single Gaze Necessary (SGN)	161.03 ms	149.89 ms
All Enemies	268.48 ms	194.82 ms

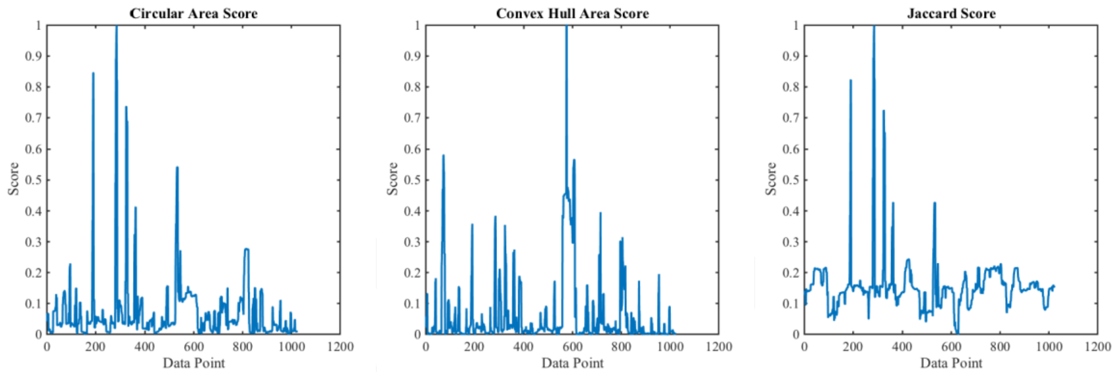
As we can see from the Table 1, in gaze cueing, participants respond slower than the non-gaze cueing. Also within the gaze cueing and the non-gaze cueing, 3 gaze necessary condition (3GN) has the highest response time when compared with other group conditions. In gaze cueing, the participants can see eye movements of each other. Gaze cueing might have resulted in slowing down the participant’s gaze thus divergence. Because it is basically a visual clutter. Therefore, in gaze cueing, the response time is higher than the non-gaze cueing. In the following section, the raw scores of one group are depicted.

6.2 Raw Scores

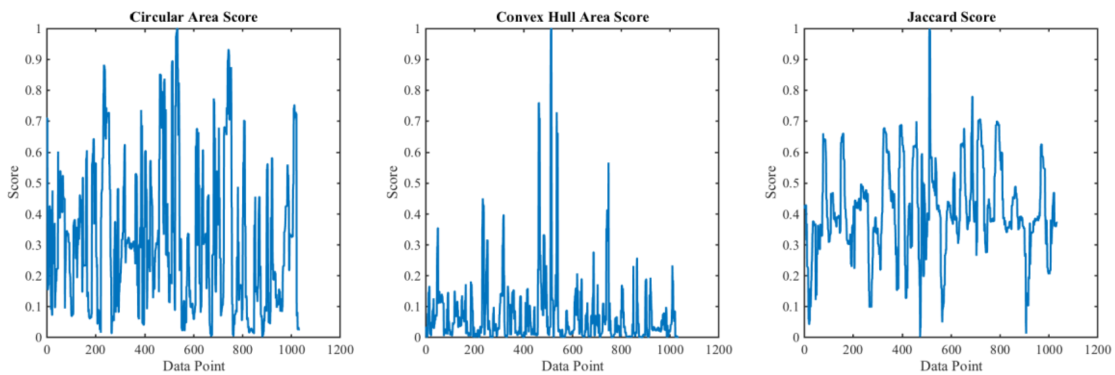
We applied three scores (circular area, convex hull area, jaccard index) to the group eye movement datasets. After that, we normalized the scores. Therefore, the scores diverge from zero to one. We analyzed the results in terms of both group conditions and gaze-cueing. In the following section, we depicted one group's raw scores.

6.2.1 Gaze Cueing

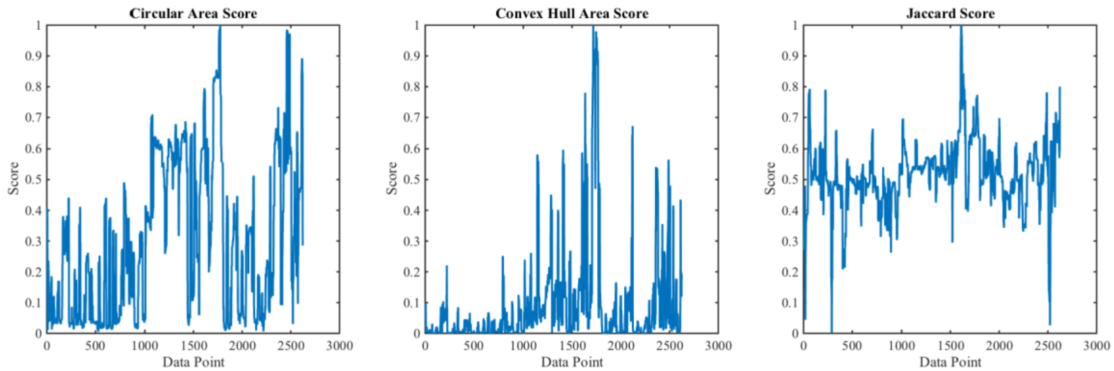
6.2.1.1 3 Gaze Group Condition



6.2.1.2 Single Gaze Group Condition

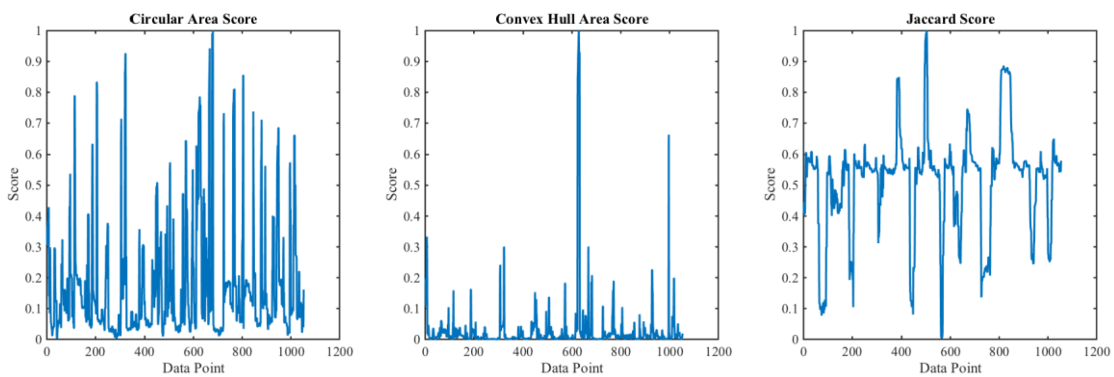


6.2.1.3 All Enemies Group Condition

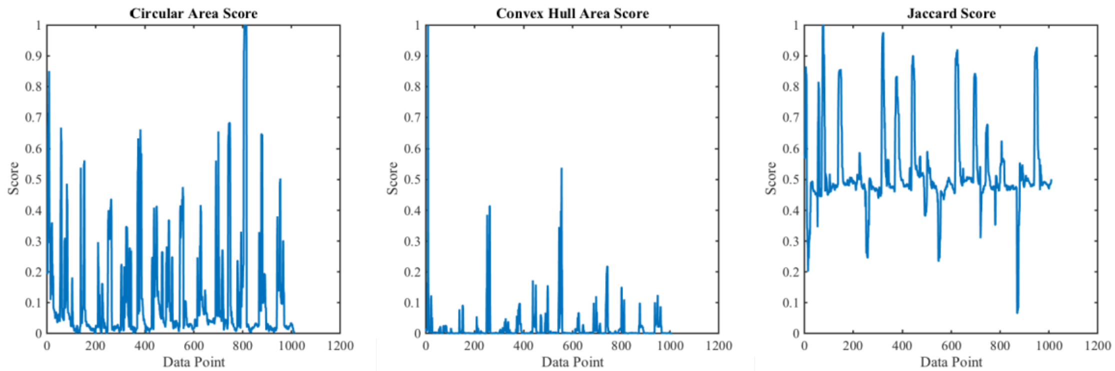


6.2.2 Non-Gaze Cueing

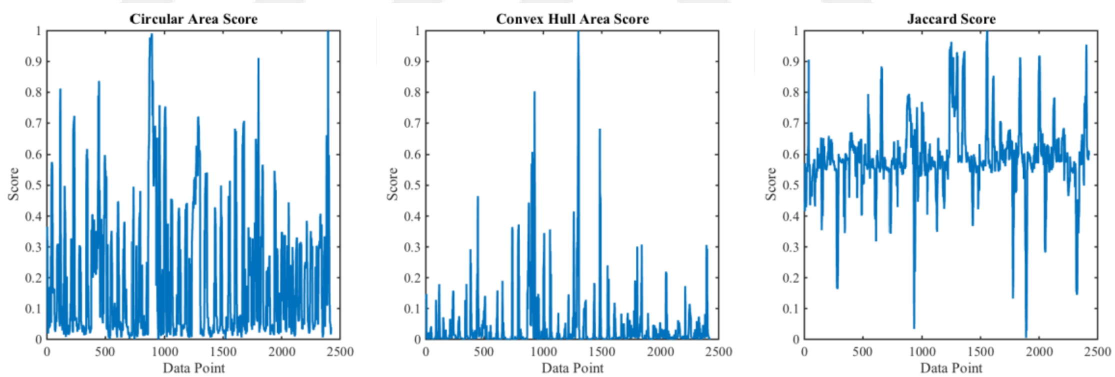
6.2.2.1 3 Gaze Group Condition



6.2.2.2 Single Gaze Group Condition



6.2.2.3 All Enemies Group Condition



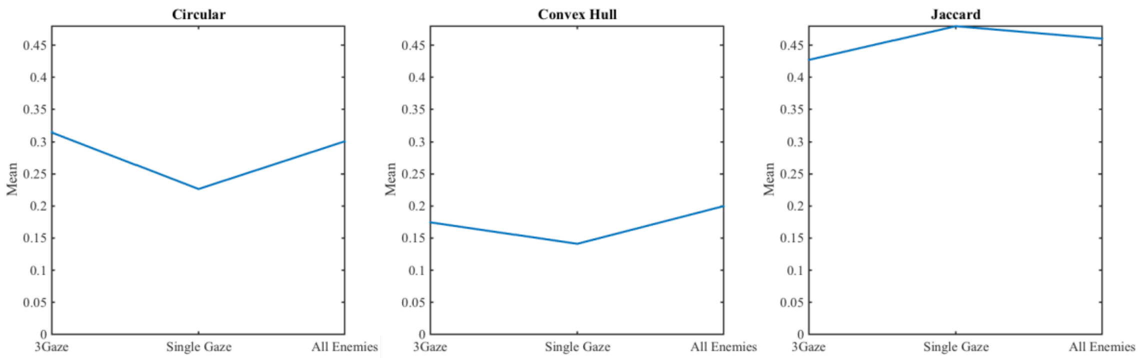
6.2.3 Results of the Raw Scores

As we can see from the figures which were depicted above, the raw scores do not give sufficient and descriptive information about the group eye movement data. There is no trend observed in raw scores. Because the raw score data is too noisy. To reduce the noise, we bin(bucket) the raw score data and analyze the bin counts. Also, we calculate the mean and the median of the scores to observe a significant change in different group conditions.

In section 6.2 and 6.3, we analyze the mean and the median of the scores respectively. In section 6.5 and 6.6, we will report the analysis of the binned raw score data.

6.3 Mean of the Scores

6.3.1 Gaze Cueing



3Gaze	Single Gaze	All Enemies
0.31	0.23	0.30

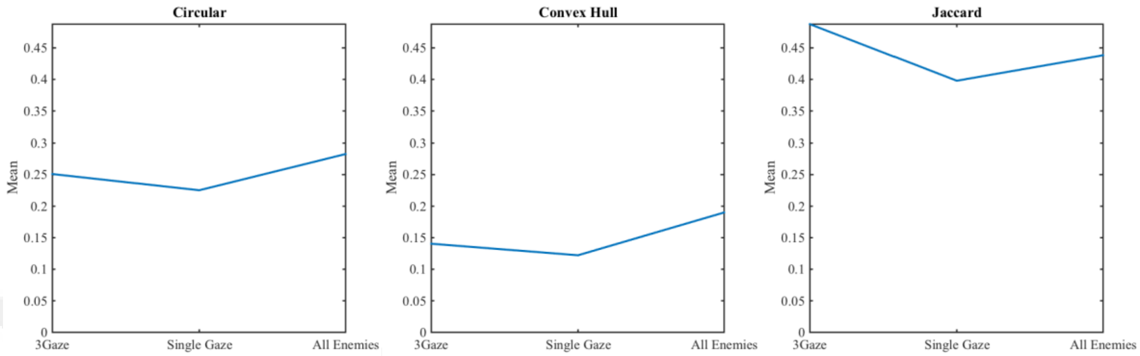
3Gaze	Single Gaze	All Enemies
0.17	0.14	0.20

3Gaze	Single Gaze	All Enemies
0.43	0.48	0.46

Based on the mean of the circular area score, we observe 3GN (3 gaze necessary) and AE (all enemies) group conditions have higher scores compared by SGN (single gaze necessary) group condition. This means that in 3GN and AE group conditions, the gaze divergence is more than the SGN group condition.

According to the mean of the convex hull area score, we observe a similar trend as in circular area score. However, there is a conflict in the mean of the jaccard index. According to the jaccard index, SGN group condition has the highest score compared by other two group conditions. This means that the divergence in SGN is more than the 3GN and AE.

6.3.2 Non-Gaze Cueing



3Gaze	Single Gaze	All Enemies
0.25	0.23	0.28

3Gaze	Single Gaze	All Enemies
0.14	0.12	0.19

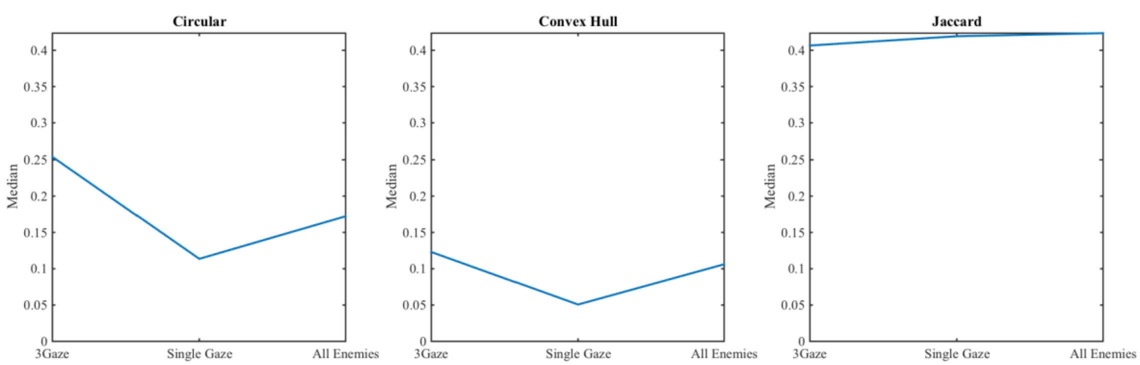
3Gaze	Single Gaze	All Enemies
0.49	0.40	0.44

Based on the mean of the circular area score, we observe AE (all enemies) group condition has higher score compared by SGN (single gaze necessary) and 3GN (3 gaze necessary) group conditions. This means that in AE group condition, the gaze divergence is more than 3GN and SGN group conditions.

According to the mean of the convex hull area score, we observe similar trend like circular area score. However, there is a conflict in jaccard index with respect to the mean scores. According to the jaccard index, 3GN group condition has the highest score compared by other two group conditions. Whereas circular and convex hull area scores, jaccard score indicates that in 3GN, the divergence is more than SGN and AE group conditions.

6.4 Median of the Scores

6.4.1 Gaze Cueing



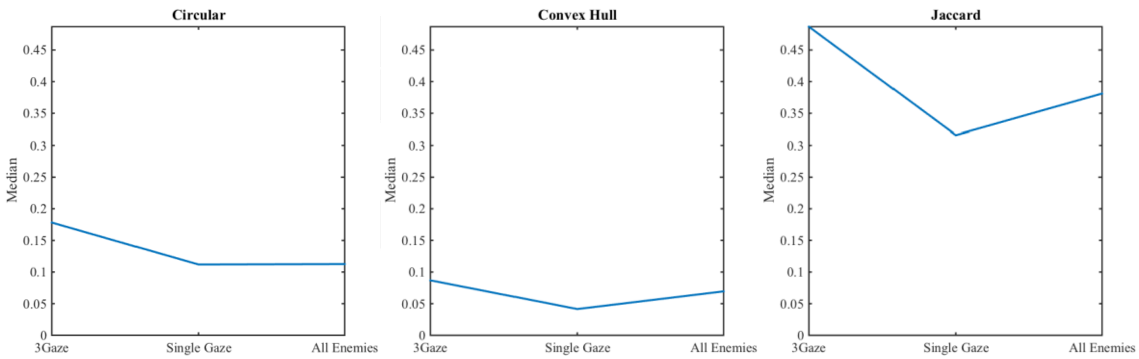
3Gaze	Single Gaze	All Enemies
0.25	0.11	0.17

3Gaze	Single Gaze	All Enemies
0.12	0.05	0.10

3Gaze	Single Gaze	All Enemies
0.41	0.42	0.47

Based on the median of the circular and convex hull area scores, we can observe similar trend as mean of these scores. However, in jaccard index, we observe opposite trend in which AE (all enemies) group condition has the highest score and 3GN (3 gaze necessary) group condition has the lowest score. This means that in 3GN group condition, the divergence is less than SGN (single gaze necessary) and AE group conditions and the most divergence is in AE group condition.

6.4.2 Non-Gaze Cueing



3Gaze	Single Gaze	All Enemies
0.18	0.11	0.11

3Gaze	Single Gaze	All Enemies
0.09	0.04	0.07

3Gaze	Single Gaze	All Enemies
0.48	0.32	0.38

Based on the median of the circular area score, we observe 3GN (3 gaze necessary) group condition has higher score compared by AE (all enemies) and SGN (single gaze necessary) group conditions. This means that in 3GN group condition, the gaze divergence is more than SGN and AE.

According to the median of the convex hull area score and jaccard score, we observe similar trend like circular area score.

6.5 Results of the Mean and the Median of the Raw Scores

Mean and median of the scores give relatively more descriptive results than the raw scores. Whereas the raw scores, in mean and the median of the scores, we can find the least and the most converged group conditions. To make more detailed analysis, we bin the raw score data and analyzed the bin counts.

We analyzed the binned scores in terms of hit, miss and overall (hit + miss). We created 320 bins starting from 0.1 to 0.9. We defined the interval between two bins as 0,0025 (The bins are 0.1, 0.10125, 0.10250, 0.10375, 0.10500, …,0.8975, 0.9). After that, we calculated the number of scores for each bin. We normalized the scores. Therefore, the scores diverge from zero to one. We didn’t include zero and one scores. Because it is nearly impossible that the participants look at the same point at the same time (divergence is zero) and it is also impossible that the participant diverges infinitively (the maximum divergence is limited with the resolution of the screen [1920 x 1080]).

Score is a measurement of divergence of participant gaze location. When the score is low, the divergence of the group participants’ eye movement is low. The convex hull area score is computed by calculating the triangular area generated with three eye movement data points (Three eye movement dataset is compared because in our experiment there are three participants in one group.). The circular area score is computed by calculating the circular area generated with three eye movement data. The jaccard index is computed by dividing the intersection of three eye movement dataset by union of them. Consequently, lower threshold (0.1) stands for the lower divergence and higher threshold (0.9) stands for the higher divergence. For instance, if the circular area score is 0.1, in that moment, the circular area constructed with the group members’ eye movements is 0.1. If they move away from each other, the area of the circle gets higher.

As a result, high bin count for low score (0.1) stands for the low divergence. High bin count for low score (0.1) means that in the score dataset, there are more low scores (0.1) when compared the other.

In the analysis, we first calculate the mean and median of the scores under different group conditions. After mean and median calculation to make further and descriptive statistical analysis, we define several bin values for scores.

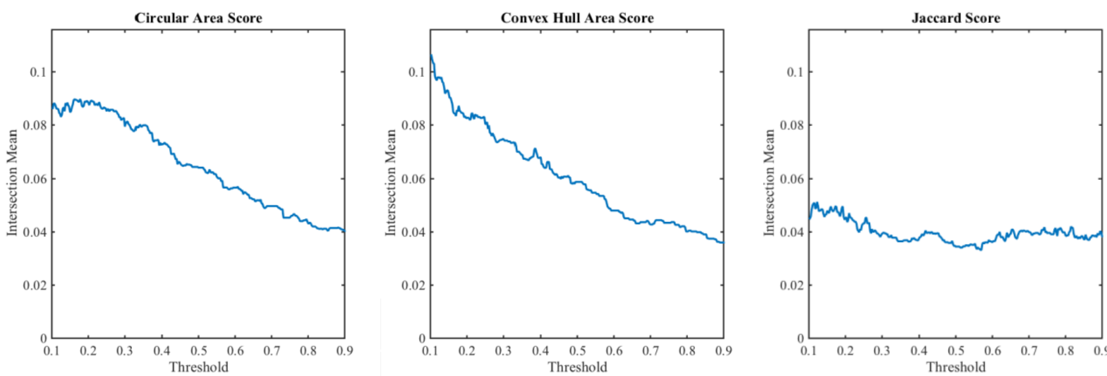
We divide the bin count (for instance 0.325) by the number of scores in the dataset. In the following analysis report, we will use the term **intersection mean** for this division result and **threshold** for the bin value.

6.6 Findings with the Presence of Gaze-Cueing

In this section, we report the analysis of the results of gaze-cueing. In gaze-cueing, participants can see each other’s eye movements on their screen during the experiment.

6.6.1 3GN Group Condition

6.6.1.1 Hit Session

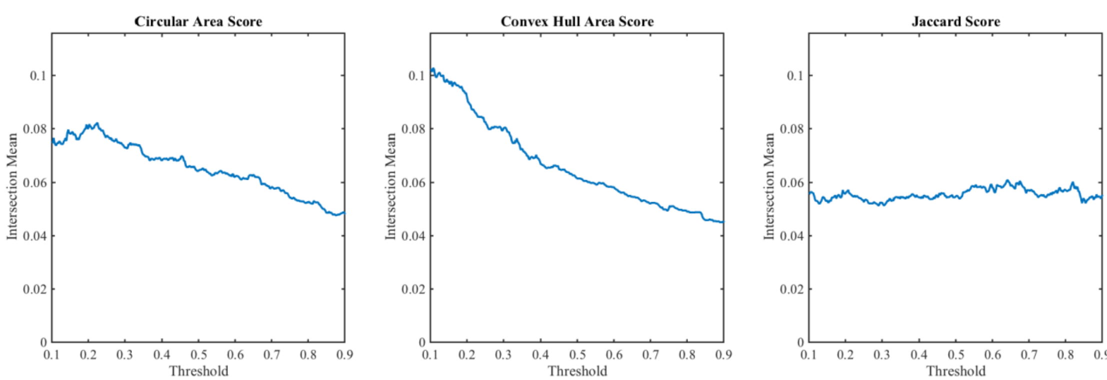


0.1	0.3	0.5	0.7	0.9
0.087	0.077	0.061	0.047	0.041

0.1	0.3	0.5	0.7	0.9
0.103	0.067	0.054	0.044	0.037

0.1	0.3	0.5	0.7	0.9
0.045	0.037	0.036	0.039	0.039

6.6.1.2 Miss Session

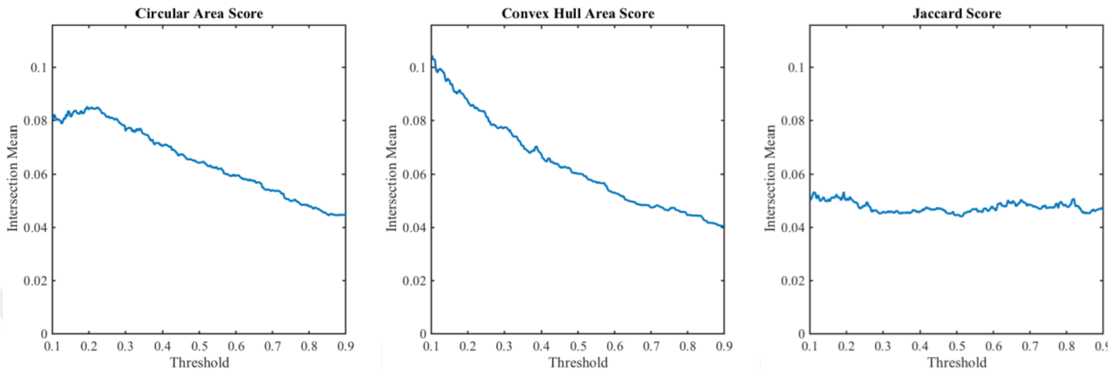


0.1	0.3	0.5	0.7	0.9
0.077	0.069	0.063	0.057	0.057

0.1	0.3	0.5	0.7	0.9
0.102	0.070	0.059	0.050	0.044

0.1	0.3	0.5	0.7	0.9
0.055	0.054	0.060	0.056	0.055

6.6.1.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.082	0.073	0.073	0.052	0.044

0.1	0.3	0.5	0.7	0.9
0.103	0.069	0.056	0.047	0.040

0.1	0.3	0.5	0.7	0.9
0.050	0.045	0.048	0.047	0.047

For circular area score, in hit session we can see that there is more intersection mean in lower threshold (0.1) and less intersection mean in higher threshold (0.9) compared by miss session. This means that the participants are more converged during hit session in 3GN group condition.

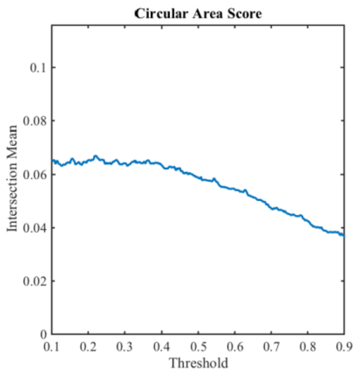
For convex hull area score, there is a similar founding as in circular area score. According to convex hull area score, participants are more converged during hit session in 3GN group condition.

According to the jaccard index, we can say that in hit session, there is less intersection mean in lower threshold (0.1) and also less intersection mean in higher threshold (0.9) compared with miss session. Less intersection mean in higher threshold (0.9) indicates that the participants tend to stay close each other. But, we cannot see a significant effect as in circular and convex hull area scores.

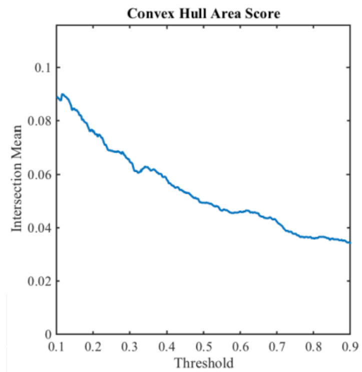
The difference is more observable in circular and convex hull area scores. According to the circular and convex area scores, the intersection mean difference between the lowest threshold (0.1) and highest threshold (0.9) is higher in hit session than in miss session.

6.6.2 SGN Group Condition

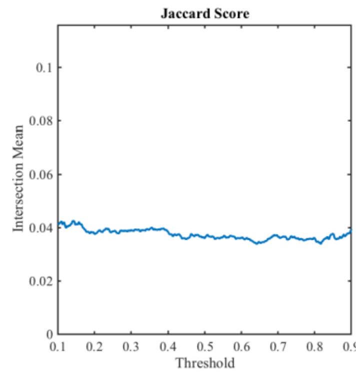
6.6.2.1 Hit Session



0.1	0.3	0.5	0.7	0.9
0.067	0.064	0.058	0.047	0.037

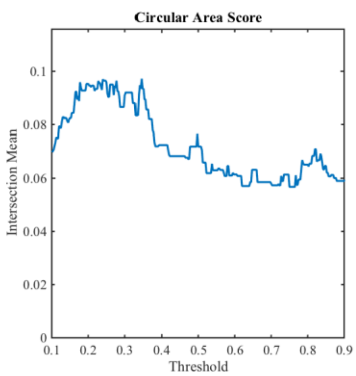


0.1	0.3	0.5	0.7	0.9
0.088	0.088	0.047	0.039	0.035

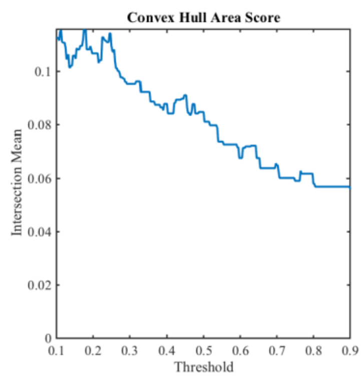


0.1	0.3	0.5	0.7	0.9
0.043	0.040	0.038	0.036	0.038

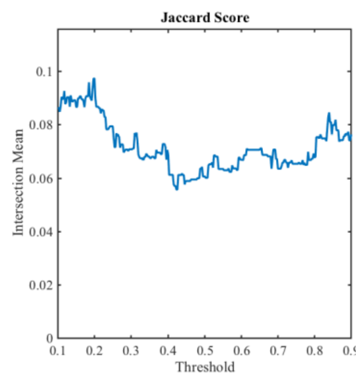
6.6.2.2 Miss Session



0.1	0.3	0.5	0.7	0.9
0.073	0.083	0.066	0.057	0.058

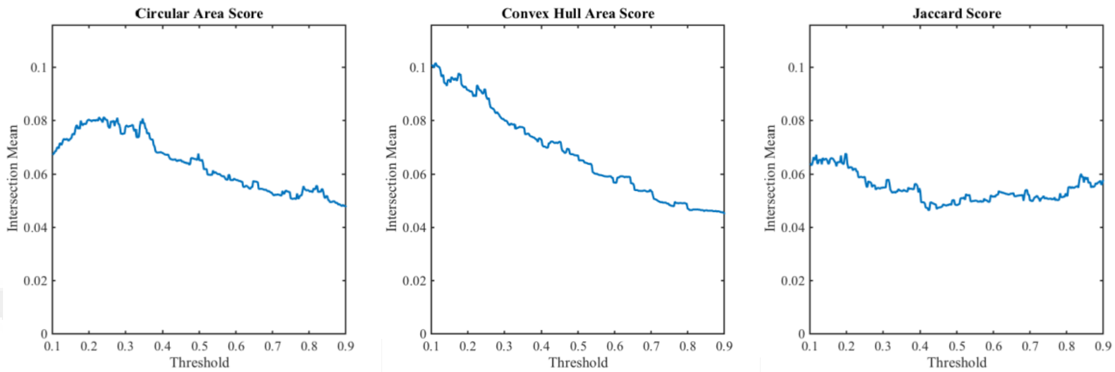


0.1	0.3	0.5	0.7	0.9
0.121	0.083	0.071	0.058	0.057



0.1	0.3	0.5	0.7	0.9
0.057	0.074	0.068	0.071	0.071

6.6.2.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.070	0.074	0.062	0.052	0.047

0.1	0.3	0.5	0.7	0.9
0.105	0.072	0.059	0.049	0.046

0.1	0.3	0.5	0.7	0.9
0.063	0.057	0.053	0.053	0.055

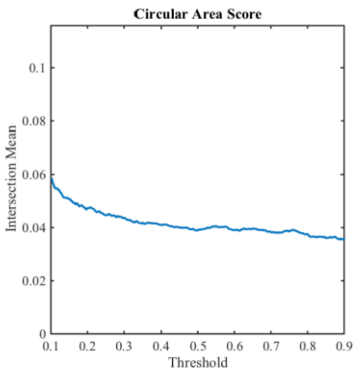
Based on the circular area score, we can say that there is less intersection mean in hit session’s lower (0.1) and higher (0.9) threshold values compared by miss session’s. According to the intersection means in hit and miss sessions, we can say that the participants stay away from each other not too much in hit session when we compare it with miss session. Because hit session’s higher threshold (0.9) intersection mean is lower than miss session’s. Higher threshold’s (0.9) intersection mean stands for the divergence of group members.

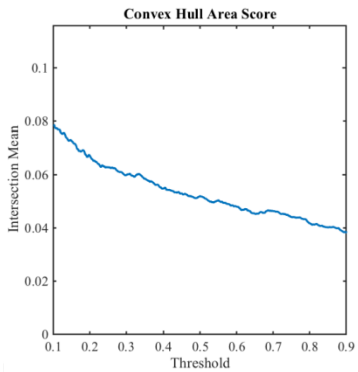
However, in 3GN group condition, hit session’s lower threshold (0.1) intersection mean is higher than miss session’s. Lower threshold’s (0.1) intersection mean stands for the grouping of members. We can conclude that in 3GN group condition, members are converged more in hit session compared with SGN group condition.

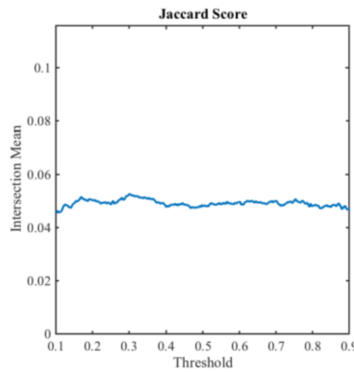
For convex hull area score and jaccard index, there is a similar trend as in circular area score.

6.6.3 AE Group Condition

6.6.3.1 Hit Session





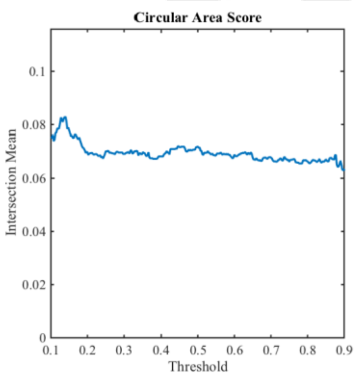


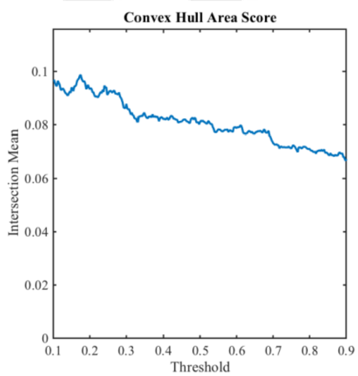
0.1	0.3	0.5	0.7	0.9
0.057	0.042	0.041	0.041	0.036

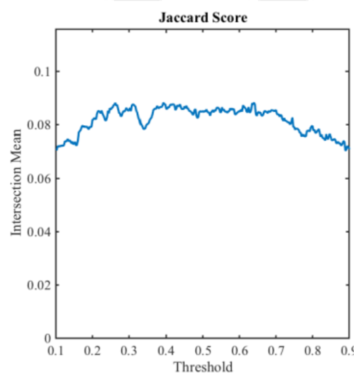
0.1	0.3	0.5	0.7	0.9
0.077	0.057	0.050	0.045	0.038

0.1	0.3	0.5	0.7	0.9
0.046	0.049	0.050	0.050	0.047

6.6.3.2 Miss Session





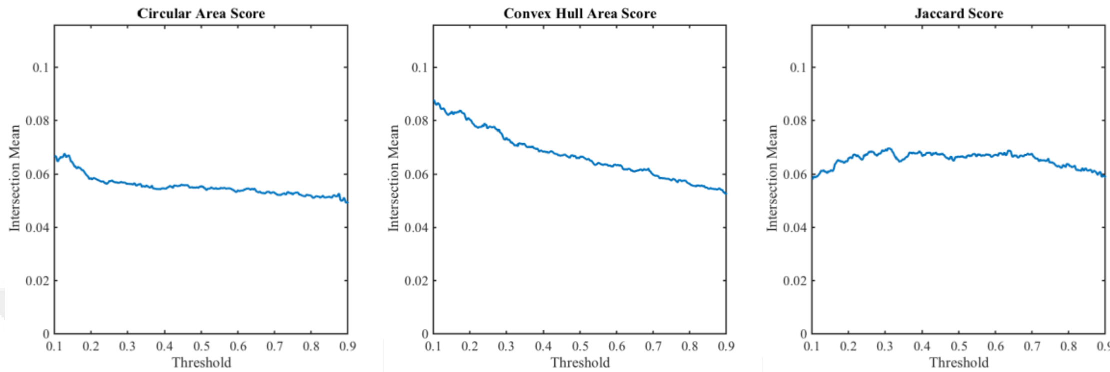


0.1	0.3	0.5	0.7	0.9
0.077	0.066	0.067	0.067	0.062

0.1	0.3	0.5	0.7	0.9
0.096	0.082	0.076	0.069	0.065

0.1	0.3	0.5	0.7	0.9
0.069	0.085	0.085	0.078	0.069

6.6.3.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.067	0.054	0.054	0.052	0.049

0.1	0.3	0.5	0.7	0.9
0.086	0.070	0.063	0.057	0.052

0.1	0.3	0.5	0.7	0.9
0.059	0.059	0.067	0.064	0.058

Based on the circular area score, we can say that there is less intersection mean in hit session's lower (0.1) and higher (0.9) thresholds compared by miss session's. According to the intersection means in hit and miss sessions, we can say that the participants stay away from each other not too much in hit session. Because the intersection mean is lower in hit session's higher threshold (0.9) values.

For convex hull area score and jaccard score, there is a similar trend like circular area score.

Overall Findings;

Intersection mean in hit session's lowest threshold value (0.1) can be seen below;

Table 2: Intersection mean in hit session's lowest threshold value

	Circular Area	Convex Hull Area	Jaccard Index
3GN	0.087	0.103	0.45
SGN	0.067	0.088	0.43
AE	0.057	0.077	0.46

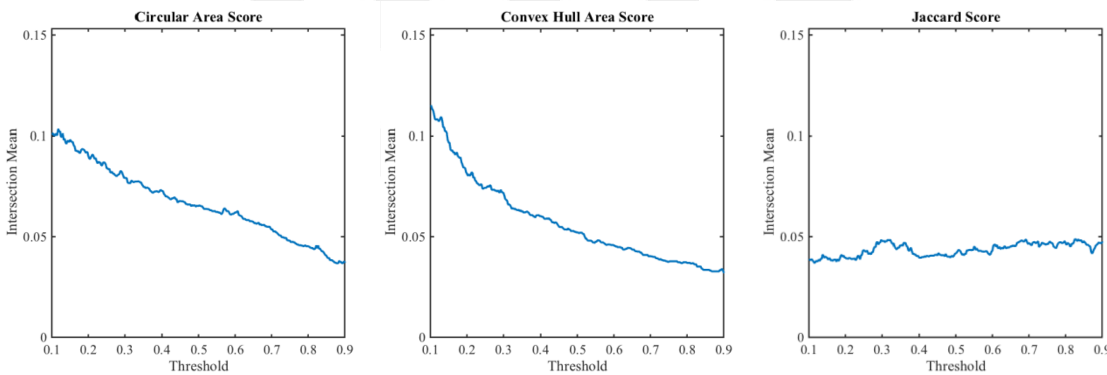
There is a decreasing trend in both circular area and convex hull area scores with respect to the group conditions. This means that the participants tend to stay close each other most in 3GN group condition. But we cannot observe the same trend in jaccard index.

6.7 Findings with the Presence of Non-Gaze Cueing

In this section we analyze the results of non-gaze cueing. In non-gaze cueing, participants do not see each other's eye movements on their screen. We have three group conditions for non-gaze cueing. These group conditions are 3 gaze necessary (3GN), single gaze necessary (SGN) and all enemies (AE). Below, we will report the analysis of non-gaze cueing group conditions.

6.7.1 3GN Group Condition

6.7.1.1 Hit Session

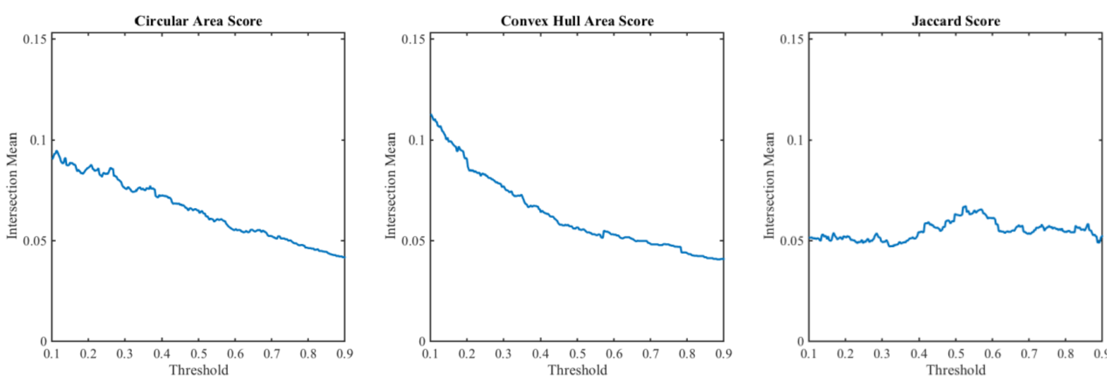


0.1	0.3	0.5	0.7	0.9
0.093	0.067	0.058	0.046	0.046

0.1	0.3	0.5	0.7	0.9
0.104	0.056	0.056	0.036	0.030

0.1	0.3	0.5	0.7	0.9
0.035	0.040	0.040	0.042	0.043

6.7.1.2 Miss Session

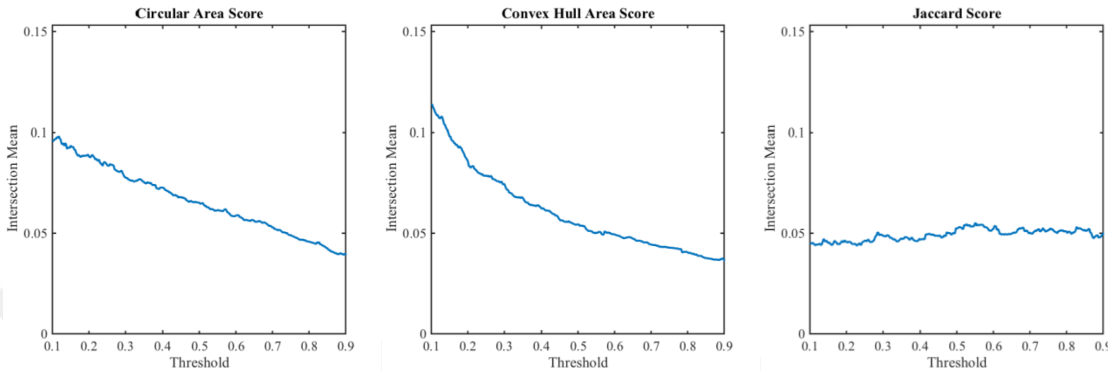


0.1	0.3	0.5	0.7	0.9
0.085	0.070	0.056	0.047	0.038

0.1	0.3	0.5	0.7	0.9
0.103	0.062	0.049	0.044	0.038

0.1	0.3	0.5	0.7	0.9
0.048	0.047	0.060	0.052	0.052

6.7.1.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.089	0.068	0.057	0.047	0.035

0.1	0.3	0.5	0.7	0.9
0.104	0.059	0.046	0.040	0.034

0.1	0.3	0.5	0.7	0.9
0.041	0.043	0.051	0.047	0.046

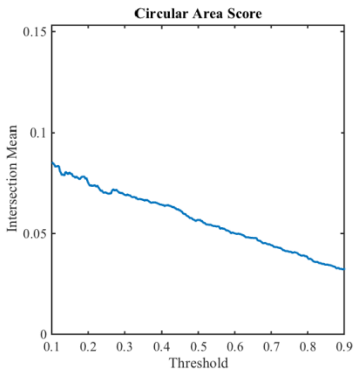
For circular area score, we can say that there is more intersection mean in hit session’s lower threshold (0.1) compared by miss session’s. However, there is more intersection mean in hit session’s higher threshold (0.9) compared by miss session’s. Based on the intersection means in hit and miss sessions, we can say that the participants stay away from each other not too much in hit session because of the hit session’s lower threshold (0.1).

For convex area score, we can say that there is less intersection mean in hit session’s lower threshold (0.1) compared by miss session’s. However, there is also less intersection mean in hit session’s higher threshold (0.9) compared by miss session’s. Based on the intersection means in hit and miss sessions, we can say that the participants stay away from each other not too much in hit session because of the hit session’s higher threshold (0.9).

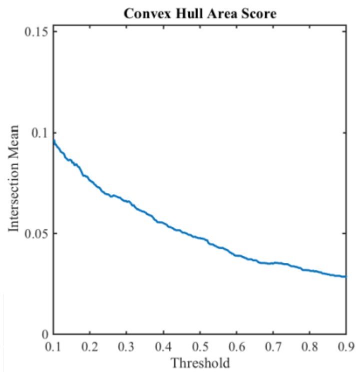
For jaccard index, we can say that there is more intersection mean in hit session’s lower threshold (0.1) compared by miss session’s. However, there is also more intersection mean in hit session’s higher threshold (0.9) compared by miss session’s. Based on the intersection means in hit and miss sessions, there is no trend such as circular area score.

6.7.2 SGN Group Condition

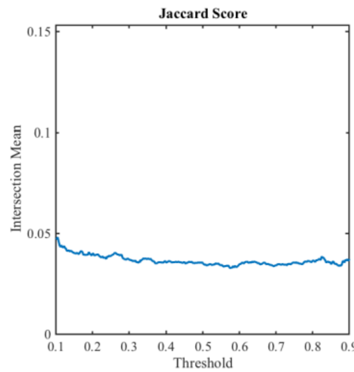
6.7.2.1 Hit Session



0.1	0.3	0.5	0.7	0.9
0.078	0.060	0.049	0.039	0.030

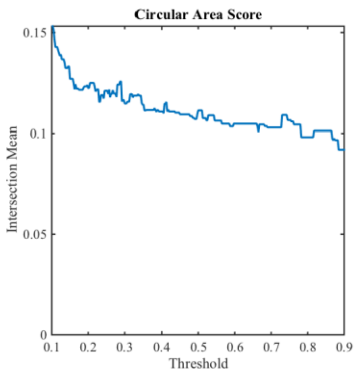


0.1	0.3	0.5	0.7	0.9
0.087	0.053	0.040	0.032	0.026

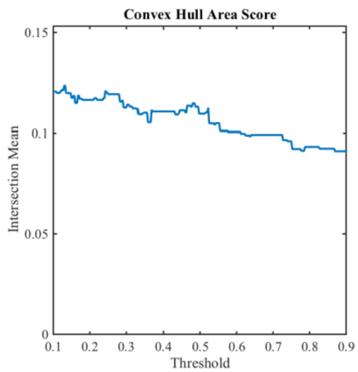


0.1	0.3	0.5	0.7	0.9
0.044	0.033	0.032	0.032	0.034

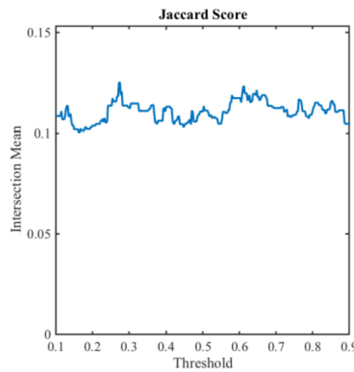
6.7.2.2 Miss Session



0.1	0.3	0.5	0.7	0.9
0.140	0.103	0.098	0.101	0.085

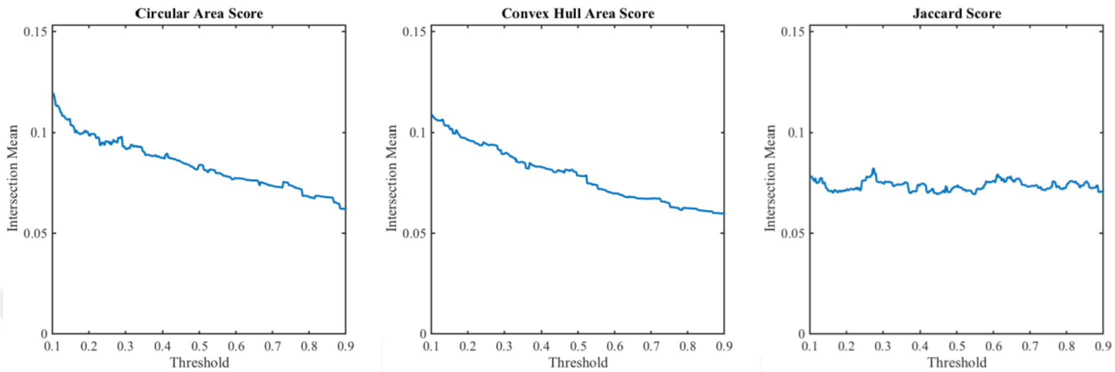


0.1	0.3	0.5	0.7	0.9
0.111	0.102	0.097	0.097	0.084



0.1	0.3	0.5	0.7	0.9
0.100	0.097	0.099	0.100	0.096

6.7.2.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.109	0.082	0.074	0.070	0.057

0.1	0.3	0.5	0.7	0.9
0.099	0.078	0.068	0.061	0.055

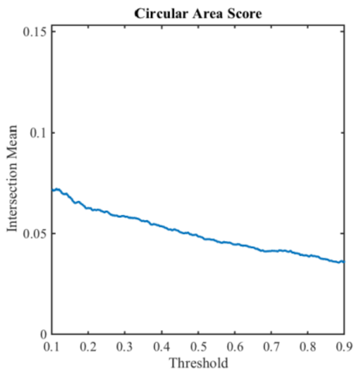
0.1	0.3	0.5	0.7	0.9
0.072	0.065	0.065	0.066	0.065

For circular area score, we can say that there is less intersection mean in hit session's lower threshold (0.1) compared by miss session's. However, there is also lower intersection mean in hit session's higher threshold (0.9) compared by miss session's. Based on the intersection means in hit and miss sessions, we can say that the participants stay away from each other not too much in hit session because of the hit session's higher threshold (0.9).

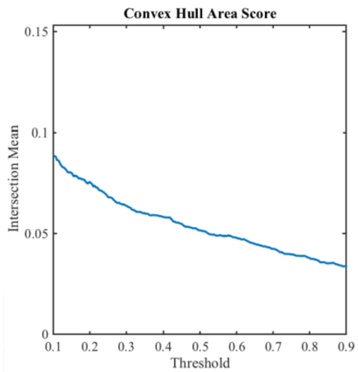
For convex hull area score and jaccard score, there is a similar trend like circular area score.

6.7.3 AE Group Condition

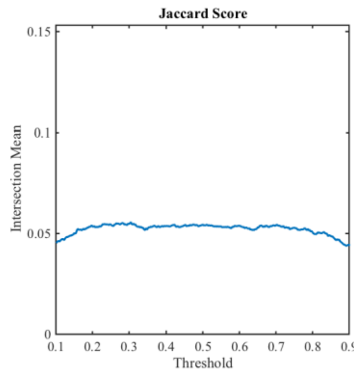
6.7.3.1 Hit Session



0.1	0.3	0.5	0.7	0.9
0.066	0.050	0.042	0.038	0.033

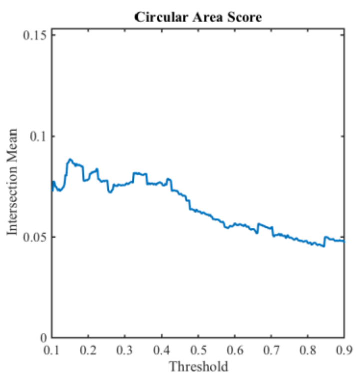


0.1	0.3	0.5	0.7	0.9
0.082	0.055	0.045	0.037	0.031

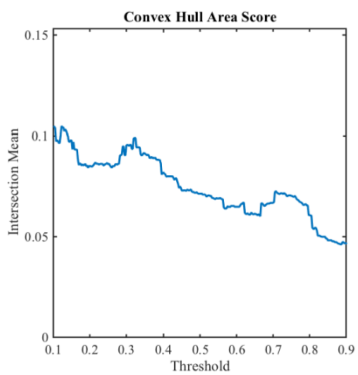


0.1	0.3	0.5	0.7	0.9
0.043	0.049	0.049	0.049	0.041

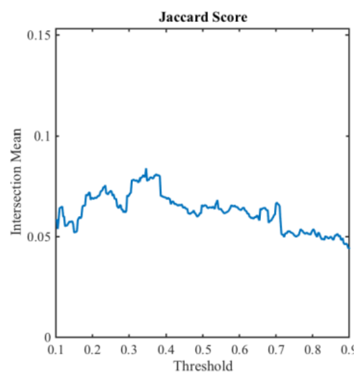
6.7.3.2 Miss Session



0.1	0.3	0.5	0.7	0.9
0.072	0.071	0.054	0.047	0.044

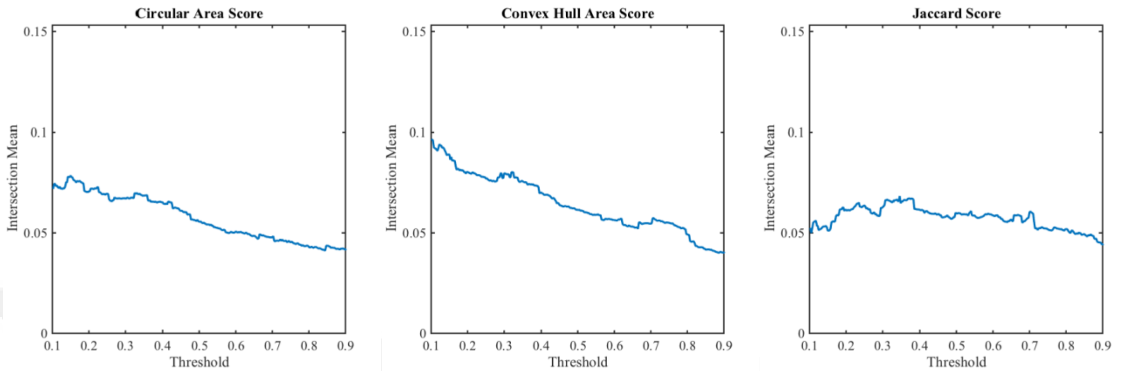


0.1	0.3	0.5	0.7	0.9
0.096	0.082	0.064	0.066	0.043



0.1	0.3	0.5	0.7	0.9
0.043	0.058	0.058	0.049	0.041

6.7.3.3 Overall Session



0.1	0.3	0.5	0.7	0.9
0.069	0.060	0.048	0.042	0.039

0.1	0.3	0.5	0.7	0.9
0.089	0.069	0.055	0.051	0.037

0.1	0.3	0.5	0.7	0.9
0.046	0.062	0.054	0.049	0.041

Overall Findings;

Intersection mean in hit session’s lowest threshold value (0.1) can be seen below;

Table 3: Intersection mean in hit session's lowest threshold value

	Circular Area	Convex Hull Area	Jaccard Index
3GN	0.093	0.104	0.35
SGN	0.078	0.087	0.44
AE	0.066	0.082	0.43

There is a decreasing trend in both circular area and convex hull area scores. This means that the participants tend to behave close each other most in 3GN group condition. But we cannot observe the same trend in jaccard index.

6.8 Findings with the Model Fitting

In section 6.1 and 6.2, we have analyzed the effect of different group conditions (3GN, SGN, AE). We have found different gaze patterns for different group conditions as stated above. This analysis is made for our first research question.

The second research question is about analyzing the effect of gaze awareness. There are two eye movement dataset groups. First one is collected under the condition where participants can see each other’s gaze (gaze-cueing). Second one is collected under the condition where participants cannot see each other’s gaze (non-gaze-cueing).

For comparing gaze-cueing and non-gaze-cueing, we compute four regression models. These regression models are first, second, third, fourth order functions respectively which are depicted below;

First order polynomial function: $f(x) = p_1x$

Second order polynomial function: $f(x) = p_1x + p_2x^2$

Third order polynomial function: $f(x) = p_1x + p_2x^2 + p_3x^3$

Fourth order polynomial function: $f(x) = p_1x + p_2x^2 + p_3x^3\ p_4x^4$

We fit each group condition's eye movement data to each regression model. After model fitting, we calculate the residuals of each model then we take the root mean square of the computed residuals. We aim to estimate the polynomial feature of gaze-cueing depending on the root mean square of residuals. Below there are three tables calculated with three scores (jaccard index, circular area, convex hull area).

Table 4: Jaccard Distance – Overall Condition

Order	Non Gaze-Cueing	Gaze-Cueing
1st	0.0026	0.0029
2nd	0.0020	0.0016
3rd	0.0019	0.0014
4th	0.0017	0.0014

Table 5: Circular Area – Overall Condition

Order	Non Gaze-Cueing	Gaze-Cueing
1st	0.0018	0.0026
2nd	0.0015	0.0024
3rd	0.0012	0.0016
4th	0.0012	0.0012

Table 6: Convex Hull Area – Overall Condition

Order	Non Gaze-Cueing	Gaze-Cueing
1st	0.0035	0.0029
2nd	0.0020	0.0011
3rd	0.0016	0.0010
4th	0.0015	0.0009

We see that for both gaze-cueing and non-gaze-cueing, the mean square of residuals decreasing as we increase the order of polynomial function. However, for gaze-cueing the decreasing rate is higher than non-gaze-cueing. We observe that non gaze-cueing's eye movement data fits to first polynomial order better than gaze-cueing's eye movement data except the convex hull area score. Because the mean square of the residuals is lower in non-gaze-cueing than gaze-cueing. Therefore, we can say that non-gaze-cueing scores are more linear than gaze-cueing scores.

We found that seeing the other participants eye movement cursor can lead the polynomial decrease in scores. In non-gaze-cueing, participants cannot see each other's gaze. The decrease in non-gaze-cueing is more linear than gaze-cueing. We can conclude gaze awareness may have caused the polynomial characteristic of gaze-cueing gaze condition.

CHAPTER 7

7 DISCUSSION AND FUTURE WORK

7.1 Discussion

In this thesis, we analyzed two effects in group eye tracking. The first is the effect of the group condition in comparison to the condition in which individuals perform the eye tracking tasks. The second one is the effect of the awareness of the gaze location of other participants (as represented by a gaze cue on the screen). We further defined three group conditions and two gaze conditions in a game where the participants popped up a balloon by gaze. The group conditions were 3 gaze necessary (3GN), single gaze necessary (SGN), all enemies (AE). As stated in section 5, in 3GN group condition, participants must move coordinated to hit the target. In SGN group condition, participants have same color. They should tend to close to each other. Because the target which they aim, is the same. However, in 3GN group condition, all three participants should hit the same balloon for scoring. In SGN condition one hit enough for scoring. In AG condition, the participants are enemy each other. They have own color. If a participant hits a balloon with a different color, the participant gets a penalty (negative score). Therefore, the participants should be away from each other. The gaze conditions were the presence of a gaze-cueing and the absence of a gaze cueing (i.e., non-gaze-cueing). The experiment was conducted for 3GN with gaze-cueing, SGN with gaze-cueing, AE with gaze-cueing, 3GN with non-gaze-cueing, SGN with non-gaze-cueing, AE with non-gaze-cueing conditions.

In the current eye movement literature, there isn't a single metric to quantify the divergence of multiple eye movement datasets (more than two eye movement datasets). The current metrics in the literature can be applied only to two eye movement datasets. Therefore, for more than two eye movement datasets, there won't be a single metric to quantify the divergence. For instance, if there are three eye movement datasets, there will be three different divergence scores (1-2, 1-3, 2-3). The goal of this thesis is to develop a single metric to quantify the divergence in group eye movement data.

We analyzed the eye movement data collected on different groups by comparing alternative group eye tracking measures. We investigated three gaze-divergence metrics. These metrics are circular area, convex hull area and jaccard index. The circular area metric is computed by calculating the area of the minimal circle that spans all points in dataset. The convex hull area metric is computed by calculating the area of the minimal convex shape that spans all points in dataset. The jaccard index is computed by dividing the size of the intersection of different datasets by the size of the union of them.

From the overall results, we can conclude that circular and convex hull area metrics give better results than jaccard index. There is a similarity between the circular and convex hull area scores. Circular area score gives better results than the convex hull area score. For convex hull area, we calculate the convex area constituted by multiple gaze locations. As we have seen in section 4.1.2.4, the convex area may be low when the dispersion among the group participant is high. Because of this problem, convex hull area metric may give worse results than circular area metric. For jaccard index, we cannot observe significant variance between high and low scores. However, in convex hull area and circular area scores, we have observed significant differences between the high and low scores as expected. Moreover, the correlation between the convex hull area score and circular area score is higher than the jaccard index. However, in model fitting's residual comparison, we have observed that jaccard index gives similar results (trend) when compared with convex hull and circular area scores.

There are two gaze condition. These gaze conditions are, gaze cueing and non-gaze cueing. Each gaze condition has three group conditions. These group conditions are 3GN, SGN and AE. In gaze cueing group conditions, we have ordered the amount of groupings (calculated by the three scores) in each group condition. According to the scores, the group members are more converged in 3GN than SGN than AE. In non-gaze-animation's group conditions, we order the amount of groupings in each group condition. We observed similar order as in gaze-animation. According to the scores, the group members are more converged in 3GN than SGN than AE. Based on the group condition definitions, we expected a similar order.

We have also compared the gaze cueing's and non-gaze cueing's overall scores. To calculate the overall scores, we took the mean of the scores corresponding to each group condition. We have developed four different regression models and each model has different polynomial characteristic. When we fit the eye movement data to our regression models, we observed a similar trend(decreasing) for gaze-animation and non-gaze-animation. However, we observed that the gaze-animation's trend is more polynomial than non-gaze-animation's. We observed an effect of gaze awareness on group members.

7.2 Future Work

In this thesis, we have analyzed the gaze awareness' and different group conditions' effect on group members. We have focused on one input modality (gaze). However, as stated above, there are different modalities to analyze (haptic, speech, body gesture etc.). As a future work, we will combine several modalities (speech + gaze, gaze + body gesture) and analyze the effect of the different modalities on different group conditions as future work.

We have done the experiments with three member groups. If we had crowded groups gaze cueing is more useful because we would be able to observe dynamically moving gaze locations. Feature research should address extending the number of group participants. Further research direction will be extending the group member number to 5 and 7 to observe the member count effect in groups.

References

Acartürk, C., Freitas, J., Fal, M., & Dias, M. S. (2015). Elderly speech-gaze interaction: State of the art and challenges for interaction design. *Universal Access in Human-Computer Interaction. Access to Today's Technologies* (pp. 3-12). Los Angeles: Springer.

Allport, G. (1968). The historical background of modern social psychology. In G. L. (Eds.), *Handbook of social psychology* (2nd Edition ed., Vol. 1, pp. 1-80). Reading, MA: Addison-Wesley Press.

Austin, J. S., Mateo, J. C., Hansen, J. P., & Villanueva, A. (2009). Evaluation of the Potential of Gaze Input for Game Interaction. *PsychNology, 7*(2), 213-236.

Bednarik, R., Shipilov, A., & Pietinen, S. (2011). Bidirectional Gaze in Remote Computer Mediated Collaboration: Setup and Initial Results from Pair-Programming. *Computer Supported Cooperative Work.*

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106*(3).

Cherubini, M., Nüssli, M.-A., & Dillenbourg, P. (2008). Deixis and Coupling of Partners' Eye Movements in Collaborative Work at Distance. *Eye Tracking Research & Application Symposium.* Georgia.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010, August). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42*(3), 692-700.

Deniz, O., Fal, M., Bozkurt, U., & Acartürk, C. (2015). GET-Social: Group Eye Tracking environment for social gaze analysis. *XVIII. European Conference on Eye Movements.* Vienna: Journal of Eye Movement Research.

Djamasbi, S., & Mortazavi, S. (2015). Generation Y, Baby Boomers, and Gaze Interaction Experience in Gaming. *48th Hawaii International Conference on System Sciences (HICSS).* Computer Society Press.

Donovan, J. O., Ward, J., Hodgins, S., & Sundstedt, V. (2009). Rabbit Run: Gaze and Voice Based Game Interaction. *The 9th Irish Eurographics Workshop.* Dublin.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics, 4*(6), 631-645.

Holmqvist, K. N. (2011). *Eye Tracking: a comprehensive guide to methods and measures.* Oxford, UK: Oxford University Press.

Isokoski , P., Joos, M., Spakov, O., & Martin, B. (2009). Gaze controlled games. *Universal Access in the Information Society, 8*(323).

Isokoski, P., & Martin, B. (2006). Eye tracker input in first person shooter games. *Communication by Gaze Interaction – COGAIN.* Turin.

Istance , H., Vickers , S., & Hyrskykari , A. (2009). Gaze-Based Interaction with Massively Multiplayer On-line Games. *Conference on Human Factors in Computing Systems.* ACM.

Jarick, M., & Kingstone, A. (2015, September 23). The duality of gaze: eyes extract and signal social information during sustained cooperative and competitive dyadic gaze. *Frontiers in Psychology*.

Johnson, D. H., & Sinanovic, S. (2000). Symmetrizing the Kullback-Leibler Distance. *IEEE Transactions on Information Theory*.

Kleinke, C. L. (1986). Gaze and Eye Contact: A Research Review. *Psychological Bulletin, 100*(1), 78-100.

Kleinke, C. L., & Pohlen, P. D. (1971). Affective and Emotional Responses as a Function of Other Person's Gaze and Cooperativeness in a Two-Person Game. *Journal of Personality and Social Psychology, 17*(3), 308-313.

Levandowsky, M., & Winter, D. (1971). Distance between Sets. *Nature, 234*(5), 547-579.

Liversedge, S., Gilchrist, I., & Everling, S. (n.d.). *Oxford Handbook of Eye Movements.* Oxford Library of Psychology.

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2- D images. *Spatial Vision, 9*, 363–386.

Meijering, B., van Rijn, H., Taatgen, N. A., & Verbrugge, R. (2012). What Eye Movements Can Tell about Theory of Mind in a Strategic Game. *7*(9).

Menon, T., & Phillips, K. W. (2011). Getting Even or Being at Odds? Cohesion in Even and Odd-Sized Small Groups. *Organization Science, 22*(3), 738 - 753.

Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., & Theraulaz, G. (2010, April 7). The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLOS ONE*.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443–453.

Nystörm, M., Novak, M., & Holmqvist, K. (2004). A novel approach to image coding using off-line foveation controlled by multiple eye-tracking measurement. *Picture Coding Symposium.* PCS - Tektronix, Inc.

Olsen, J. K., Ringenberg, M., Aleven, V., & Rummel, N. (2015). Dual Eye Tracking as a Tool to Assess Collaboration. *17th International Conference on Artificial Intelligence in Education, 3.*

Patrick, J., Nüssli, M.-A., & Weifeng, L. (2010). Using dual eye-tracking to unveil coordination and expertise in collaborative Tetris. *24th ACM BCS Conference on Human Computer Interaction.*

Pfeiffer, U., Schilbach , L., Jording, M., Timmermans, B., Bente, G., & Vogeley, K. (2012, December). Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in Psychology, 3*(537).

Rajashekar, U., Cormack, L. K., & Bovik, A. C. (2004). Point-of-gaze analysis reveals visual search strategies. *Human Vision and Electronic Imaging.*

Serano, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*(11), 489-493.

Soegaard, M., & Friis, R. (2013). *Encyclopedia of Human Computer Interaction* (2nd Edition ed.). The Interaction Design Foundation,.

Starker, I., & Bold, R. A. (1990). A Gaze-Responsive Self-Disclosing Display. *Conference on Human Factors.* ACM.

Swalwell , M. (2006). Multi-Player Computer Gaming: 'Better than playing (PC Games) with yourself'. *Reconstruction: Studies in Contemporary Culture*.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1972). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1*(2), 149-178.

Tatler, B., Baddeley, R., & Gilchrist, I. (2005, March). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, 643-659.

Wand, M. P., & Jones, M. C. (1995). *Kernel Smoothing* (First Edition ed.). New York, US: Chapman & Hall.

Wilcox, T., Evans, M., Pearce, C., & Sundstedt, V. (2008). Gaze and voice based game interaction: The revenge of the killer penguins. *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH.*

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

Z. I. Botev, J. F. (2010). Kernel density estimation via diffusion. *Annals of Statistics, 38*(5), 2916-2957.

Zhang , X., & MacKenzie, I. S. (2007). Evaluating Eye Tracking with ISO 9241 - Part 9. *International conference on Human-computer interaction: intelligent multimodal interaction environments* (pp. 779-788). Springer-Verlag.

Zheng, B., Hajari, N., & Atkins, M. S. (2016). Revealing team cognition from dual eye-tracking in the surgical setting. *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (pp. 321-322). New York: ACM.